Understanding the Values in the Context of Wi-Fi Access Data

Graduation Thesis

Xueyao Wang

Master Thesis | Graduation Project

May, 2019 - September, 2019

Xueyao Wang

4735587
xueyaowang.com
schultzwang@gmail.com
+31 619 127 261

Supervisory team

Chair
Jacky Bourgeois
Department Design Engineering
Section Internet of Things
Industrial Design Engineering
Delft University of Technology

Mentor
Kars Alfrink
Department Design Engineering
Section Internet of Things
Industrial Design Engineering
Delft University of Technology

Delft University of Technology
Faculty of Industrial Design Engineering
Landbergstraat 15
2628 CE Delft
The Netherlands
www.tudelft.nl

Acknowledgements

This graduation project marks an end to my 2-year study experience in TU Delft. The past two years was such an adventure full of learning and other delightful surprises. Here I've got a completely different mindset of design thinking, I've seen more possibilities design could bring to society, and I've positioned myself with a clear goal to achieve. The experience redefines what design is to me, and it will definitely influence my future career and life.

Many thanks to my graduation project chair Jacky Bourgeois and mentor Kars Alfrink, you helped me finish this project in the field where I complete had no knowledge before. Your professional knowledge guide me through most of the struggling parts in this project, and your input always pointed out a very clear direction to continue with. I've always got new inspirations after each discussion, and the feedback helped me greatly improve my research (when I had absolutely no idea of how to perform a research project).

And great thanks to all the experts and students who participated in my interviews, questionnaires and generative sessions. Without your time and insights I could not accomplish any part of this project. Thanks to all the experts, you all showed great patience to support me, sharing your experience and knowledge to help someone like me who was totally innocent in this field. Thanks to all the students who gave me feedback and inspiration, each of your input pushed my concept a step forward.

Also big thanks to all the fellow students who encouraged me with my project. All of you were busy with your own graduation projects, but you still spent time listening to my confusions and struggles when I was stuck. And special thanks to the Chinese students at TU Delft , belonging to this small community made me feel like home when I'm in the Netherlands. The warmth and happiness (and dinners) you guys brought to me were the best thing I ever had in the past two years.

There are still some regrets in this project, I hope it could give a lesson to myself. My academic career finally comes to its end, but there are still so many things I shall learn and improve as a designer. I'm looking forward to the future, with all my memories here in TU Delft.

Contents

1 Introduction

2 Literature review on design theories

3 Literature review on Wi-Fi access data and GDPR

4 Context and stakeholder research

5 Application

6 Conclusion

References

Appendix (see seperate file)

Abstract

For data-intensive research, Wi-Fi access data has become an important source since it’s low-cost, convenient and rich in information. It is mainly used to study people’s movement and then used to infer individuals’ or a group of people’s behavior pattern. The application of the research result is also common. It could be used to do interior positioning, crowd management, facility management and more.

As the Global Data Privacy Regulation (GDPR) became enforceable in 2018, it set strict rules for personal data collection and usage. Since Wi-Fi access data contains personal information and it’s hard to be anonymised for research purposes, GDPR brings more limitations for people like researchers when they deal with Wi-Fi access data.

The GDPR triggers the development of a data platform in TU Delft, which aims at sharing data to researchers in a legal way. This data platform will legitimate the management of such personal data, helping the researchers to get and process the data in a GDPR-compliant way. Three stakeholders are involved in this context: the ICT Department as the data controller, researchers and designers as the data requester, and the students and employees in the TU campus as the data subject. All their values need to be taken seriously when developing the platform. Thus, the aim of the research is to bridge the gap between the stakeholder’s values and the implementations of the platform developers.

Literature review is conducted first, discussing why ethics are important in platform design. Then the background information about Wi-Fi access data and GDPR are researched, so as an overview of the data platform. Then the research uses Friedman’s Value Sensitive Design methods to explore the values of different stakeholder entities. Methods like interview, questionnaire and generative session are used to collect the stakeholders’ insights. Then the insights are analysed, being defined as different values according to Schwartz’s theory of Basic Human Values.

Then, a toolkit is designed through two minimum viable product (MVP) iterations. It communicates the values to the platform developers, then inspire them to come up with functionalities that fulfill the values. The form of the toolkit are two sets of cards. The first card set are ‘values cards’. On the frontside of each card, a defined value, a sub-value and one quote from the stakeholder that help to understand the value are shown. On the backside of each card, questions that help the developers to measure whether this value is met are asked. The second card set are ‘inspiration cards’. It shows one or more functionality that could be developed to fulfill each value, triggering the developers coming up with more tangible features. It is expected that the toolkit will be used in the group discussions when deciding what functionalities will be implemented into the platform, to speak for the absent stakeholders as well as to trigger discussion.

Finally, the limitations of the research are discussed, the whole project is concluded and personal reflection is made.

Chapter

1

Introduction

This chapter gives a brief introduction to the project. It introduces the context and the stakeholders involved, defines the research problem, and sets the goal of the research project.

For researchers in many fields, Wi-Fi access data is a rich source to conduct a variety of research. The Wi-Fi access data contains information that could be used to identify a unique device, and how far this device is from a certain Wi-Fi access point (AP), and when is it when the device is at such location. Then the device's location in a given period could be inferred. This means that the data helps the researchers to recognize a device's movement routine, which could be further researched as an individual or a group of people's behavior patterns. The research of such data brings positive impacts in many fields, for example, crowd management, building design, facility usage and much more.

Though this data is derived from public space, it contains personal information and may concern privacy issues. To handle this data, the researchers must comply with data management rules. In 2018, the Global Data Privacy Regulation (GDPR) became enforceable. The introduction of the GDPR aims to give all the EU citizens more control over their personal information in this digital age. The GDPR set strict rules for dealing with personal data, meaning that for researchers and other people who are dealing with personal data, there's more limitations for collecting, storing and processing such data.

Apart from the laws, for ethical considerations, the people from whom the data is collected should have the right to have more control over their own data. This means that if they want, they should have the right to know where their data flows to, what it is being used for, if it is well protected, or even more.

Under such circumstance, there is an expectation that the researchers could handle the personal data in both legal and ethical manner. Being legal means being GDPR compliant, while being ethical means that give the original owners of the data control over their data, considering their needs and expectations even if it's not a must according to GDPR.

For researchers, to support them to acquire data in a more standardized procedure, the Information and Communications Technology (ICT) Department of TU Delft is envisioning a data platform which stores GDPR-compliant data only for research. This platform is expected to be designed not only in a way of legitimating research data usage, but also in an ethical way that meeting all the stakeholders' values.

This research project deals with the personal data that is collected via Wi-Fi access points in the TU Delft campus. The aim of this research is to explore the ethical aspects during the collecting and processing of Wi-Fi access data.
It explores the insights of different stakeholders around the data and defines their values. The final outcome of this research will hopefully be a support at the starting stage to the development of the data platform. It is expected that the final design concept will help the developers to incorporate different stakeholder's values into the future data platform, to achieve a fully value-based platform.

1.1 Context

In the TU Delft campus, the ICT Department provides wireless local area networking (WLAN) for students, employees and visitors. When a mobile device and its owners approach to different Wi-Fi access points (AP), the session, the strength of the signal (the distance between the device and the AP) and other data is collected at the same time. This data could be used for Internet maintenance, for example, to put more APs at the location where there are usually more people. For this purpose, data from those resources could be collected without informed consent. That means that when the students connect their devices to the campus Wi-Fi, they do not need to agree to any terms or conditions that the university is collecting their data.

This kind of data is usually referred to as 'opportunistic sensor data'. For researchers, these 'opportunistic' data could be used to analyse people's location and movement patterns across the campus. Also this data can be combined with other non-sensor personal data to be further studied.

When the Global Data Privacy Regulation (GDPR) was introduced in 2018, there are more limitations for ICT department and researchers to handle this data, since the data contains personal information. The regulations do not only influence them in the legal aspect, but also raise ethical concerns that the providers of those data shall also have more knowledge and control over their own data. At the moment, the researchers apply for those data directly from the ICT Department, without any standardized procedures. To legitimate the data management, also to support both researchers and ICT Department itself, the ICT is envisioning a data platform that stores all kinds of research data. The researchers can then obtain data directly from this platform, in a GDPR-compliant manner.

1.2 Stakeholders

There are three groups of stakeholders involved in a personal data usage scenario: ***the data controller, the data requester and the data subject***.

The scope of the research is in the TU Delft campus, and the personal data is the Wi-Fi access data. So in this research project, the data controller is the ICT Department, who collects, manages the data and keep the data usage in compliance with GDPR. The data requesters are the researchers and designers in the university, who request the data to complete their working goals. The data subjects are students and employees who use the campus Wi-Fi, their data is collected from Wi-Fi APs and will later be used by the other two stakeholders. The three stakeholder entities have different needs towards the data, and there might be potential conflicts between their needs.

1.3 Research problem

As the GDPR carries out, there is not only the need to legitimate the data usage, but also the ethical issues to construct such a data platform. It is not enough to develop a platform that only serves to comply with laws and regulations. In the meanwhile, the development of the envisioning data platform should take the stakeholders' hopes and concerns into consideration, to fulfill the practical and ethical demands. So the researchers, the ICT Department, the original owners of the data and many other stakeholders' insights are vital for the developing process.

In this multi-stakeholder context, there is a potential conflict between the three stakeholder entities: the data controller and requester want to collect and process the data as much as possible for operation and research use, while the data subject are reluctant to share their personal data, because they want to protect their privacy well.

To improve data management and leverage data usage, the ICT department is envisioning a data platform that manages the data in a better way. It is expected to store and manage a variety of data in a GDPR-compliant manner, so that the researchers can request data from the platform conveniently and lawfully.

The platform should aim at not only keeping data management in compliance with GDPR, but also balancing all the stakeholders expectations, which means to incorporate their values into the platform design. Thus, before the functionalities and features of this platform is decided, its (potential) users' needs and concerns shall be explored. As a designer, it's important to understand the users, and convey the results of the user research to the developers of the platform.

Thus, the topic of this research project is clear, that is Understanding the Values in the Context of Wi-Fi Access Data. Based on this topic, a problem statement is formed:

"What are the expectations and concerns of the stakeholders, during the Wi-Fi collecting and usage process?"

This problem statement leads a direction for the research, and triggers a research problem:

"How to help the developers incorporate values into the design of the data platform, that balance the expectations of different stakeholders?"

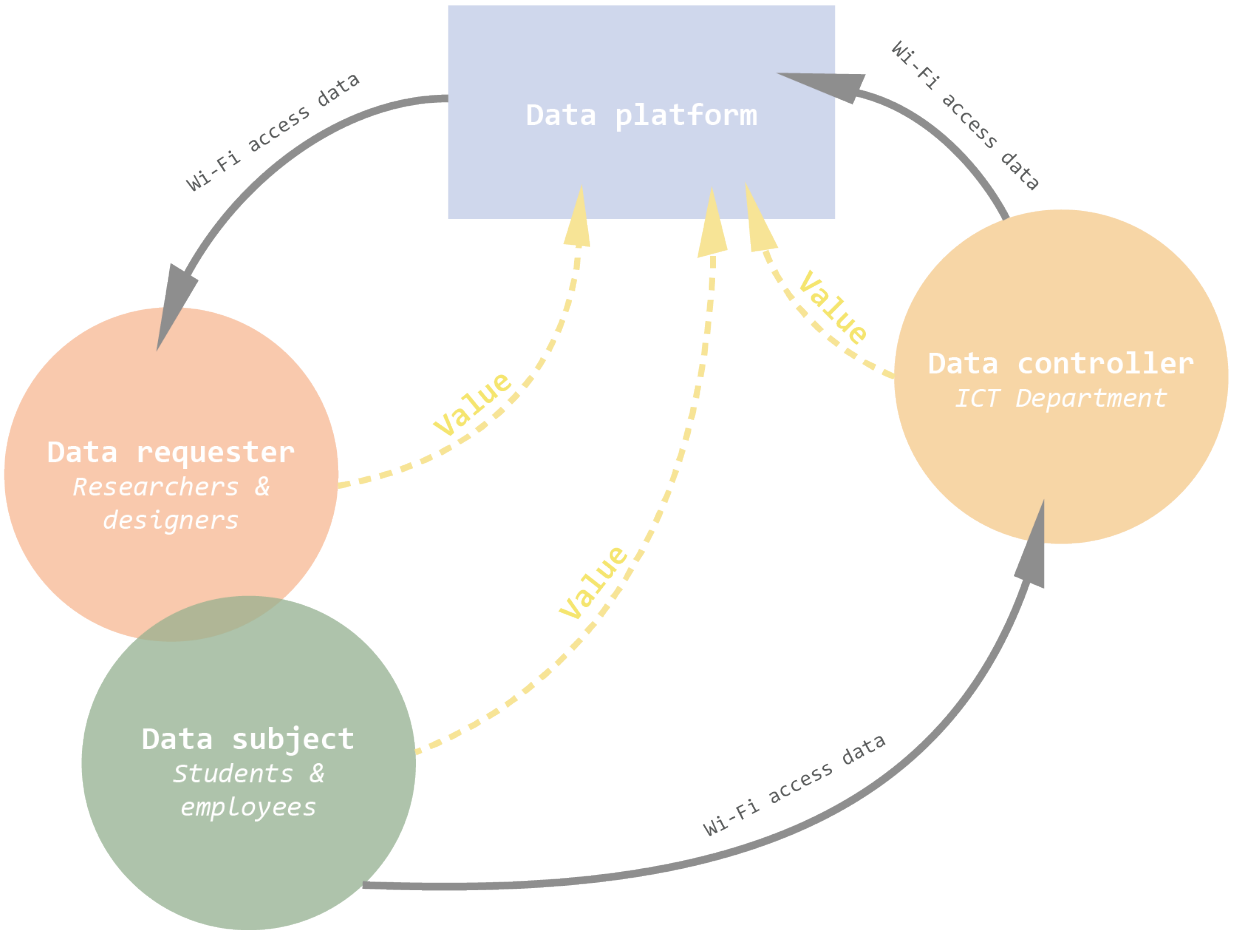
The figure 1-1 shows a simplified relationship of the three stakeholder entities and the platform. The pink circle on the left is the data requester, and the green circle below represents data subject. These two groups are overlapped, since the data subject could also participate in the research, and the data requesters could also use the campus Wi-Fi. The yellow circle on the right represents data controller. The Wi-Fi access data flow from the data subject to the data controller, later it will be transported into the data platform. Then the data requesters can request data from the data platform. The values of all the stakeholders will be considered and incorporated into the data platform.

1.4 Research goal

To answer the research question above, there are lots of questions to think about before the development of this data platform. For the data controllers, what kind of data are they going to put into this platform? What do they want to get for themselves from this data platform? For data requesters, what kind of data do they need for research?

Figure 1-1: A simplified context map with three stakeholder entities and the data platform

▼



Do they understand the personal data usage well or not? For data subjects, are they aware that their data is being used for research? Are they willing to share their data or not?

The understanding of such questions will help the developers to have an overview of what the platform should be like, what kind of functionality should be there, and how the users use the platform. But how to understand different needs, and how to transform those needs into tangible features and embed them into the platform?

Different stakeholders may have completely different values, and their values might potentially conflicting with one another. Even people from the same stakeholder entity could have totally different ideas. The input from all the stakeholders are massive and complicated, there should be a systematic way of introducing their insights to the developers. Also, for the development of the platform, there is still a gap between “what the stakeholders need” and “how to achieve the stakeholders’ needs”. This means that the developers need more tangible requirements, to realise technical functionalities. The figure 1-2 shows the gap between understanding the stakeholders and incorporating their needs into the platform.

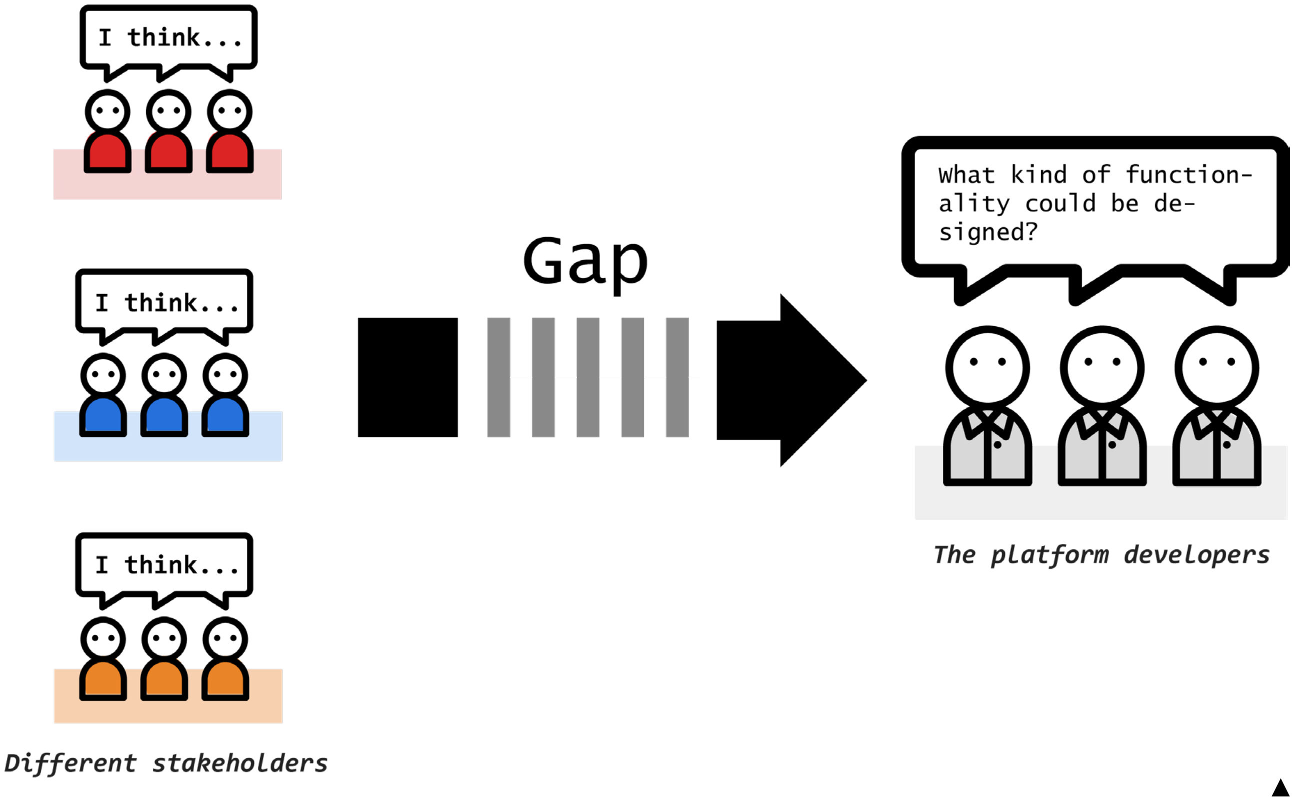


Figure 1-2: Current gap between the stakeholders and developers

Thus, this project will serve as a communicator, bridging the gap between the stakeholder's needs and how to achieve those needs, transforming the stakeholders' insights into inspirations that helps the developers to come up with solid solutions. The figure 1-3 shows how the research project works as a communicator between the two sides.

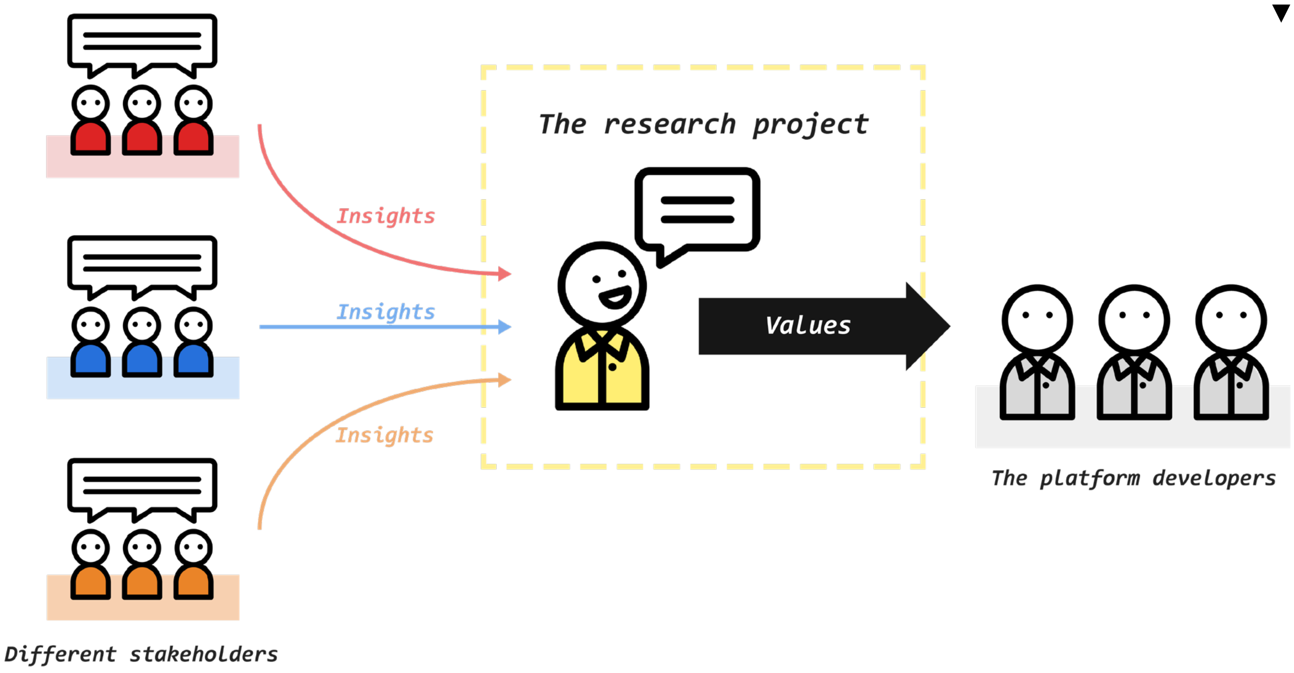
The goal of this research should help solve the research problem, that is to incorporate values into the platform. To achieve this, the result of this research should be:

Bringing communication and inspiration into the design and development of a value-based data platform in a multi-stakeholder context.

So, the final deliverable of the project will be a toolkit, that presents the values of the stakeholders, and serves as an inspiration for the platform design.

Figure 1-3: The aim of the research project is to serve as a communicator

▼



Chapter

2

Literature review on design theories

This chapter reviews literature on related design theories. Theories of design ethics and the Value Sensitive Design method will be introduced. There are also two case studies about personal data management in research.

2.1 Design ethics

In terms of designing an object, a designer’s job could be to select the design parameters of an artifact so that the artifact ‘works’ is makeable and, if possible, pleases the user as well (Baldwin et al., 2000). This suggests that the designer’s aim is not only making an object functions, but also taking ethical considerations, in this case, bring happiness to the users. Since there is always a vision that the design serves for the well-being of its users/audiences, we could say, there is an ethical nature in all kinds of design.

Thinking about the design of a data platform which deals with personal data, ethics is a fundamental thing to consider. But what does ethics to do with a data platform, and should designers take responsibilities of ethics during the developing process? Below a few theories about ethics, design, and technology are introduced, explaining why technology design is inherently ethical, and why designers should take ethical into consideration during the design process.

Technology shapes behavior

When designers envisioning a technical object, they think about the context of usage, they think about who are those direct or indirect users and they think about how they use the object. They thus define actors with specific tastes, competences, motives, aspirations, political prejudices, and the rest, and they assume that morality, technology, science, and economy will evolve in

particular ways (Akrich, 1992). This means that they are envisioning a world that their designing object will be existing, and Akrich (1992) and Latour (1992) defines the end product of the work a “script”, or in a more familiar and understandable word, a “scenario”.

Just like a film script, this technical object (also, its designers) gives its surrounding actors (its projected users) prescription on how to act. This way, new technologies may not only lead to new arrangements of people and things, they may in addition generate and “naturalized” new forms and orders of causality and new forms of knowledge about the world (Akrich, 1992). In this sense, many social behavior could be shaped by the technical objects, and ethical issues may emerge from those designs.

A famous example of how a designed object shapes social behavior is Robert Moses’ bridge design. He was known as the “Master Builder” of New York, had shaped much of New York’s infrastructure, including a number of “low-hanging overpasses” on the Long Island parkways that led to Jones Beach (Winner, 1980). These overpasses were designed lower, that the buses could not pass beneath them. This way, colored people and poor people who could only afford public transportation could not access Jones Beach. Although there has been some dispute about whether this consequence was actually intended in the design (Joerges 1999), in this case, it is proved that technologies can be used in ways that enhance the power, authority, and privilege of some over others (Winner, 1980). When object was influencing the social pattern, itself becomes racial as well. This example clearly reveals ethical issues that are not emerged from the object itself, but the actors around it and the context the designer deliberately created.

◄ ***Figure 2-1: Palmer Avenue Bridge, Bronx River Parkway, 1927 (detail). Photo: Historic American Engineering Record***

Technology as a mediation role

Studied from Akrich and Latour's Script Concept, Peter-Paul Verbeek (2006) argues that comparing to common sense that the ethics concerning technical objects is only relates to the quality and functionality of the objects, it should also be related to the role they play in their usage scenario. He brings up the concept of technological mediation, which concerns the role of technology in human action and human experience.

According to Heidegger (1927), tools could be considered as a connection between human and the real-world. That is to say, when using a technical tool, human's behavior in the context is also shaped by the tool. So this tool is a mediator between its user and the real world. Technological artifacts are not neutral intermediaries but actively co-shape people's being in the world: their perceptions and actions, experience, and existence (Verbeek, 2006).

Don Ihde (1990) mentions that technologies provide a representation of reality, which requires interpretation. He gives an example of looking at a tree with an infrared camera. One can perceive from it, seeing whether the tree is healthy or not. In this sense, technology is no longer neutral, it has intentions and shapes people's perception in the real world.

Apart from perception, technologies also mediate people's actions. For example, a traffic sign makes people slow down because of what it signifies, not because of its material presence in the relation between humans and the world. (Verbeek, 2006).

The technical mediation shows that technology has ethical issues to do with itself, and to designers, design such technical object should also take ethics into consideration. Verbeek (2006) call this technology design "materialize morality". He suggests that designers should think about the mediation role of technologies, how it will eventually play in society could be integrated in the design process. Designers can not only help to shape the role of the technology itself, but also the context it's going to perform and the interaction between it and its users.

The figure below shows Verbeek's sources of mediation. It shows that designers and technology itself are not the only source that decide the mediation; the users can also interpret the technology on their own. Then the mediation will lead to user's perception or action.

Figure 2-2: Verbeek's sources of mediation (2006) ▸

Designers' moral responsibility

To summarize from above, technology shapes the behavior of its projected users, and it influences their perception and action. Technology itself is not a neutral object, and the designers should take moral responsibility during the design process, incorporating ethics into the technology design.

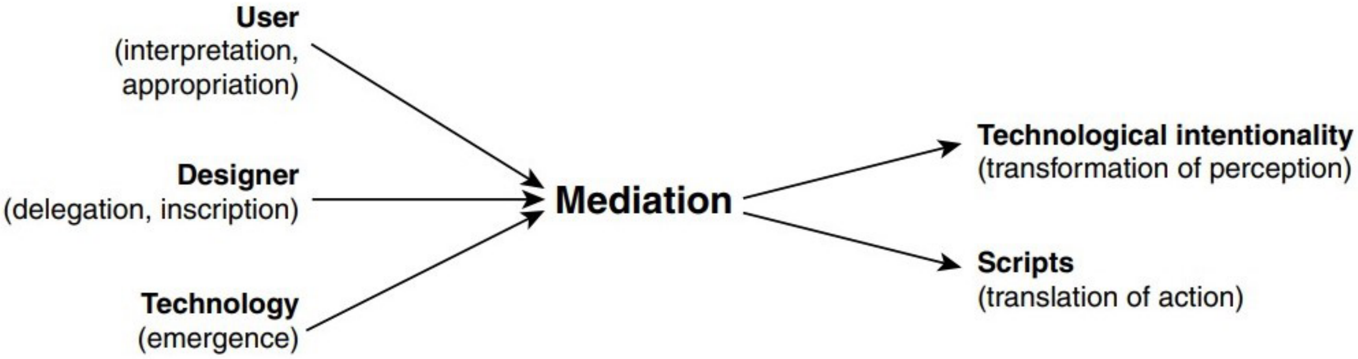
During the design process, it is not enough to just look at what humans intend and do, "Ascribing more responsibility to persons who act with technology requires coming to grips with the behavior of the technology (Johnson et al., 2005)." Moral responsibility is, thus, not only about how the actions of a person or a group of people affect others in a morally significant way; it is also about how their actions are shaped by technology. Moral responsibility from this perspective is not located in an individual or an interpersonal relationship, but is distributed among humans and technologies (Noorman, 2018).

So the designer's moral responsibility shall not be limited with the quality and the usability of the object alone, they shall take a step further, looking at the whole context where the designed object influence people's perception and action, and inspect their interpretation of their responsibility in the light of their view of the good life (Burg et al., 1970).

Ethical Design Model

From the above literature review, we already have the conclusion that ethics should be implemented in design. But what exactly are these ethics? According to Merriam Webster, ethics is "the discipline dealing with what is good and bad and with moral duty and obligation." From a designer's perspective, ethical design should have good impact to the users as well as the society.

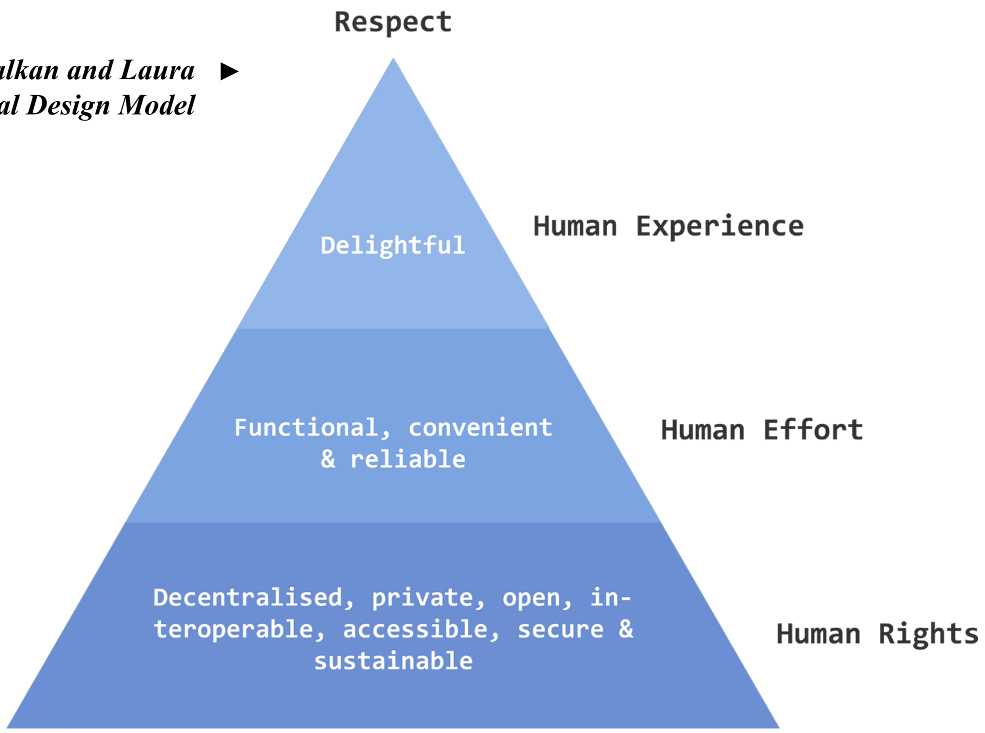
We here use Aral Balkan and Laura Kalbag's Ethical Design Model to see



what could ethical design bring to the actors around it, to be more specifically in this case, what will be brought to the stakeholders of the Wi-Fi access data in the TU camus. The model is formed in a pyramid hierarchy, and there are three layers in the ethical hierarchy of needs: human rights, human effort and human experience. The human rights lays on the bottom, meaning that the basis of an ethical design is to respect civic rights. It will benefit its user/audience as a person. In this case, it means that the data platform should protect data subject's privacy and keep the data secure. The layer in the middle is human effort. It means that the design is functional, convenient, and reliable. In this case, it means that the data platform provides the data requesters a convenient and reliable way of accessing data, and it also helps the data controller to manage data in an easier way. The layer on top is human experience, it means the product makes its users/audience smile. In this case, it means that the data platform will balance all the stakeholder's values, making their lives better.

As long as the developer implement GDPR to the platform, it is predictable that some the bottom layer of Human Rights will be fulfilled. For example, privacy, security will be assured by the regulations, and accessibility of the data is the fundamental functionality of the platform. Then, the middle layer Human Effort needs more its direct user's input, namely the data requesters. Finally, to delight all the stakeholders lives need more effort. To understand the stakeholders' needs will help to create a technically functional platform for them to use, and eventually put delightful impact on the stakeholders' experience.

Figure 2-3: Aral Balkan and Laura Kalbag's Ethical Design Model ▶


Respect
Delightful
Human Experience
Functional, convenient & reliable
Human Effort
Decentralised, private, open, interoperable, accessible, secure & sustainable
Human Rights

Summary of Chapter 2.1

Design is an ethical process that leads to better lives of the human beings. While new technology emerges, it does not only serves its users with its functionality; the impact of it could be broader and deeper. It transmits knowledge and other information to people, and shapes the society's behavior. The ethics within the technology design could have a huge impact in the society.

Technology is also a mediator between the users and reality. It shapes people's perception as well as action. It's not neutral, so the designers should take ethical considerations when designing, thinking beyond the functionalities and qualities, having a greater picture of how this technology could do good to the whole context. It's designers' moral responsibilities to achieve the goal of leading the users to better lives, by delivering ethical technology designs to the society.

Finally, an Ethical Design Model is introduced, help the designers to achieve ethical design in three levels: Human Rights, Human Effort and Human Experience. An ethical design shall first fulfill basic human rights like privacy, security and sustainability, then saves human effort by delivering functional, convenient and reliable objects, and finally improves human experience by bringing delightfulness to the users.

2.2 Value Sensitive Design

What is Value Sensitive Design?

Thinking about bringing ethics into design, one of the most famous theories is Value Sensitive Design. Value Sensitive Design is a theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process.' (Friedman et al., 2002) It highlights the way in which technology both shapes society and is shaped by social factors (Friedman et al., 2002). Such complex socio-technical systems involve intertwined interactions between humans and technology and cannot be designed in a value vacuum (Gispen, 2017). Since the research project aims at exploring different stakeholders' needs and incorporate moral considerations into a technology platform design, Value Sensitive Design is a suitable approach to target the stakeholders' moral values and help to integrate them in the later platform design.

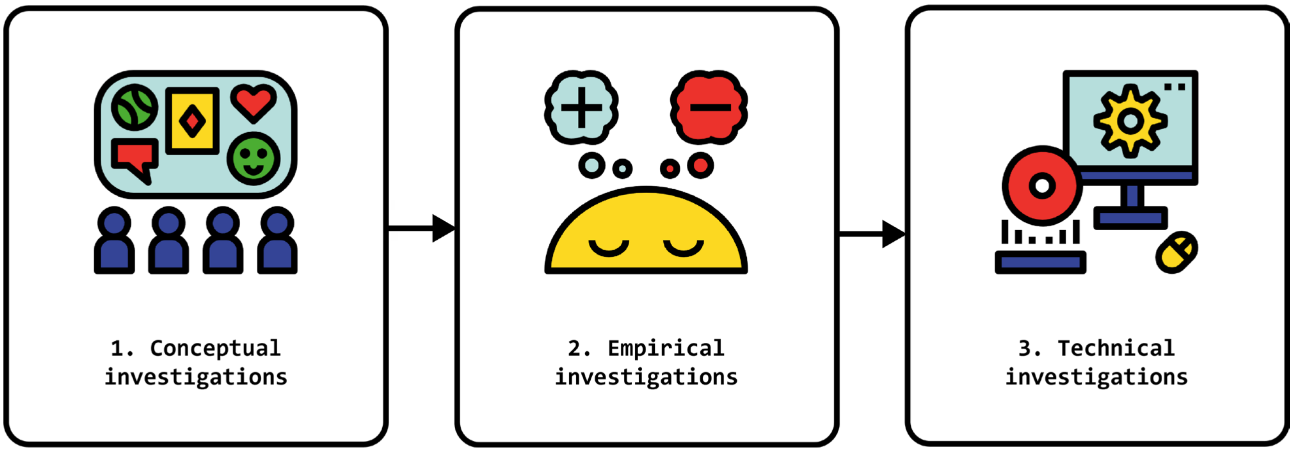
Before stepping into this approach, the first question needs to be answered is: what is value?

Friedman (2002) broadly defines the term wherein ***a value refers to***

what a person or group of people consider important in life. So both a person's hopes and concerns can be traced back to his values. The values are closely connected to ethics. Sometimes ethics has been subsumed within a theory of values, and other times conversely, with ethical values viewed as just one component of ethics more generally (Friedman, 2009). Ethics are a set of prescriptions, while values are tied to action (Knobel et al., 2011). A key difference between ethics and values is that ethics is a uniform that is widely accepted in one region or is universal, while values differ from person to person (Oluwatosin, 2017). In this project, a stakeholder's values will be collected and analysed, and the process that the designers and developers try to fulfill these values is ethical, or say, this process is incorporating ethics into the platform design.

Three steps in iteration

The goal of Value Sensitive Design is to influence the design of technology by explicitly attending to human values and integrating them into and throughout the design process (Friedman and Kahn, 1997). It addresses design issues within the fields of information systems design (Friedman et al., 2013). When applying Value Sensitive Design in a design practice, there is a three-step iterative methodology: conceptual, empirical, and technical investigations. Conceptual investigations raises questions of fundamental issues within a project, for example, what values are implicated and is there any value more important than others? During this step, values are somehow defined. Empirical investigations are then carried out to analyse the stakeholders' attitudes towards those values. Social science research method


1. Conceptual investigations
2. Empirical investigations
3. Technical investigations

▲

Figure 2-4: The three-step methodology in VSD

could be used in this phase, exploring how stakeholders consider these values and if they find one more important than others. Finally, technical investigations focus on how existing technological properties and underlying mechanisms support or hinder human values (Friedman, 2009).

Value Sensitive Design methods

The core concern of value sensitive design is to address human values in the technical design process (Friedman, 2017). According to van den Hoven (2015), there are four key claims of Value Sensitive Design: values can be expressed and embedded in technology, technologies have real and sometimes non-obvious impacts on those who are directly and indirectly affected, explicit thinking about the values that are imparted in technical design is morally significant, and value considerations should be surfaced early in the technical design process. These elements makes it possible to use Value Sensitive Design from the beginning of the research, from defining the values to incorporating them. Friedman (2017) also provides a variety of methods used to investigation of values in technology, serving such purposes as stakeholder identification and legitimation, value representation and elicitation, and values analysis. Below a few methods will be introduced.

1. Direct and Indirect Stakeholder Analysis

This method encourages not only to explore direct stakeholders of the technology, but also consider indirect stakeholders in the context, whose data or presence may be implicated by the technology (Czeskis et al., 2010). An example of use of such method is when Czeskis et al. (2010) defining the challenge in the case of mobile parenting technology, they do not only analyse parents and children, but also include children’s friends and friends’ parents as indirect stakeholders, comparing different attitudes towards the same technology when acting in different roles.

2. Value Scenario

Value scenario is just like a common scenario, while it emphasizes societal impact of technology and context, namely: (a) implications for direct and indirect stakeholders, (b) key values, (c) widespread use, (d) indirect impacts, (e) longer-term use, and (f) systemic effects (Friedman, 2017). In different cases, this method can be used by different stakeholders for different purposes. It can either be written down by the researchers and shown to respondents as a means to explore design space (Czeskis et al., 2010), or it could be written down by respondents to express their insights (Woelfer et al., 2011).

3. Value-oriented Semi-structured Interview

Semi-structured interviews provide a means to tap into stakeholders' understandings, views and values (Kahn, 1999; Piaget, 1929/1960). The value-oriented semi-structured interview sets questions to elicit insights about values towards technology, and the semi structure makes it possible to dig deeper with the insights.

4. Scalable Information Dimensions

This method is used when assessing the values and concerns in scalable dimensions. Interviewees can set questions with scalable impact, and ask the respondent to choose from different scales. The form of the method varies, interviews, surveys,value scenario and other method could be applied. (Friedman, 2017)

5. Value-oriented Coding Manual

After using methods like value scenario or semi-structured interview, the value-oriented coding manual could be used to analyse the result. Coding categories will be generated from the qualitative research data, and then a label, a definition and sample responses will be given to each category. (Friedman, 2017)

Summary of Chapter 2.2

Value Sensitive Design is a fundamental approach when dealing with ethics in the design of human-technology interaction. Value is what people find important in their lives. Taking stakeholders' values into consideration will ultimately help designers deliver ethical designs.

There is a three-step iteration method in Value Sensitive Design: conceptual, empirical, and technical investigations. This three-step method first raise fundamental societal questions that leads to defining the values in the context. Then stakeholders' attitudes towards those values are explored, weighing preferences of those values. And finally the study of current technology will help designers understand what kind of values this certain technology supports or hinders

And five Value Sensitive Design methods are introduced. They are: Direct and Indirect Stakeholder Analysis, Value Scenario, Value-oriented Semi-structured Interview, Scalable Information Dimensions and Value-oriented Coding Manual. They will later be used in the research of the stakeholders, help to elicit values from stakeholders' insights.

2.3 Related project about data-intensive research

There are lots of research and design using personal data to achieve different goals. For example, McCormick et al. (2017) collect data from Twitter for demographic research, Blumenstock et al. (2015) use mobile phone metadata to predict immeasurable wealth in developing countries, and Tang and Lansky (2005) try to use Personal Health Records to improve the engagement of patients in the healthcare system. Nowadays, online activities of individuals, for example on mobile phones, also allow the continuous collection of health-related and other data (Costa, 2013). Ethical concerns about privacy has always been an issue of these projects. Here are two case studies of Apple ResearchKit and Open Humans, of how they collect the data and how they protect the users' privacy.

Apple ResearchKit

Apple is stepping into the medical field, using the open source framework ResearchKit to conduct medical research. The researchers can use this framework to build their own apps, to recruit participants and collect their insights as well as their data.

The Apple ResearchKit solves one big problem for researchers: to enroll participants and gather data. It's always been a challenge for researchers to get enough data for research, but the huge amount of iPhone users makes it possible for researchers to reach suitable participants all around the world. ResearchKit does not only connect more data subjects, but also makes it convenient to collect data on a frequent basis with the built-in sensors in iPhones. For researchers, the result is that there are more data subject with a great diversity, and their data are collected more frequently with the help of iPhone apps, so in the end there are more data a more accurate representation of the population (Apple, 2019).

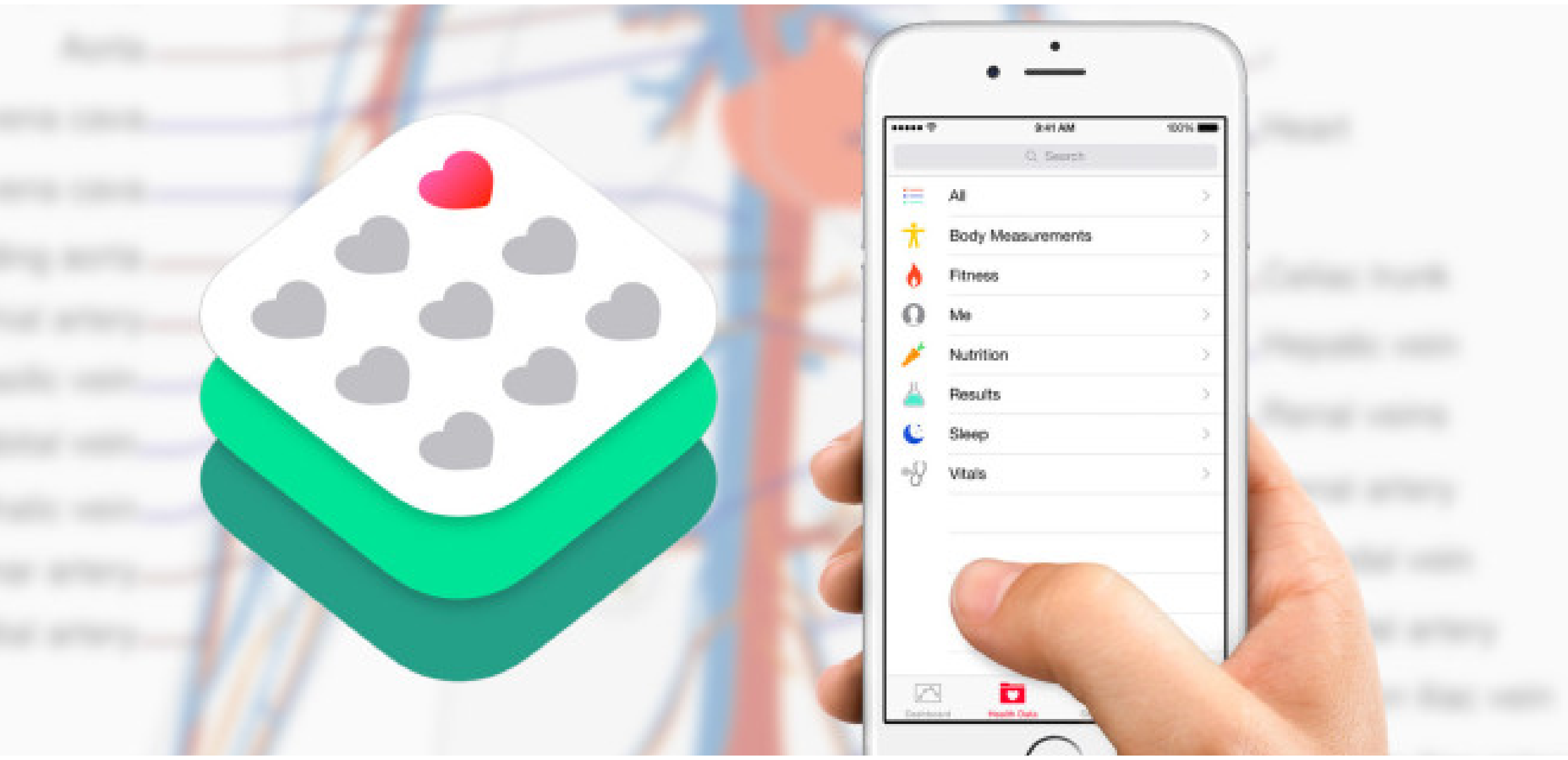
For data subjects, they can also conveniently share their data with researchers without going out to hospitals. They also have the autonomy to share their data with the research they're interested in by downloading different apps. In return, they understand their health conditions better, and the patients can manage their symptoms and medications better (Apple, 2019).

For privacy concerns, the participants can choose the research they are interested, and they can choose what kind of data they are sharing. There's informed consent in every app for data subjects to understand the risks of using their data. However, there's no more information for the participants, for they don't know about the research carried on using their data, they don't know the process or the result.

Even though Apple doesn’t collect participants’ data, the researchers and institutes who collect and use the data is completely out of data subjects’ control. Even their data is de-identified before sending to the researchers, it is possible to reconstruct someone’s identity from a relatively small amount of data, even after that data has technically been made anonymous (Ouellette, 2016). In the meanwhile, to read the informed consent is also not an easy task on iPhone. Some participants just choose ‘agree’ without reading it, meaning that they may still have no idea of the potential risks and their rights.

To summarize, the large-scale medical data collection like Apple ResearchKit favors the researchers as well as the participants. However, only provide informed consent and de-identify the information for research is still not enough for ethical considerations.

> **"We can't promise perfect anonymity. The biggest risk in these studies is to your privacy; we're going to de-identify it, but because we're going to make it available for lots of research, there exists a chance that someone could re-identify you."**
>
> **——John Wilbanks,**
> **Sage Bionetwork's**
> **Chief Commons Officer**



Open Humans

Open Humans is a community-based platform that enables its members to share a growing number of personal data types. It does not only allow data subjects participate in research projects, but also let them create their own projects, and facilitates the exploration of personal data for the individual member (Tzovaras et al., 2017).

It's designed as a web platform with the goal of easily enabling connections to existing and newly created data sources and data (re-)using applications. The goal of the platform is to enable members to import data into their accounts from various sources and use the data to explore it on their own and share it with citizen science and academic research projects alike (Tzovaras et al., 2017).

Open Humans provides a platform where collecting data and donating data become convenient. For data donators, they have a lot of power controlling their data. They know what kind of data they are sharing with a certain project, and they know whether their usernames are pseudonymised or not. The data they donate to different projects will be stored into their accounts, and they can import data from other services. If they want, they can make their data publicly available as well (Tzovaras et al., 2017).

For the users who want to collect data, they can create their own projects on this platform, then others can donate their data by joining into the project. The participants usernames will be pseudonymised by default while joining in a project, and the project has to require the data from the participants by specifying what kind of data they want to access. In addition to specifying the access permissions, projects also need to clearly signal whether they are a research study that has been approved by an Institutional Review Board (IRB) or equivalent, or whether they are a project not performing such research (Tzovaras et al., 2017).

Since its launch, Open Humans has been growing rapidly, and has supported a variety of research projects. For example, it supports The Quality of Life (QoL) Technologies Lab to collect data from multiple sources to better understand the health implications of lifestyle behaviors; it supports Imputer to help its project members to augment their existing genetic data sources; and it supports The Human-Computer Interaction for Personal Genomics (PGHCI) project at Wellesley College and New York University to do genetic data visualization research (Tzovaras et al., 2017).

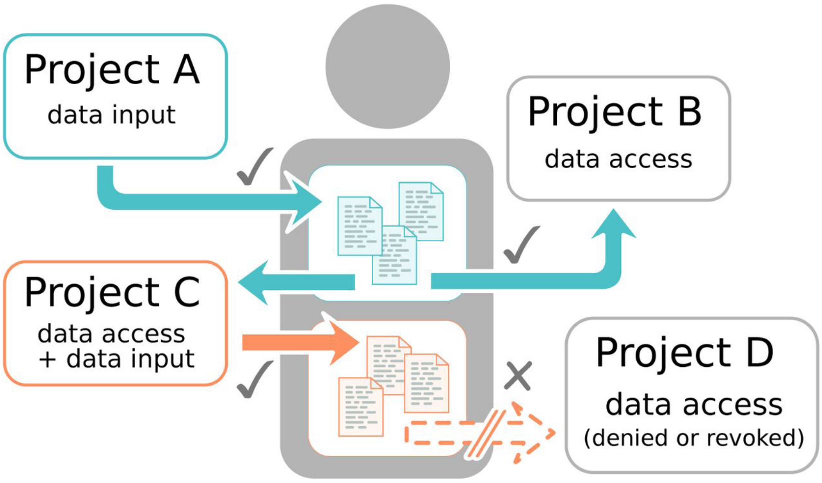
The platform gives both the data requesters and data subjects high

◄ ***Figure 2-5: Managing health data using ResearchKit app*** autonomy. On the one hand, the data requesters can choose which data subject they want to collect data and what kind of data to be collected. On the other hand, the data subject can choose to upload the data they would like to donate, and only give permission to the projects they are interested to use their data. The data requesters shall make it clear for the data subject what kind of data they want to access, for what kind of purpose they are going to use the data and what kind of security measures they will take to protect privacy.

Comparing to Apple ResearchKit, Open Humans does more to give control back to the data subjects. The data subjects in this case knows well about the project, and understand how their data is protected. Their data is used in an ethical way where they remain the owner of their data.

Figure 2-6: The Open Humans authorization flow (Tzovaras et al., 2017)

▼



Chapter

3

Literature review on Wi-Fi access data and GDPR

This chapter reviews literature on Wi-Fi access data and GDPR, which explains the background information and introduces basic knowledges to the topic.

3.1 Wi-Fi access data

3.1.1 What is Wi-Fi access data

In the TU Delft campus, the ICT Department provides wireless network in most areas of the campus, including all the buildings and squares. The students, employees and visitors can log on to the Wi-Fi with their NetID accounts to use the campus Wi-Fi for free.

While the mobile devices turn on Wi-Fi, they transmit Wi-Fi signals to be connected to a Wi-Fi access point. Meanwhile certain data is being collected. These Wi-Fi access data includes (Braggaar, 2018):

- ***Timestamp***: *the time at which the probe request arrives at the Wi-Fi base station*

- ***MAC address***: *the Media Control Address or unique identifier of the device on the network*

- ***RSSI***: *an indicator for signal strength*

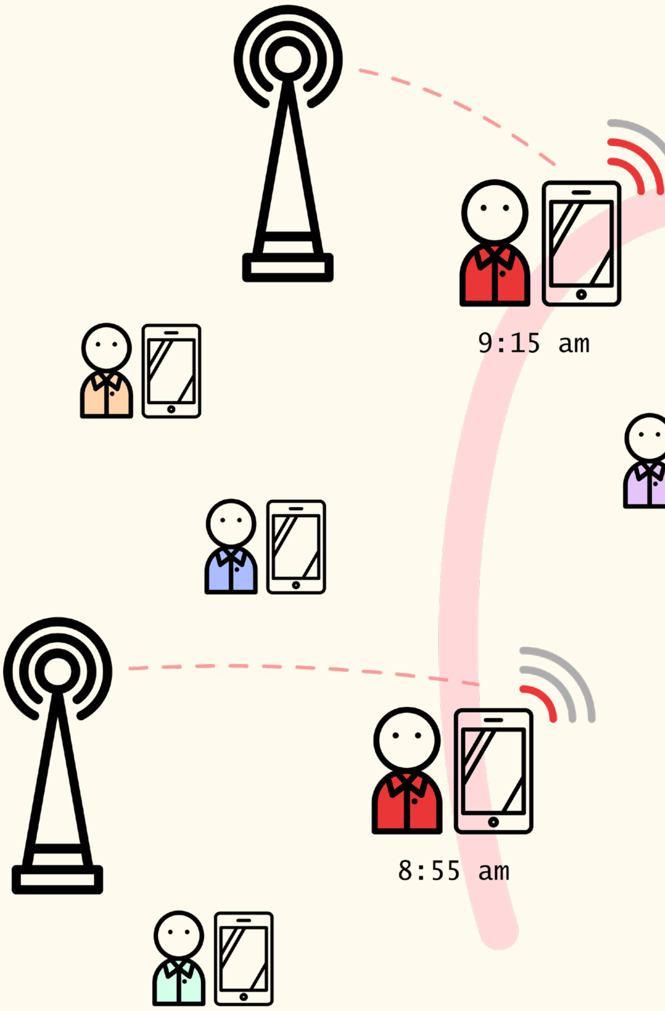
- ***SSID***: *the name of the network the client is currently connected to*

- ***Requested SSID***: *the name of the network the client station wants to connect to*

- ***Vendor***: *the manufacturer info of the device's chipset, which can be inferred from MAC address*

3.1.2 What information can be inferred from Wi-Fi access data

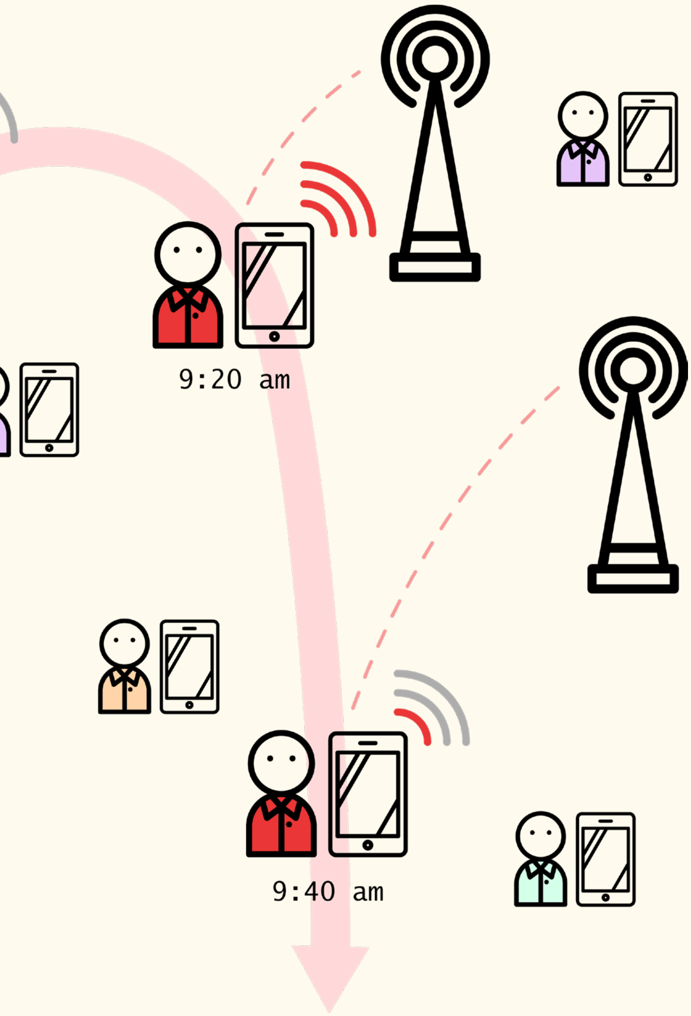
When a device connects to an access point, the distance from the device to the access point can be inferred from the device signal strength. When the device moves, it approaches other access points. By analysing the signal strength and connection session data together with the access point map, the device's movements are tracked, and people's behavior pattern can be further studied.



The figure 3-1 shows how a person's movement pattern is inferred from the Timestamp and RSSI. The people with different color represents for devices with different MAC addresses, the unique MAC address helps to identify a certain person's device. The signal strength helps to infer the distance between a device and its nearby access points. The time below the person shows when the person is at a certain location. And the pink route is the person's movement pattern.

Figure 3-1: How Wi-Fi access data locate devices

▼



3.1.3 Wi-Fi access data in research

Comparing with other ways of tracking people and studying their behavior, using Wi-Fi access data has more advantages. First, the Wi-Fi access points are already well-established in most public buildings, which means that no extra devices need to be installed to collect data. It's almost free and easy to obtain movement data. Second, users of smartphones with Wi-Fi functionality is increasing, meaning that the size of the dataset is massive. Third, in the system like eduroam, data can be collected without data subjects' consent if it's anonymised. This kind of data collection is referred to as 'opportunistic data', provides a significant advantage compared to other techniques for tracking where the subjects need to actively cooperate, either by carrying a specialized tracking device or by actively and willingly sharing information about their location (Bonné et al., 2013).

A variety of studies have been done using Wi-Fi access data to study people's behavior pattern. Bonné et al. used the real-time gathered data for crowd control in a music festival. They visualize the real-time data to see how people flows, monitoring crowd density, analyzing how long they spend at the festival and which artist has more audiences (2013). A. Danalet et al. used Wi-Fi traces to measure catering choices on campus, then forecasting the average number of visits after the opening of a new self-service (2016). Zhu et al. by studying campus and dormitory Wi-Fi user distribution, found out that students generally have regular diet only on weekdays and they have tardiness behavior in the weekday mornings (2015). Kalogianni et al. used passive Wi-Fi monitoring the occupation and movement in campus buildings to improve future use of the campus (2015).

For commercial purposes, commercial Wi-Fi deployments may be transformed into powerful tools for conducting market research and gauging insights from customers, as Wi-Fi traffic can reveal information on first time vs. frequent visitors at shops, customer loyalty, dwell times, walking paths, real-time heat-maps, customer gender and age (Redondi et al., 2018).

3.2 GDPR

3.2.1 GDPR and personal data management

The EU General Data Protection Regulation (GDPR) is the most important change in data privacy regulation in 20 years. The aim of the GDPR is to protect all EU citizens from privacy and data breaches in today's data-driven world (GDPR, 2018).

According to Dutch Data Protection Act, the MAC address and location data are personal data (DAPA, 2015). To comply with GDPR, the data controller should have legal grounds to process personal data. According to GDPR Article 6(1), one principle for processing personal data is that the person concerned consents to the processing for specific purposes. However, from the technical perspective, it is almost impossible to ask every data subject in advance for their permission to process their data. In the case of TU Delft, when the researchers want to use a group of people's Wi-Fi access data, their information is encrypted. And it's not possible to ask for anonymous people's information. If they want to have the consent beforehand, that means that every student, employee and visitor has to give permission for each of their device, which is also very difficult to accomplish.

Another principle to process personal data is that the processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority. In the context of TU Delft, research is a legal ground that fulfills this principle.

When the ICT Department of TU Delft collects data through Wi-Fi access points, they use the data to evaluate whether there are enough access points in the campus, and to maintain the Internet infrastructure. However, using the data for providing Internet service has a time limitation for storing the data. That's why a data platform is now expected. If the ICT Department builds such a platform in the future to share data for research purposes, then there is the legal ground to store the data longer and manage it more properly. It's also good for researchers to get data such as Wi-Fi access data efficiently and legally.

3.2.2 GDPR-compliant data flow

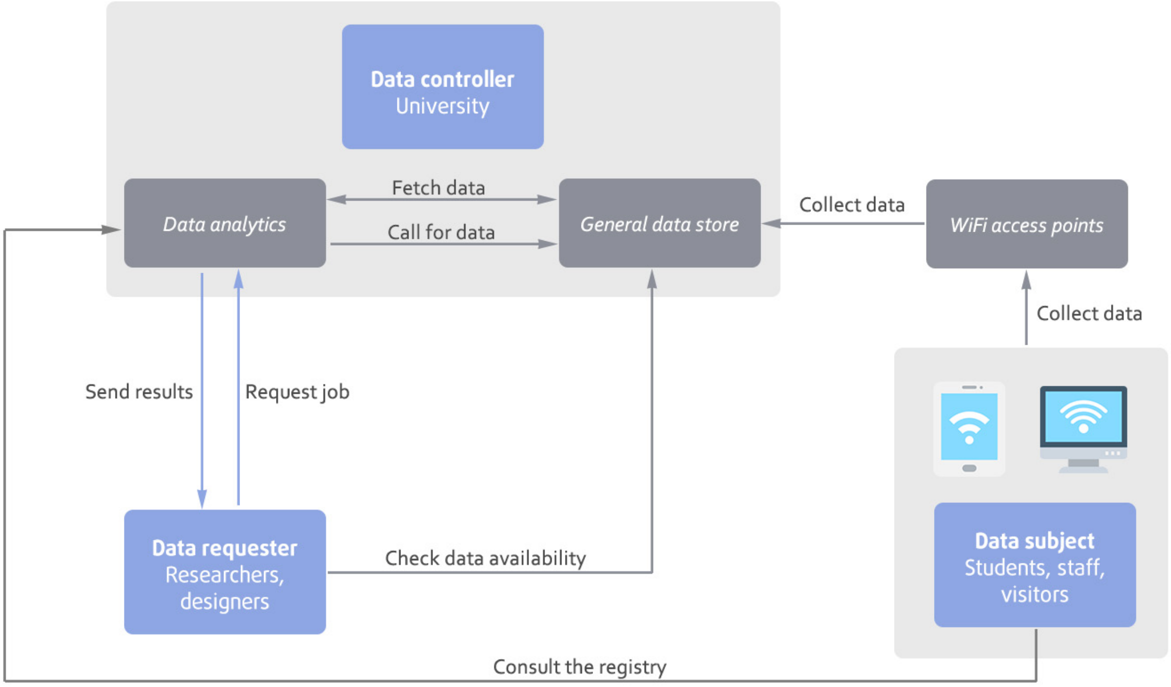
According to Bourgeois et al. (2018), the figure 3-2 shows an example of a GDPR-compliant data flow between the three stakeholder entities. On the top of the figure is the Data controller, that is the university ICT Department. There are two data boxes. The one on the right is the General Data Store, that is to temporarily store all the data collected. The one on the left is the Data Analytics Hub. It request data from the general data store, analyze it and publishes the result.

On the bottom left are the data requesters who are researchers and designers. The data requester can check if their data is available in the General Data Store first. If so, they can request data analytics job from the Data Analytics Hub, and then the Data Analytics Hub will fetch data from General Data Store according to their requirements. Once the data analytics is done, the result will be sent to data requesters.

On the bottom right are the data subjects who are students, employees and visitors. When they use their mobile device to connect the Wi-Fi in the campus, their data is collected via Wi-Fi Access Points. This data is stored in the data controller’s General Data Store.

Figure 3-2: An example of GDPR-compliant data flow (Bourgeois et al. 2018)

▼



3.2.3 Pseudonymization and anonymization

According to GDPR Recital 26, the principles of data protection only applies to personal data, which data can be used to identify a natural person (data subject). On the other way around, GDPR does not apply to anonymous information. When using personal data for research, there are two ways to protect the data subject's privacy: pseudonymization and anonymization.

Pseudonymization

According to GDPR, pseudonymization is defined as the processing of personal data in such a way that the data can no longer be attributed to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organizational measures to ensure non-attribution to an identified or identifiable individual (Article 4(3b)).

Pseudonymization is a reversible process, meaning that there is a way to trace back to the person that the data belongs to. So pseudonymised data is still personal data, which must comply with GDPR.

Anonymization

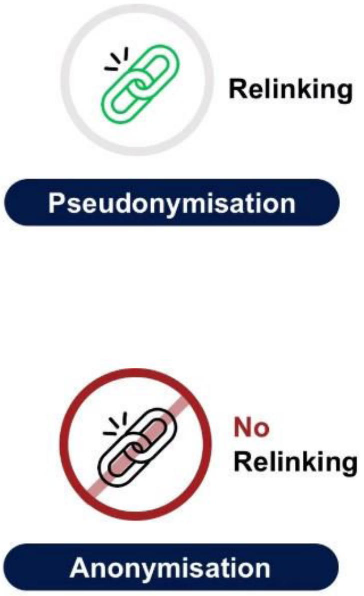
Comparing to pseudonymization, anonymization removes any data that leads to an identifiable person permanently. There's no way to track down a person using anonymised data. Thus, anonymised data doesn't fall into the scope of GDPR.

However, to completely anonymise data sets is difficult, and doing so may decrease the value of the data sets. For example, when the researchers in TU Delft get Wi-Fi data from the ICT department, the MAC addresses are encrypted. This means that with this data set, the researchers can not track a specific device down, they can just recognise the routines of certain devices, without knowing whose device they are. But if someone compares the Wi-Fi data set with other information, for example, the teachers' open schedule, it is highly possible to match the movement pattern with the teacher's teaching classroom, and identify the device's owner. So this way of encrypting MAC address is only pseudonymisation of the data.

If the ICT department want to completely anonymise the data, they have to remove any predictable individual movement as well. So for example, instead of revealing certain persons' movement routine, they may only provide data concerning a group of people, like “200 people are at the library at 9 am”. However, by completely anonymising the data, the data loses a lot of value for research as well.

Besides, GDPR also states that the data subject shall have the right to withdraw his or her consent at any time (Article 7(3)). If a data subject’s data is completely anonymized, there’s no way for the withdrawal. If the data subjects want more control over their data, it’s better that their data is still identifiable. That means anonymisation should be avoided in practice (O'Brien, 2009). Not only since anonymisation excludes data linkage or update, but also because anonymisation takes away most legal obligations to protect the data or respect individual rights or interests, while the (hypothetical) risk of re-identification remains (Greely, 2007).

In summary, if the data requesters want to process the data more flexible without GDPR, they can only get anonymised data which could contain less information than the original data. In the current situation of TU Delft, it is almost impossible to completely anonymise the Wi-Fi access data if it’s going to be used in research. Meanwhile, thinking about the ethical issues that to give more control to data subjects, it’s also not possible to anonymise the data as well.



▲
Figure 3-3: The different between pseudonymisation and anonymisation

Chapter

4

Context and stakeholder research

This chapter introduces the process of doing context and stakeholder research. The whole research and design process and methods of the context and the stakeholders are shown.

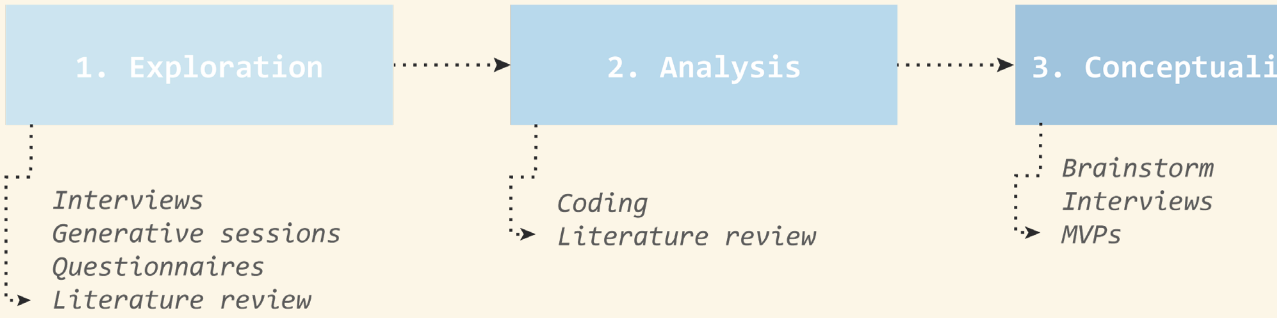
4.1 Research and design process and approach

The whole research and design process of this project contains four phases: exploration, analysis, conceptualisation and evaluation. The methods used in different phases is shown in Figure 4-1.

In the exploration phase, interviews, generative sessions, questionnaire and literature review are used as methods. Since there are three stakeholders in this context, it's necessary to use a variety of methods to get to know their insights. Interviews are used when understanding the experts, including research experts who work closely with Wi-Fi access data for research, ICT expert who has rich knowledge about ICT data management and a vision of the data platform, data steward who support the researchers to use personal data legally and ethically. Generative sessions are held together with students as data subjects. It involves students from faculty of Industrial Design Engineering, Architecture, Civil Engineering and Electrical Engineering, Mathematics and Computer Science. The setting of the generative session enables the students to express their insights in a half-structured way, about the sensitivity of the personal data and the attitudes towards data collection. Considering the sensitivity of some questions, an anonymised questionnaire is used to encourage the data subject to speak out their thoughts more honestly. Also, the form of the questionnaire helps to get more feedback than the generative session. Besides, literature review is used to explore some basic knowledge of the context and the stakeholders.

After the exploration phase, the information collected are analysed. The quotes from interviews and generative sessions are coded and categorized. Literature review is also done during this phase to help analysing the result.

When the result of the analysis is summarized, the conceptualisation of the toolkit starts. A brainstorm is firstly organized. Participants are industrial design students and the values of different stakeholders are shown to them, and they brainstorm on how to transform these values into tangible functionalities. The result of the brainstorm provides an embryonic form of the final design. Then a minimum viable product (MVP) of the toolkit is made, to be tested with the expecting audience who come from system design fields to see if the toolkit is clear and self-explanatory. During the iteration of the MVP, there are also interviews with people from ICT department and EEMCS department, to understand the workflow of a platform developer, and the technique problems with platform building. The iteration of the MVPs are based on the feedback of the students from Faculty of Computer Science, they give feedback after they tried to use the toolkit, and the comments are on both form of the toolkit as well as the content.



4.2 The data platform

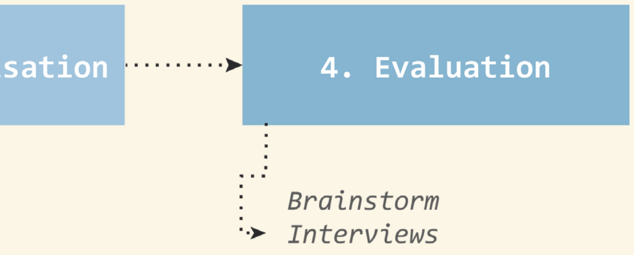
The aim of this project is to support the design and development of the future data platform in TU Delft, where data controllers and data subjects can manage and process the data more conveniently and legally. During the first stage, the data shared on the platform will be the Wi-Fi access data whose owner is the ICT Department. After the platform is well-developed, other data from different data controllers will be put into the platform as well. Considering the scope of this research, only Wi-Fi access data will be considered during the design stage of the platform.

The Wi-Fi access data will first be collected by the ICT Department for operational use. For this usage, they do not need to ask for informed consent from data subjects. Then, they will put the data into the data platform, and the data requesters can directly access those data from the platform. There are two key points for designing this platform.

1. Research as the legal ground

Comparing to the original source of the Wi-Fi data which is for operational use, the data on this data platform will only used in research. That means that once the data is moved to the platform, the legal ground for the data changes. The legal ground is crucial when keeping the platform in compliance with GDPR.

As analysed in Chapter 3.2, there are two difficulties when implementing GDPR into the platform. The first is that it is not feasible to ask for every data subject's consent to process their Wi-Fi access data, because the Wi-Fi data comes from different devices. That means when someone buys a new device or changes a device, they are asked for informed consent once again. The second difficulty is anonymisation. A conventional method to protect data and avoid consent or other legal requirements is anonymisation. Yet, there seems to be a broad consensus that it is impossible to guarantee anonymity (Mostert, 2016). It is not the practical reality that a clear



▲

Figure 4-1: The research and design process and methods

distinction between pseudonymous and anonymous data can always be made (Sethi, 2014).

So to find an alternative besides informed consent and anonymisation, the legal ground for research is created for the data platform. When the data is being used for research that serves a high public interest, consent is not required.

2. Implementing values into the platform

GDPR's aim is about giving data subjects more control over their data. Apart from the law, as discussed in Chapter 2, technology design is an inherent ethical process, that all the stakeholders' values should be considered. In this case, the direct users of the platform are data controllers and data requesters, but it's also important to think about the data subjects as indirect stakeholders. Their values about data are also important for developers of the platform.

4.3 Understanding the stakeholders

To complete a total value-based data platform, the most important thing is to explore the stakeholders' values. A variety of research has been done to understand the three different stakeholder entities: the data controllers, data requesters and data subjects.

4.3.1 Data controllers

1. What do data controllers do?

The data controller in this context refers to the ICT Department of TU Delft, who maintains Wi-Fi infrastructures and owns the Wi-Fi access data. The whole ICT Department for operation has around 150 people, while the innovation department that supports research has about 8 or 9 people. The operation staff help the students and employees with technical problems, for example, the printing not working or their laptop failing to connect to eduroam.

Meanwhile, the ICT Department also do educational support for students and employees. For example, its innovation team supports researchers with special needs, such as providing special network connection and data management for research groups, enabling online courses techniques for students.

In the context of the research, the ICT Department is the initiator of the data platform. People from this department will mainly be the developers of this platform. When the platform is built, the first data they will put into is their own data, for example, the Wi-Fi access data. They will use the platform to do better educational support and improve the data management.

2. Method

To understand the data controller's value, an expert interview was conducted. Lolke Boonstra, who's doing ICT support for research in the ICT innovation department was interviewed. Lolke is also the initiator of the data platform, so his vision about the platform was also asked.

Apart from Lolke who works in the innovation department where they focus on supporting the researchers, to understand the value of the whole department better, an ICT technical staff from service desk was also invited to have a casual chat about his personal goals and values.

3. Key insights

Internet maintenance as the main value

The main value of the ICT Department is without a doubt to provide smooth Internet experience for all the campus Wi-Fi users. They do not intend to track any user with the collection of the Wi-Fi data. The data collected by them are used for analysing user distribution and session, in order to provide enough Wi-Fi access points, and provides seamless internet experience even when an error occurs.

Technical support

Apart from internet, they also provide support for other hardware things for students and employees, for example, fixing laptop and printer issues. As a technical person at the service desk, their goal of the work is to help the students and employees solve technical issues and support them to be fully devoted to their study and research. This is the same with ICT's main goal, that is to maintain the whole ICT infrastructure.

"Maintaining the whole ICT infrastructure, that's the main value. But there also should be a shift, to also shift more to ICT support for education and for research. "

——Lolke Boonstra,
research expert from ICT
Innovation Department

Educational support as the new goal

Currently, ICT is providing educational support to students and teachers. For example, developing e-learning tools and applications that improves teaching and self-study experiences, also provide data for researchers if they need it for research. Comparing to their image as a service provider, there is going to be a shift in the strategic policy of the board, that is to shift more to support for education and for research. The envisioned data platform is one step for them to give more support for researchers.

Implementing GDPR in TU Delft

At the current stage, GDPR is not implemented in the TU Delft yet, and there is not a standard procedure for researchers to obtain data from ICT Department. They rely on personal contact to get the data, and the ICT Department will pseudonymize the data before giving to the researchers. This way of passing data actually interferes with ICT's operational works. Also, it is not fully GDPR compliant.

With the development of the data platform, it will help the researchers not to interfere with ICT operation. And it also fits ICT's current strategy to have more support for research. It is also one important step during the process of implementing GDPR in the university, that get the data management more standardized.

"My goal is to help the students solve all the technical problems, so they can focus on their study."

——ICT staff at service desk

"TU Delft didn't implement GDPR. We're trying to implement it now, but it cost effort."

**——Lolke Boonstra,
research expert from ICT
Innovation Department**

4.3.2 Data requesters

1. What do data requesters do?

The data requester in this context refers to the researchers and designers in TU Delft who need to obtain Wi-Fi access data for their project. They could be researchers, teachers, PhD candidates and students. When they need certain data set for research, they either apply from data controllers, or if they cannot do so, they have to collect the data themselves.

At the current stage, to obtain data could be difficult for researchers. Usually there's one or more persons responsible for applying for the data, and other people in the research group can just focus on their work. When it comes to sensitive private data, and they have no idea if it's ethical or legal to publish the result, they go to the data steward in every faculty for help.

In the context of the research, the researchers are the direct users of the data platform. Since the platform is only responsible for its data, the researchers also have to know GDPR well to handle those data in a proper way.

2. Method

For MSc Geomatics students, there is a Geomatics Synthesis Project to accomplish, which includes quite a few topics using Wi-Fi access data to analyse students' movement pattern. To understand the data requesters values, two teachers and one former master student who participated in one of those projects were interviewed.

> **"Wifi access points are there to transmit data. Wifi is available everywhere. You can use that already. You don't have to do anything, and it's free."**
>
> **——Edward Verbree, expert of geoinformation from GIS Department**

> **"It's hmmm… You just never thought about it [ethical issues behind Wi-Fi access data]."**
>
> **——Balázs Dukai, research software engineer from 3D Geoinformation Research Group**

These projects that uses Wi-Fi access data to analyze students' movement pattern and building occupancy are initiated by the Faculty Management Department. They want to use the result to have an overview of people's movement and distribution in and between the buildings, in order to improve building space usage. The teachers who organize the projects find the clients, and one of the clients is facility management department. So the projects matches perfectly with this research.

Besides the insights of the data usage, technical questions were also asked during the interviews, for example, how does Wi-Fi locating works and what's its application.

3. Key insights

Passion for the research

It's not surprising that the data requesters are more interested in the technical part of their research. When talking about their projects, they were passionate introducing how they used the data to accomplish their goals, and how these data could be used.

Dedication only to the research goal

Comparing to the research goals, they don't care a lot about how do they get the data. They focus on achieving their research goals, and how was the data collected and if it's pseudonymised properly doesn't matter a lot to them. When being asked "have you ever considered about the ethical issues when using the data?", they all mentioned that they know that the data could be used to identify a certain person, but they never thought about the ethical thing before. It feels like ethical things shouldn't be the issue for them researchers to handle, because they only uses the data to get a result. They don't think much about the ethical issues behind it simply because it's not within their scope.

Privacy data

They are also aware of the fact that these Wi-Fi access data may cause privacy issues, including recognizing or locating a certain person with extra details. For example, even though all the MAC addresses in the dataset are encrypted when being given to the researchers, there is still a possibility to recognize a specific person's MAC address. With additional information like a teacher's open schedule, his routine can be found. If a researcher compares his routine with different device's movement, it's highly possible to find out which MAC address/device belongs to him.

It also happens when a member of the project identify a specific person from the Wi-Fi access data. Additional information was used to compare to identify a MAC address, and they recognise the person. Surely they are aware of the data subjects' privacy, but it's not their top priority.

Obtaining the data

The Wi-Fi data is free and the Wi-Fi infrastructure is already there, so it's natural for them to use the data to do research. When they fail to get the Wi-Fi access data from ICT Department, they can walk around it, for example to install their own Wi-Fi devices to collect the data. To get the data this way is not a good choice for both them and ICT. For them, it's time and effort consuming, and the data collected is less than the campus Wi-Fi data. For ICT Department, the other Wi-Fi installation is interfering with the campus Wi-Fi access point, may lead to a bad impact on the Wi-Fi service.

"We only used the data to accomplish our assignment. I don't know why we didn't have access to the eduroam dataset, so we used our own Wi-Fi infrastructure. Maybe it's about privacy issue, I don't know. I never thought about this."

——Kaixuan Zhou, former master student participated in the Geomatics Synthesis Project

4.3.3 Data subjects

1. Who are the data subjects?

The data subjects in this campus are students and employees who use eduroam. Comparing to other two stakeholder entities, this group of people have less knowledge and awareness of personal data. When the data platform is created, they might not be the direct stakeholder of the platform, depending on how much control they will have over their own data, and how willing are they to manage their own data.

2. Methods

Due to the accessibility to those data subjects, only students are researched in this exploration phase. Most respondents come from Faculty of Industrial Design, a few from Faculty of Architecture, Faculty of Civil Engineering and Faculty of Electrical Engineering, Mathematics and Computer Science. Two methods are used to understand them: questionnaire and generative sessions. The first method aims to understand their awareness of privacy issues and to gain an overall picture. The second method aims to encourage them to tell their insights in a detailed way.

Hypotheses

Before the 'real' exploration took place, an informal conversation with 4 students from Faculty of Industrial Design took place. It helps to understand how normal students think about privacy and campus Wifi. Based on the discussions, 6 hypotheses were raised, based on which further research will be conducted. The hypotheses are:

1. Most people naturally don't want to share their data.
2. People will agree to share their data when they really need the service.
3. Students at campus would like to know what their data is being used for.
4. Students at campus are more likely to share their data for research purposes.
5. Anonymized data collection will give students a sense of privacy.
6. "Big data" will give students a sense of privacy.

Questionnaire

To understand student's overall attitude towards the campus Wi-Fi eduroam, a questionnaire consists of 7 questions were sent out. It's a very short questionnaire, to collect intuitive answers from the respondents. And it is designed to understand three aspects of the students' insights:

- ***Awareness (of Wi-Fi access data)***
- ***Privacy sensitivity (towards Wi-Fi access data)***
- ***Willingness (of sharing personal data)***

As a result, a total of 65 respondents answered the questionnaire, all of them are students in TU Delft, mainly coming from the faculty of Industrial Design, Architecture and Civil Engineering. The result of the three aspects are shown below.

Results of the questionnaire

1) Awareness (of Wi-Fi access data)

Most students don't know of the fact that their personal data is being collected through campus Wi-Fi. They also have no idea what kind of information is exactly being collected by the ICT. It means that they have little awareness

2) Privacy sensitivity (towards Wi-Fi access data)

After they are told that some of their personal data is being collected through Wi-Fi, they are asked to rate how secure about the campus Wi-Fi. The result shows that they're quite neutral about the Wi-Fi security: they either don't find it's extremely safe using campus Wi-Fi, or they don't feel distrustful using the Wi-Fi. However, there are few things that they don't feel comfortable doing so under the campus Wi-Fi. These things are:

1. Online payment, which relates to property safety
2. Private websites that has nothing to do with study, which relates to privacy
3. (Pirate) downloading, which relates to privacy

3) Willingness (of sharing personal data)

A majority of the respondents give consent to share their information for research purposes, before the questionnaire started. Even though they have been told that choosing disagree to share the information will not affect answering the questionnaire, most of them still give consent, without knowing what kind of research it is, and what kind of information will be given out. This means that most students don't mind to share their information for a research in the TU Delft campus. Even some of them are not willing to share when they don't know the purpose, they finally agree when the purpose and other details are explained.

Generative session

While the questionnaire above helps to grasp an overall picture of the students' insights about campus Wi-Fi, a generative session is at the same time being carried on for detailed and deeper insights. The generative session involves 12 students from Faculty of Industrial Design, and the students come from the Netherlands, China, Korea and Portugal. Most sessions were carried out separately, so the respondents wouldn't affect each other.

The session use Friedman's VSD method Scalable Information Dimensions. It provides the respondents different scenarios (3 conditions in total) where different kinds of data (7 kinds of data in total) are being collected. In different scenarios, they are asked to grade on a scale from low to high, on how much they would like to share their data in different conditions and for different purposes (5 purposes in total). And then they are asked why do they think so. The important quotes are recorded for further analysis.

The *3 different conditions* are based on the degree of (de-)identification, which are:

Condition 1: ***Data collected as real-name, traceable data (Real-name)***

Condition 2: ***Data collected as anonymous data that can be traced to one user profile (Psedonymised)***

Condition 3: ***Data collected as anonymous data that can only be reviewed as 'big data' (Anonymised)***

7 kinds of data will be collected, which are:

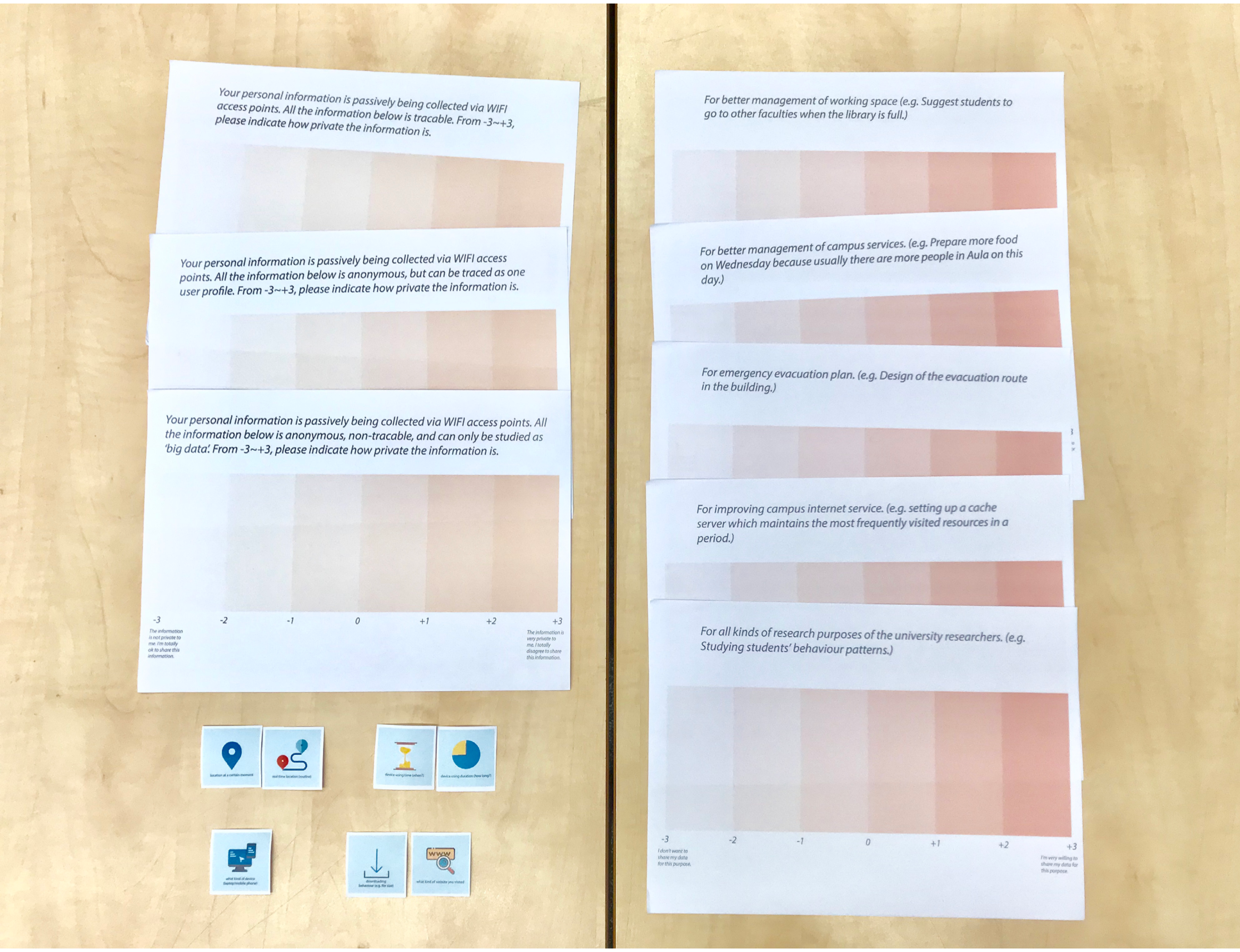
1. real-time location
2. location at a certain time
(less specific than real-time location)
3. session of using the device
4. duration of using the device
(less specific than session)
5. what kind of device you are using
(mobile phone/laptop?)
6. what kind of website you visited
7. downloading behavior
(e.g. size of the files)

And *5 purposes* to collect the data are given (under condition 2), which are:

1. for better management of working space
2. for better management of campus services
3. for emergency evacuation plan
4. for improving internet service
5. for research purposes of the university researchers

Figure 4-2: The setting of the generative session

▼


-3
-2
-1
0
+1
+2
+3
The information is not private to me. I'm totally ok to share this information.
The information is very private to me. I totally disagree to share this information.
I don't want to share my data for this purpose.
I'm very willing to share my data for this purpose.
location at a certain moment
real-time location (routine)
device using time (when?)
device using duration (how long?)
what kind of device (laptop/mobile phone)
downloading behaviour (e.g. file size)
what kind of website you visited

Your personal information is passively being collected via WIFI access points. All the information below is tracable. From -3~+3, please indicate how private the information is.

Your personal information is passively being collected via WIFI access points. All the information below is anonymous, but can be traced as one user profile. From -3~+3, please indicate how private the information is.

Your personal information is passively being collected via WIFI access points. All the information below is anonymous, non-tracable, and can only be studied as 'big data'. From -3~+3, please indicate how private the information is.

For better management of working space (e.g. Suggest students to go to other faculties when the library is full.)

For better management of campus services. (e.g. Prepare more food on Wednesday because usually there are more people in Aula on this day.)

For emergency evacuation plan. (e.g. Design of the evacuation route in the building.)

For improving campus internet service. (e.g. setting up a cache server which maintains the most frequently visited resources in a period.)

For all kinds of research purposes of the university researchers. (e.g. Studying students' behaviour patterns.)

Results of the generative session

1) Being sensitive about content information

When the information is real-name, most people find the content information (browsing history and downloading behaviour) most private. Though these contents are not part of the Wi-Fi access data, the respondents still have doubts once they learn that some of their data is being collected.

"Even though you tell me it's just the size of the file, not the content, I still have some concerns. Because I don't know what's being collected."

2) Location concerns personal safety

Almost all the students put location data on a very private degree. While the respondents pointed out that the content information (website viewing & downloading history) will make them feel privacy being violated, it's the location information (real-time & certain time location) that make them feel unsafe, because others could find them using this information. They don't want to be located and find out, because they feel the personal safety being threatened. The fact that the data is only in the campus could be a good thing because it's safer in the university. But on the other hand, it's also easier to target and locate someone when in a campus.

"Location is different, especially real-time location. It means others can find me using this information. It's scary."

3) Device information may reveal financial situation

Some students especially care about their device information. First, they fear the information could be sold to others and they may receive advertisements. Second, the device information may reveal their brand preferences and financial situation, so it's also private to them. Third, they think that what kind of device they are using is a complete personal thing. They don't feel it's related to a university research.

"I'm afraid my device information could be sold to sales company. I don't want any salesman to call me. For example, I use Apple device, but I don't want any Apple advertisements."

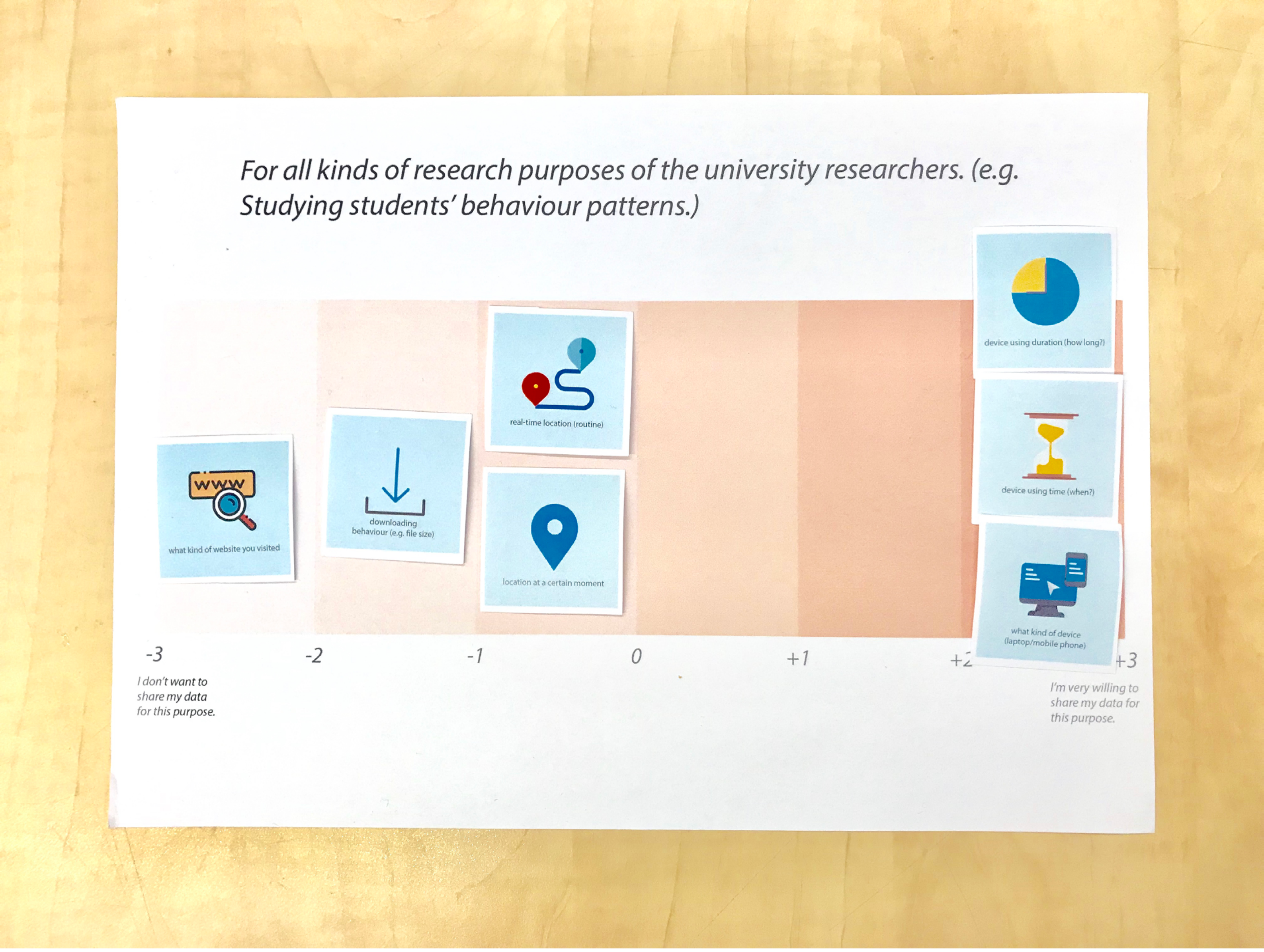

For all kinds of research purposes of the university researchers. (e.g. Studying students' behaviour patterns.)
www
what kind of website you visited
downloading behaviour (e.g. file size)
real-time location (routine)
location at a certain moment
device using duration (how long?)
device using time (when?)
what kind of device (laptop/mobile phone)
-3
-2
-1
0
+1
+2
+3
I don't want to share my data for this purpose.
I'm very willing to share my data for this purpose.

Figure 4-3: An example of generative session result

4) Pseudonymization is not enough for some of the respondents

While some of the respondents find it completely secure when their data is being de-identified, some others still find it risky when all their data could be studied together as one user profile. They believe that everyone has his own pattern, so it's very likely that they will eventually being recognized if their own user profile is being studied.

Some respondents mentioned that pseudonymization is even more uncomfortable than real-name. Once they know there is still a chance to find out their real identity using the pseudonymised data, they feel like it's a fraud. If they are told that their data is being used in a real-name condition, they think at least it's transparent for them. But once their data is being used pseudonymously, they don't have an idea when their identity will be found out.

“I would even prefer condition 2 than condition 1. Because I don’t fully trust the data user. If you tell me that there’s still a chance to find who I am, at least I will be prepared for it.”

5) ‘Big data’ is acceptable for most of the respondents

While some of the respondents find it completely secure when their data is being de-identified, some others still find it risky when all their data could be studied together as one user profile. They believe that everyone has his own pattern, so it’s very likely that they will eventually being recognized if their own user profile is being studied.

6) The purpose of the data collection greatly influences the willingness of sharing data

The purpose is an important factor that affects what kind of data respondents want to share. For a lot of times, they are not so willing to share certain data not because they think it’s private, but because they think it’s not relevant to the purpose. They don’t want to share their data for the sake of sharing it; they want it to be useful. And they really want to know what kind of purpose it is to collect their data. Sometimes even though the data is private to them, they may still share it if it’s for a good purpose.

“I don’t want to share my viewing history information for managing working space. It’s not a privacy matter. I don’t know why they need this for that purpose. I rate my willingness only based on relevance.”

7) They trust the university more than other places

The fact that they are in the university do encourage the respondents to share their information. Some people mentioned that their location information will be a private matter if it’s in other place, but in the campus it doesn’t matter. Some people also mentioned that they don’t want to share data is not because they don’t trust the university, it’s because they are afraid of a data leakage, or their data being sold to third parties.

Validation of the hypotheses

With the result of the questionnaire and the generative session, all the six hypotheses are validated/invalidated.

1. Most people naturally don't want to share their data.
Invalidated. It totally depends on the condition and the purpose.

2. People will agree to share their data when they really need the service.
Validated. The purpose is a strong influencer for the respondents.

3. Students at campus would like to know what their data is being used for.
Validated. They want more details if their data is being used by others, and they especially want it to be used for good purposes.

4. Students at campus are more likely to share their data for research purposes.
Invalidated. They are only willing to share the data if they think the research purpose is good/interesting. If they are just told it’s a research, they don’t feel responsible to donate their data.

5. Anonymized data collection will give students a sense of privacy.
Invalidated. To some of them, anonymizing their data is still not enough. They’re still concerned about being recognized.

6. "Big data" will give students a sense of privacy.
Validated. When all their data cannot be traced back to one specific person, they are more relieved.

4.3.4 Other stakeholders in the context

Data steward

Apart from the three stakeholders, there are also other stakeholders in the context. When the researchers having difficulty not knowing how to handle the personal data lawfully and ethically, they go to each faculty’s data steward for help. Data stewards are the persons who know the laws and regulations well and help the researchers to avoid troubles. According to Jeff Love, data steward of the Faculty of Industrial Design Engineering, his goal is to help the researchers manage and use personal data correctly. He gives several examples of how data requesters face difficulties: they don’t know how to deal with video information where respondents’ face and voices are revealed; or they don’t know if they can public research which contains personal data from 20 years ago; or they don’t know how to deal with data from other part of the world where GDPR is not effective. In this context, his goal and the goal of the data platform is to some extent overlapped, which is to simply support the researchers.

For the researchers, when they are dealing with personal data that is in a grey zone, they need to turn to someone for help. Not all of them read through all the data policies and regulations, so they need someone’s instructions for using the data.

Human Research Ethics Committee

Another stakeholder is Human Research Ethics Committee (HREC) in TU Delft. This committee is an overseer of all projects that concerns ethical human research. The researchers need to get HREC’s approval to do their research projects. The goal of HREC is to protect the welfare of the participants of the research. In this context, when the data platform is developed, it might be easier for researchers to get the approval of HREC, since all the data used from the platform will be GDPR compliant. It potentially support the researchers not only a more standardized procedure, but also more secure data sets.

Chapter

5

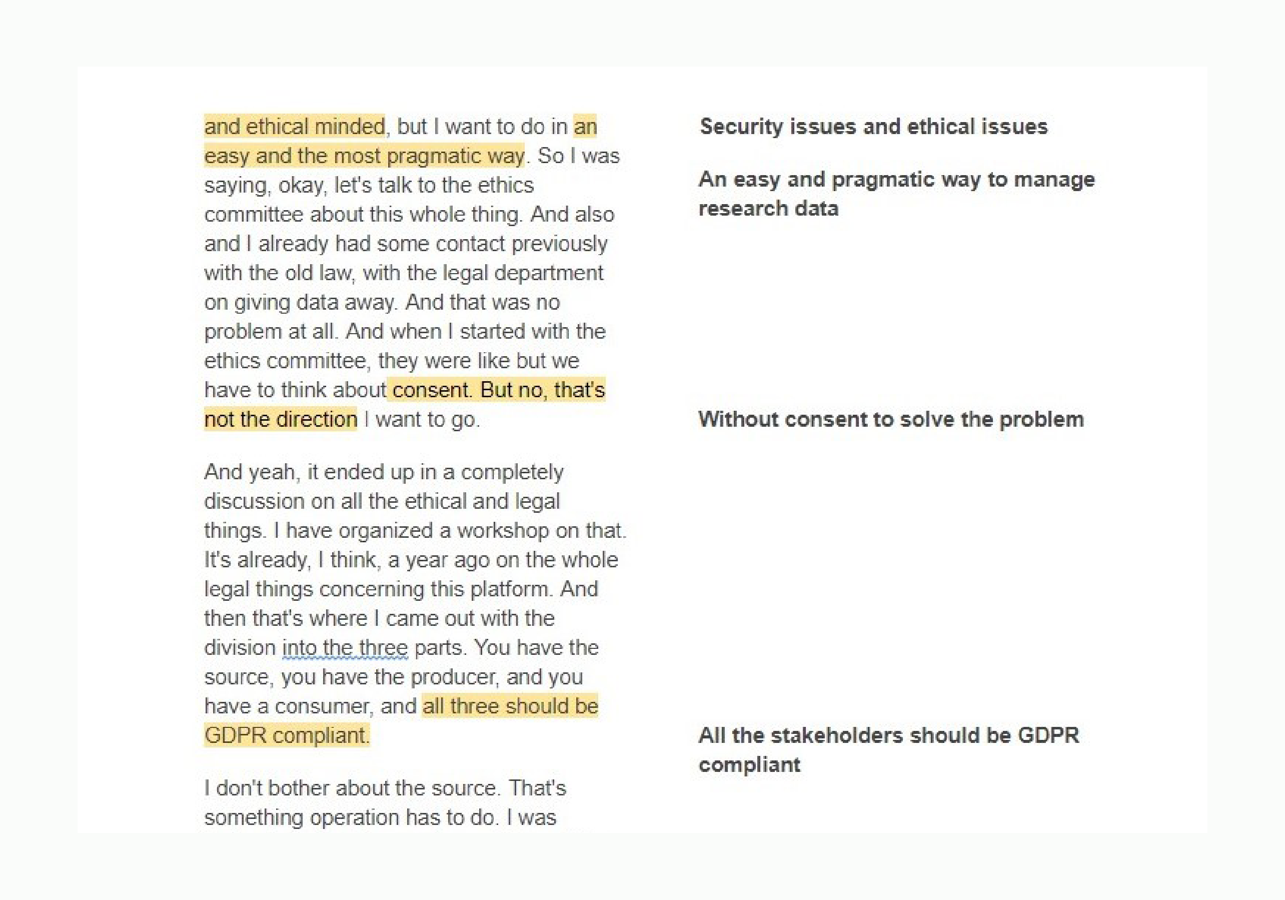
Application

This chapter analyses the result of the research in previous chapters, and started to conceptualize the final toolkit. The method used for designing the toolkit is MVP iterations.

5.1 Value analysis

After diving into all three stakeholder entities insights, it’s time to elicit their values for later on communications. First, the transcripts from interviews and quotes from generative sessions are analysed. According to Friedman’s Value-oriented Coding Manual method, the transcripts are coded with the focus on value-oriented insights. The insights that concerns personal preference of handling the data, about their concerns and worries, about their attitudes towards data collection and processing, and their envisioning of a proper way to handle the data are marked as important quotes. These quotes are highlighted and then rephrased, because they will be a firm step before analysing what kind of value it represents.

Figure 5-1: Coding of the interview



and ethical minded, but I want to do in an easy and the most pragmatic way. So I was saying, okay, let's talk to the ethics committee about this whole thing. And also and I already had some contact previously with the old law, with the legal department on giving data away. And that was no problem at all. And when I started with the ethics committee, they were like but we have to think about consent. But no, that's not the direction I want to go.

And yeah, it ended up in a completely discussion on all the ethical and legal things. I have organized a workshop on that. It's already, I think, a year ago on the whole legal things concerning this platform. And then that's where I came out with the division into the three parts. You have the source, you have the producer, and you have a consumer, and all three should be GDPR compliant.

I don't bother about the source. That's something operation has to do. I was

Security issues and ethical issues

An easy and pragmatic way to manage research data

Without consent to solve the problem

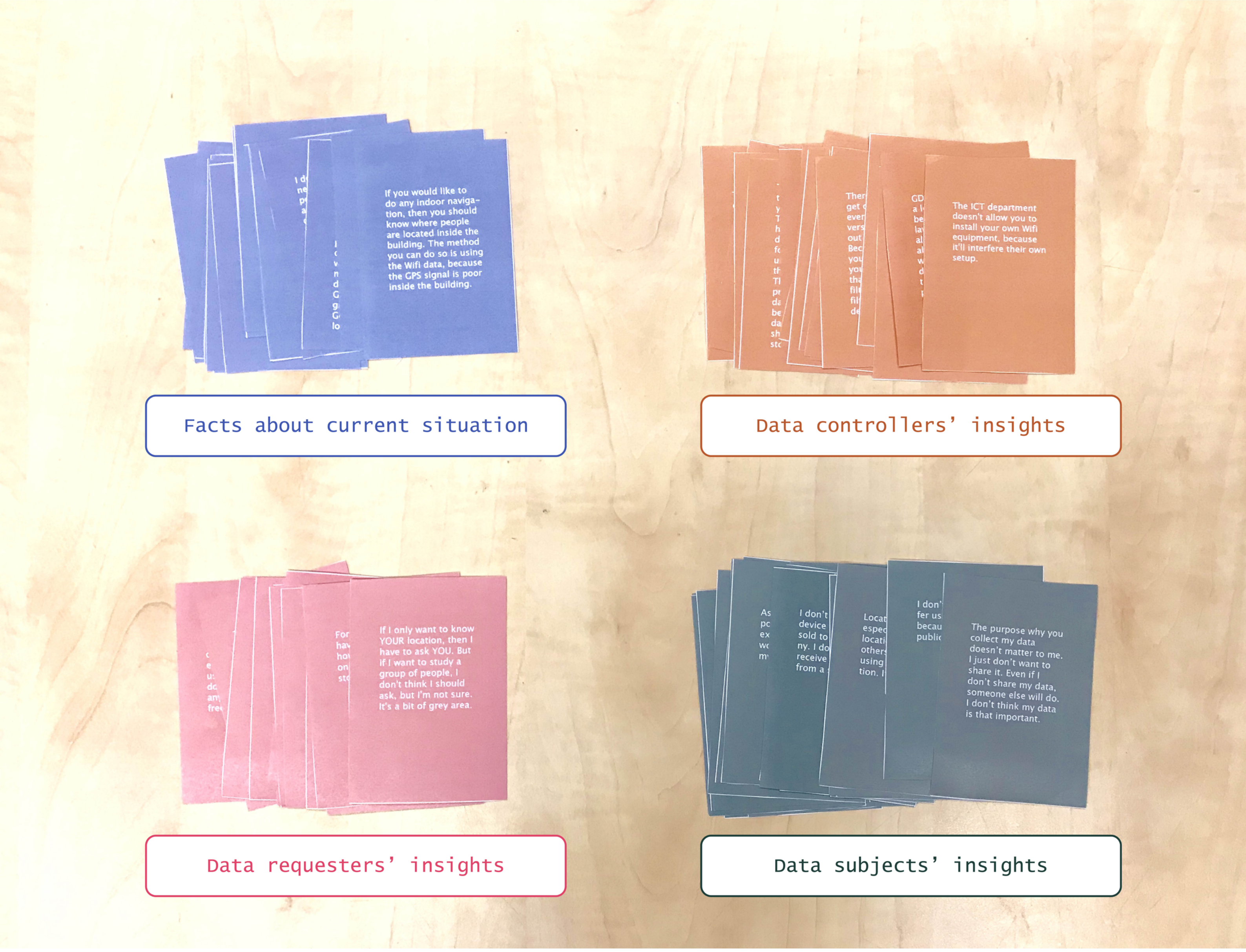
All the stakeholders should be GDPR compliant

After the value-centric quotes are selected from the exploration phase, they are printed on cards with different color background. As the figure shows below, there are 4 groups of quotes.
The blue cards show the fact about the current situation. It is about the context information, for example, how is the Wi-Fi data collected, how do the researchers get the research data at the moment. The orange cards show the data controllers' insights. The pink cards show the data requesters' insights, and the green ones shows the data subjects' insights.

Then these quotes are analysed group by group except for the Facts group. After reading all the quotes in one group, they are clustered into different groups. Each group then gets a label of value that defines the centric value of the whole group.

Figure 5-2: 4 groups of value-centric quotes

▼



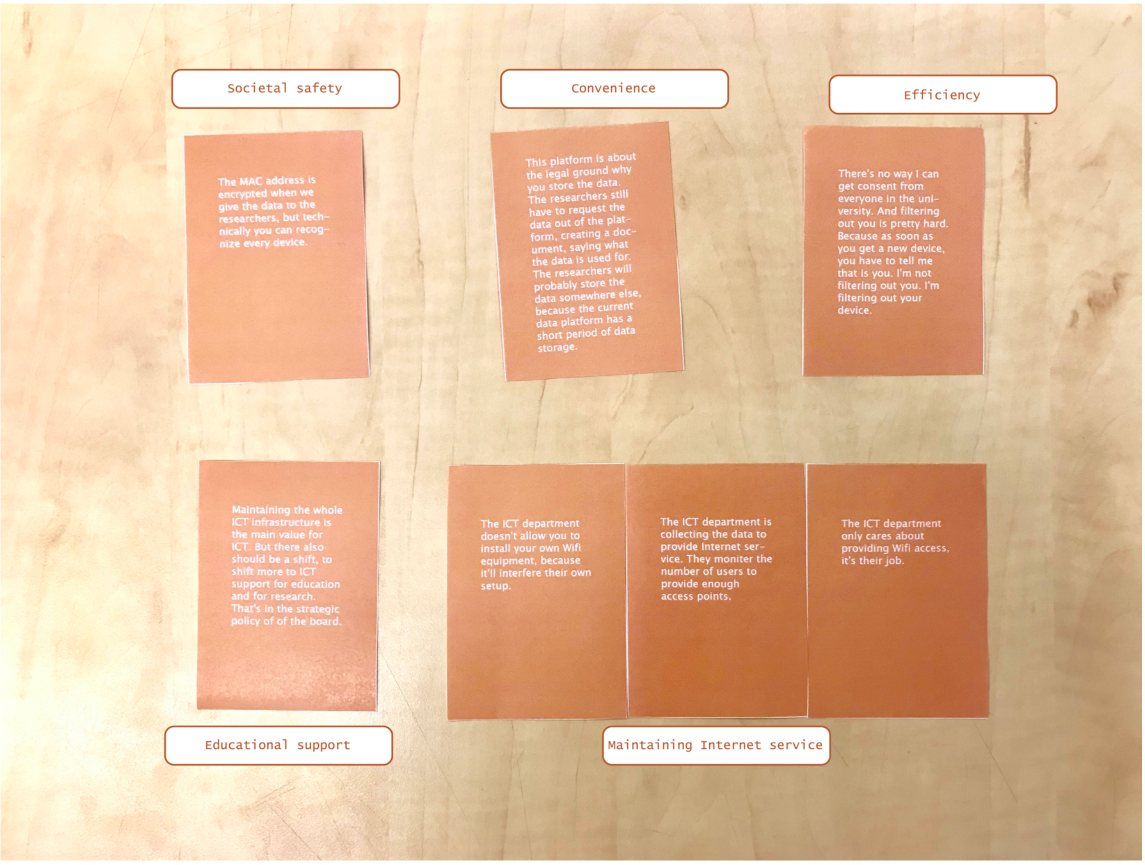
If you would like to do any indoor navigation, then you should know where people are located inside the building. The method you can do so is using the Wifi data, because the GPS signal is poor inside the building.

The ICT department doesn't allow you to install your own Wifi equipment, because it'll interfere their own setup.

If I only want to know YOUR location, then I have to ask YOU. But if I want to study a group of people, I don't think I should ask, but I'm not sure. It's a bit of grey area.

The purpose why you collect my data doesn't matter to me. I just don't want to share it. Even if I don't share my data, someone else will do. I don't think my data is that important.

Facts about current situation
Data controllers' insights
Data requesters' insights
Data subjects' insights



The MAC address is encrypted when we give the data to the researchers, but technically you can recognize every device.

This platform is about the legal ground why you store the data. The researchers still have to request the data out of the platform, creating a document, saying what the data is used for. The researchers will probably store the data somewhere else, because the current data platform has a short period of data storage.

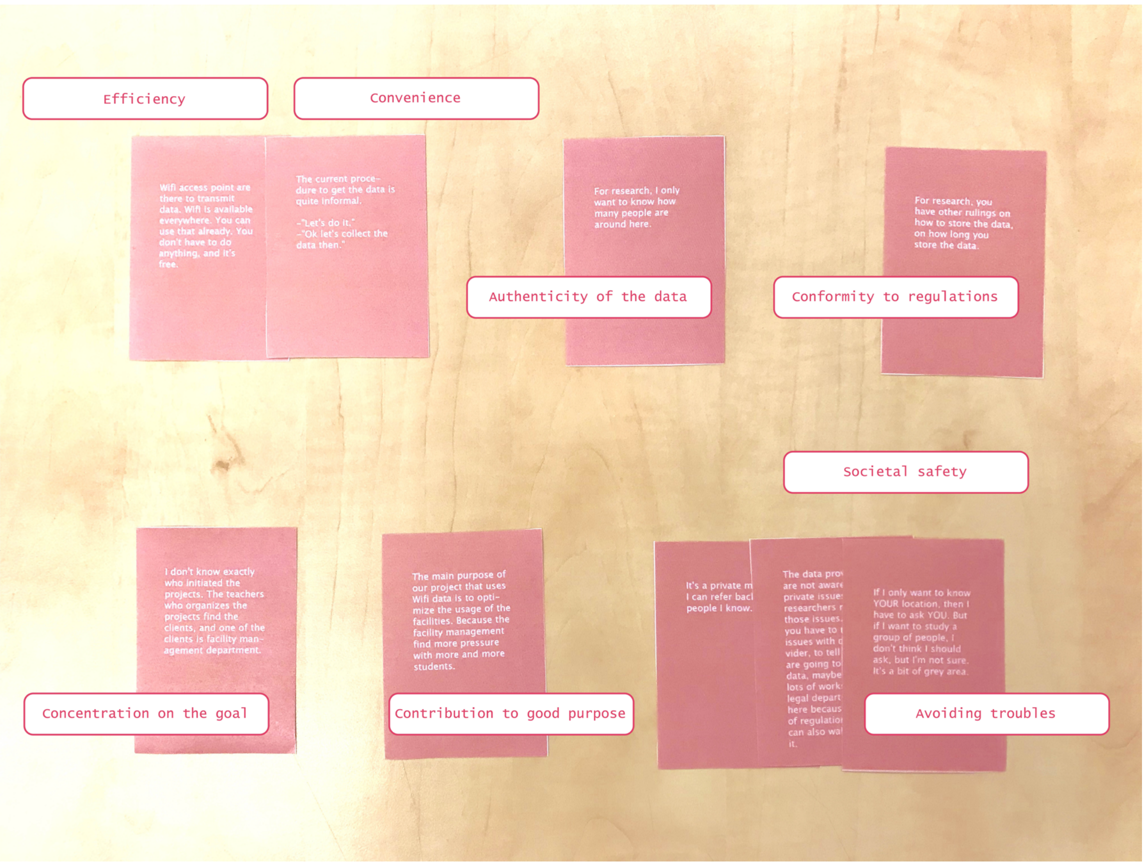
There's no way I can get consent from everyone in the university. And filtering out you is pretty hard. Because as soon as you get a new device, you have to tell me that is you. I'm not filtering out you. I'm filtering out your device.

Maintaining the whole ICT infrastructure is the main value for ICT. But there also should be a shift, to shift more to ICT support for education and for research. That's in the strategic policy of of the board.

The ICT department doesn't allow you to install your own Wifi equipment, because it'll interfere their own setup.

The ICT department is collecting the data to provide Internet service. They moniter the number of users to provide enough access points.

The ICT department only cares about providing Wifi access, it's their job.



Efficiency
Convenience
Authenticity of the data
Conformity to regulations
Societal safety
Concentration on the goal
Contribution to good purpose
Avoiding troubles

Wifi access point are there to transmit data. Wifi is available everywhere. You can use that already. You don't have to do anything, and it's free.

The current procedure to get the data is quite informal.

-"Let's do it."
-"Ok let's collect the data then."

For research, I only want to know how many people are around here.

For research, you have other rulings on how to store the data, on how long you store the data.

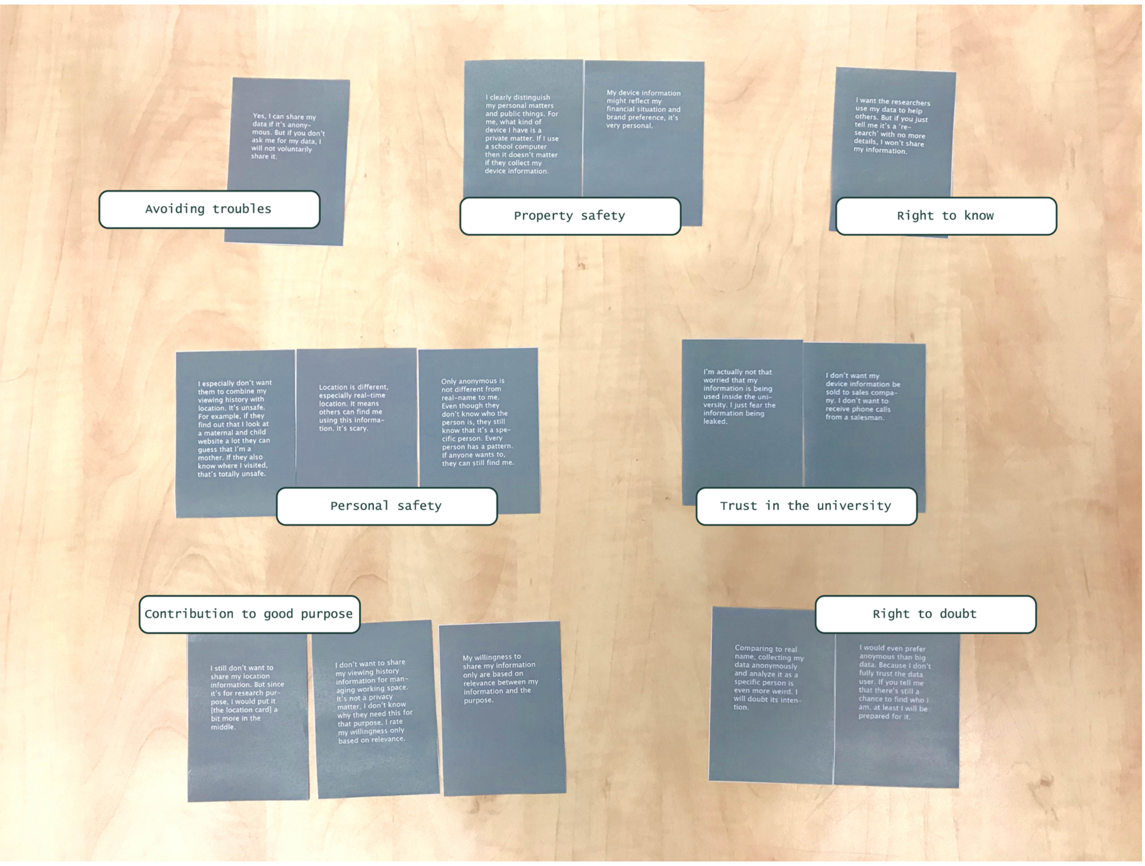
I don't know exactly who initiated the projects. The teachers who organizes the projects find the clients, and one of the clients is facility management department.

The main purpose of our project that uses Wifi data is to optimize the usage of the facilities. Because the facility management find more pressure with more and more students.

It's a private m
I can refer bacl
people I know.

The data prov
are not aware
private issue
researchers r
those issues.
you have to t
issues with d
vider, to tell
are going to
data, maybe
lots of works
legal depart
here becaus
of regulatio
can also wal
it.

If I only want to know YOUR location, then I have to ask YOU. But if I want to study a group of people, I don't think I should ask, but I'm not sure. It's a bit of grey area.



Yes, I can share my data if it's anonymous. But if you don't ask me for my data, I will not voluntarily share it.

I clearly distinguish my personal matters and public things. For me, what kind of device I have is a private matter. If I use a school computer then it doesn't matter if they collect my device information.

My device information might reflect my financial situation and brand preference, it's very personal.

I want the researchers use my data to help others. But if you just tell me it's a 're-search' with no more details, I won't share my information.

I especially don't want them to combine my viewing history with location. It's unsafe. For example, if they find out that I look at a maternal and child website a lot they can guess that I'm a mother. If they also know where I visited, that's totally unsafe.

Location is different, especially real-time location. It means others can find me using this information. It's scary.

Only anonymous is not different from real-name to me. Even though they don't know who the person is, they still know that it's a specific person. Every person has a pattern. If anyone wants to, they can still find me.

I'm actually not that worried that my information is being used inside the university. I just fear the information being leaked.

I don't want my device information be sold to sales company. I don't want to receive phone calls from a salesman.

I still don't want to share my location information. But since it's for research purpose, I would put it [the location card] a bit more in the middle.

I don't want to share my viewing history information for managing working space. It's not a privacy matter. I don't know why they need this for that purpose. I rate my willingness only based on relevance.

My willingness to share my information only are based on relevance between my information and the purpose.

Comparing to real name, collecting my data anonymously and analyze it as a specific person is even more weird. I will doubt its intention.

I would even prefer anoymous than big data. Because I don't fully trust the data user. If you tell me that there's still a chance to find who I am, at least I will be prepared for it.

▲

Figure 5-3: Value groups of three stakeholder entities

◀

The data controllers' insights are grouped into 5 values: ***Societal safety, convenience, efficiency, educational support and maintaining Internet service.***

The data requesters' values are grouped into 8 groups: ***Efficiency, convenience, authenticity of the data, conformity to regulations, concentration on the goal, contribution to good purpose, societal safety and avoiding troubles.***

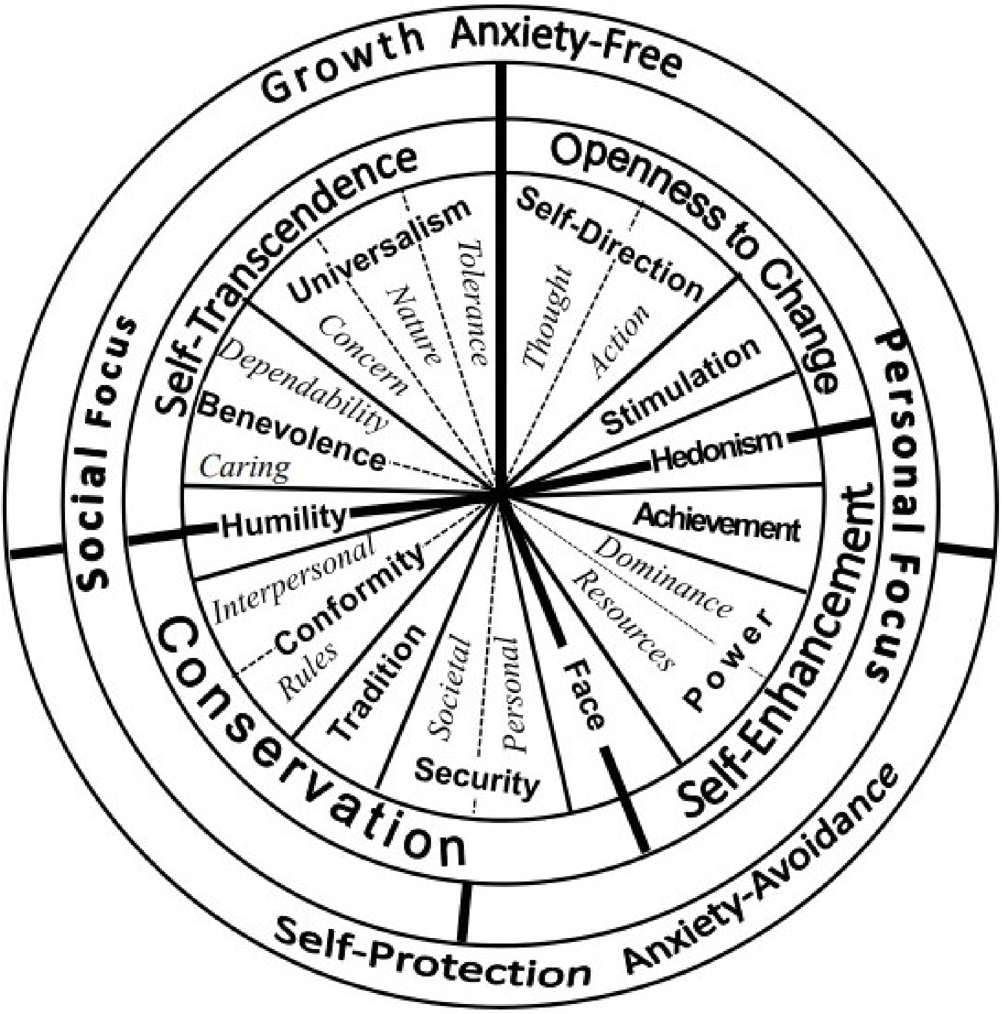
The data subjects' values are grouped into 7 groups: ***avoiding troubles, property safety, personal safety, right to know, right to doubt, trust in the university and contribution to good purpose.***

Since all the values are still very detailed, those values are grouped again into a more broader 'main value', and the label of different groups then become a sub-value that describes the value more precisely. Schwartz's theory of basic human values (2008) is used as a reference when defining the main value. The figure below shows the basic values, it's a theorized circular motivational structure of 19 narrowly defined values with ten basic values and four higher-order values.

To make the values more suitable for this project, this value model only serves as an inspiration. Some of the defined main values, for example, 'Autonomy' 'Trustworthiness' are taken from Schwartz's narrowly defined values, some others, for example, 'Achievement' 'Security' 'Conformity' 'Goodwill' (rephrased from 'Benevolence') are taken from Schwartz's basic values.

Figure 5-4: Proposed circular motivational continuum of 19 values with sources that underlie their order (Schwartz, 2012)

▼



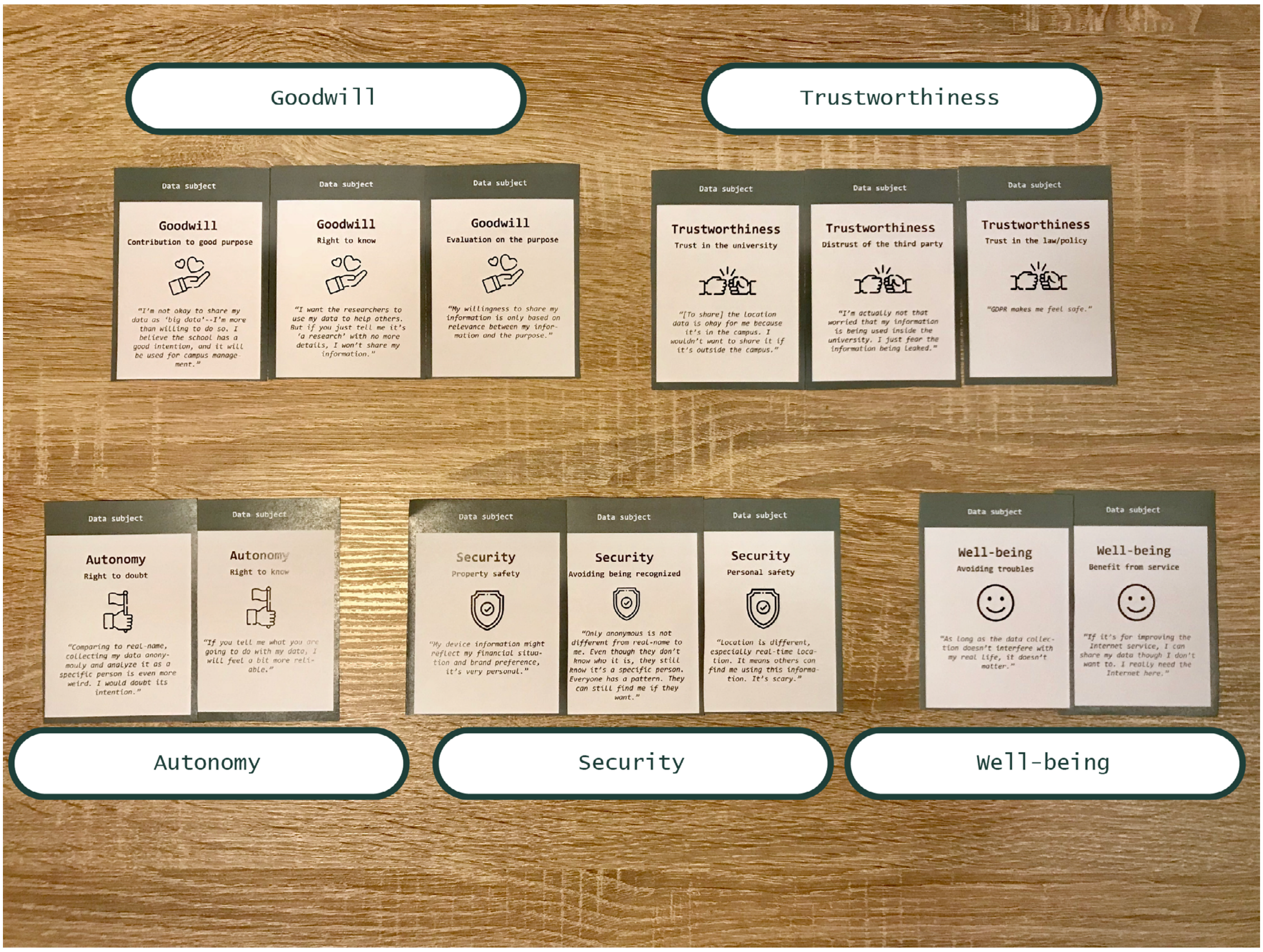
After another round of grouping based on the previous defined values (now they are grouped into a bigger main value, so they become a more detailed sub-value), there are three main values from data controller: ***achievement*, *goodwill and security***. There are five main values from data requester: ***achievement*, *security*, *conformity*, *well-being and goodwill***. There are five main values from data subject: ***goodwill*, *trustworthiness*, *autonomy*, *security and well-being***.

A new card set is made, showing all the stakeholders' values. Still, different back ground color represents for different stakeholders. On the top of the cards, it shows to which stakeholder the value belongs to. And then from top to bottom, there are one kind of value of the stakeholder, then a sub-value of the value, then a quote explaining the value.

Till this stage, the values of all the stakeholders are defined. The conceptualization of the toolkit could begin based on these value cards.

Figure 5-5: After the second round of grouping, data subjects finally has 5 main values.

▼



Data subject
Goodwill
Contribution to good purpose
"I'm not okay to share my data as 'big data'--I'm more than willing to do so. I believe the school has a good intention, and it will be used for campus management."

Data subject
Goodwill
Right to know
"I want the researchers to use my data to help others. But if you just tell me it's 'a research' with no more details, I won't share my information."

Data subject
Goodwill
Evaluation on the purpose
"My willingness to share my information is only based on relevance between my information and the purpose."

Data subject
Trustworthiness
Trust in the university
"[To share] the location data is okay for me because it's in the campus. I wouldn't want to share it if it's outside the campus."

Data subject
Trustworthiness
Distrust of the third party
"I'm actually not that worried that my information is being used inside the university. I just fear the information being leaked."

Data subject
Trustworthiness
Trust in the law/policy
"GDPR makes me feel safe."

Data subject
Autonomy
Right to doubt
"Comparing to real-name, collecting my data anonymouly and analyze it as a specific person is even more weird. I would doubt its intention."

Data subject
Autonomy
Right to know
"If you tell me what you are going to do with my data, I will feel a bit more reliable."

Data subject
Security
Property safety
"My device information might reflect my financial situation and brand preference, it's very personal."

Data subject
Security
Avoiding being recognized
"Only anonymous is not different from real-name to me. Even though they don't know who it is, they still know it's a specific person. Everyone has a pattern. They can still find me if they want."

Data subject
Security
Personal safety
"Location is different, especially real-time location. It means others can find me using this information. It's scary."

Data subject
Well-being
Avoiding troubles
"As long as the data collection doesn't interfere with my real life, it doesn't matter."

Data subject
Well-being
Benefit from service
"If it's for improving the Internet service, I can share my data though I don't want to. I really need the Internet here."

5.2 Goal of the data platform

The design of the toolkit eventually aims at helping the developers create a value-based platform. Thus, the goal of the toolkit shall comply with the goal of the platform itself. According to the values defined in the last chapter, one basic goal from each stakeholder emerges. The three goals below are defined as the basic goals of the platform, and all the functionalities shall be implemented to work towards these three goals.

The data requester want to approach the data efficiently, which means that they can be dedicated to their research goals and not waste time on getting the data and understanding how to process them lawfully and ethically. The data subject want to have their privacy protected better and have more control over their data, even though they are not the target users of the platform. The data controller wants to have a better oversight and usage of the data collection, which means that they have to know which data should be collected and how to use them in a way that they can benefit beyond Internet maintaining. Figure 5-6 shows the basic goals for three stakeholder entities.

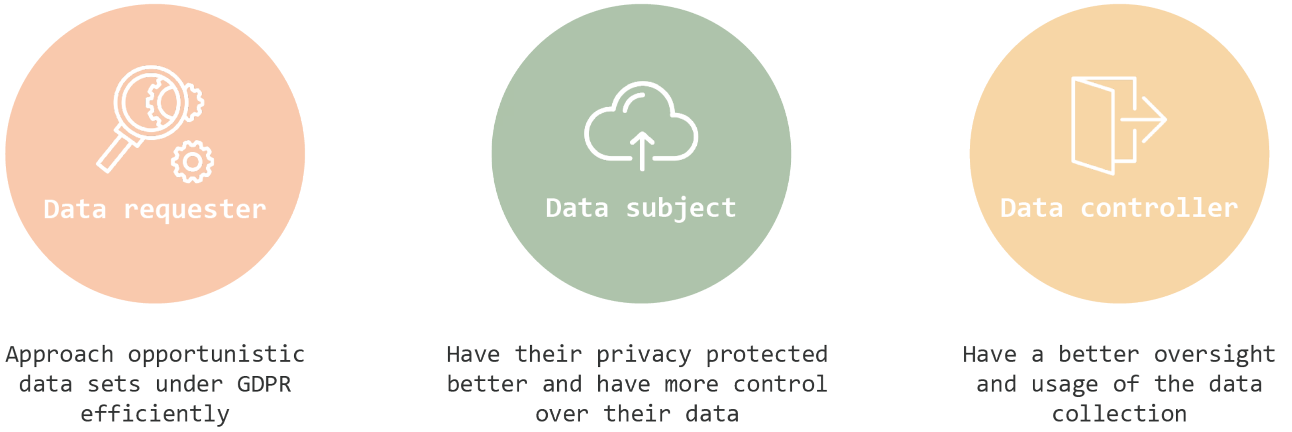

Data requester
Approach opportunistic data sets under GDPR efficiently
Data subject
Have their privacy protected better and have more control over their data
Data controller
Have a better oversight and usage of the data collection

Figure 5-6: Basic goals from each stakeholder entity when using the data platform

5.3 Conceptualisation of the toolkit

Method: Minimum viable product iterations

The approach of conceptualization is to do an iteration process. A minimum viable product (MVP) is made first, then tested with its potential users, improved based on the feedback.

To use this method is because I have little knowledge about a platform developers' knowledge, mindset and workflow. There are many things that could go wrong if the toolkit is not tested with them in time. Also, there is a lot of content in this toolkit, meaning that it's not that easy to get started with for the audience, also for me, there could also be content information going wrong if I don't work closely with those technical person. The MVPs help to improve and to test really fast. There are two rounds of MVPs and testing before the final design and evaluation.

MVP 1

Since the value cards is a good way to communicate value to other people, and it's handy to be used during the work, this form will be kept in the MVPs.
The first MVP consists of three parts: introduction of the context, value cards and value axis, and value dilemma scenario. Apart from value cards, all other parts are canvases that helps to understand the context as well as the value cards.

The aim of the MVP 1 is to help the audience to choose between so many different values. It uses axis to quantify the importance of the values, helps the audience comparing two values from one stakeholder group. It also gives the audience a way of comparing two values from two different stakeholder groups, if there is a value conflict happening.

There are three parts of the MVP1:

1) Introduction of the context

In this part, the two graphs that shows a simplified context and the three basic goals of the platform is shown. The aim of this part is to help the audience have an overview of the context quickly.

2) Value cards and value axis

In this part, the value cards created in Chapter 3 are shown to the audience. The cards consist of three parts: the value, the sub-value and a quote that explains the value. The audience will read through all the cards and have an overview of the stakeholders' insights. Figure 5-7 shows some examples of the value cards.

However, these cards at the moment are of equal importance. While some of them fit with the goal of the platform, for example, the data subject's value contribute to good purpose (goodwill) encourages them to agree to share more data for good research purpose, some others don't support the data collection of the platform, for example, the data subject's value personal safety (security) prevent them from willingly sharing their location data.

So the value axis is created to help the audience map out all the values. There are three value axis, and the only difference is that each of them belongs to a different stakeholder entity, so the vertical axis is different. The audience can put the value cards on the axis with the same stakeholder. On the horizontal axis, the left part is for those values that are negative for the platform construction, while the right part is positive for the platform construction. On the vertical axis, the upper part is for those values that complies with this stakeholder's goal, while the lower part doesn't comply with the stakeholder's goal. Figure 5-8 shows the value axis of the data subject.

After mapping out all the value cards, each value is not the same now. Some of them are more friendly to the development of the platform, and some of them match the stakeholder's main goal better. Figure 5-9 shows an example of using the value axis to map out data subject's values. Then the audience can turn to the next step.

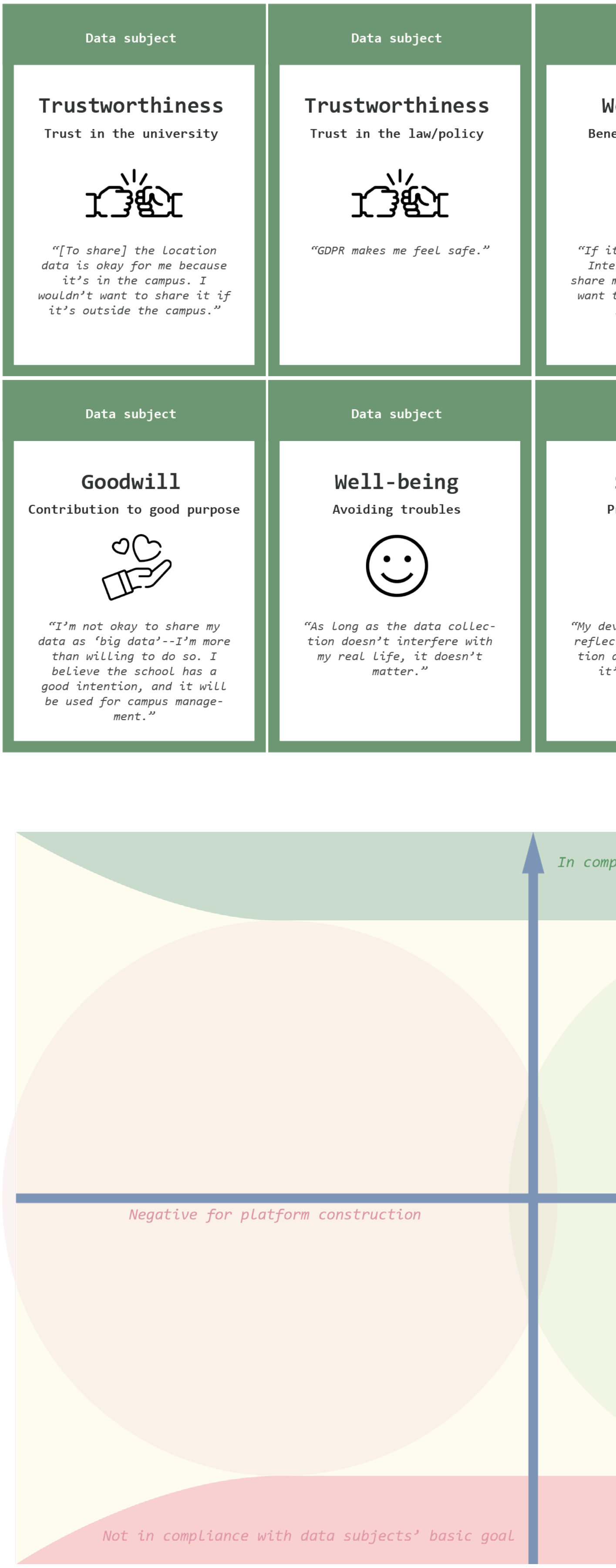
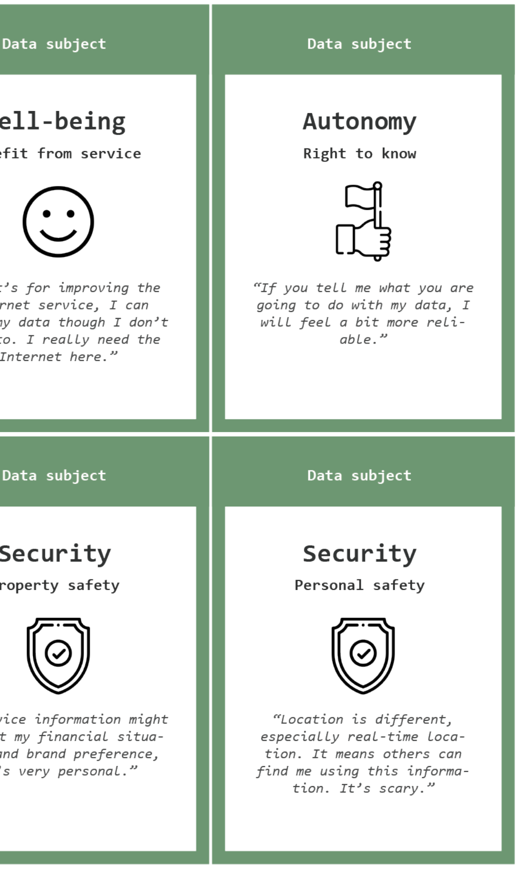
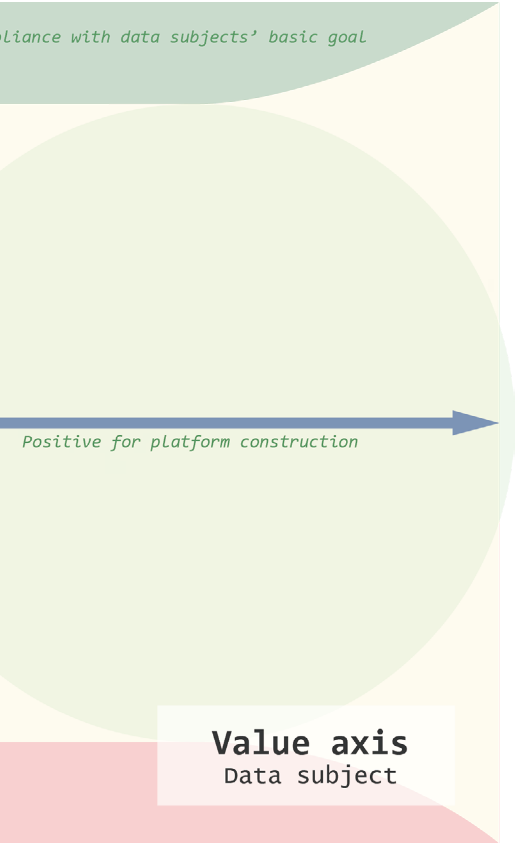

Data subject
Trustworthiness
Trust in the university
"[To share] the location data is okay for me because it's in the campus. I wouldn't want to share it if it's outside the campus."
Data subject
Trustworthiness
Trust in the law/policy
"GDPR makes me feel safe."
W
Bene
"If it
Inte
share m
want t
Data subject
Goodwill
Contribution to good purpose
"I'm not okay to share my data as 'big data'--I'm more than willing to do so. I believe the school has a good intention, and it will be used for campus management."
Data subject
Well-being
Avoiding troubles
"As long as the data collection doesn't interfere with my real life, it doesn't matter."
P
"My dev
reflec
tion c
it'
In comp
Negative for platform construction
Not in compliance with data subjects' basic goal

Figure 5-8: Value axis canvas for data subject



◄

Figure 5-7: Sample value cards of data subjects



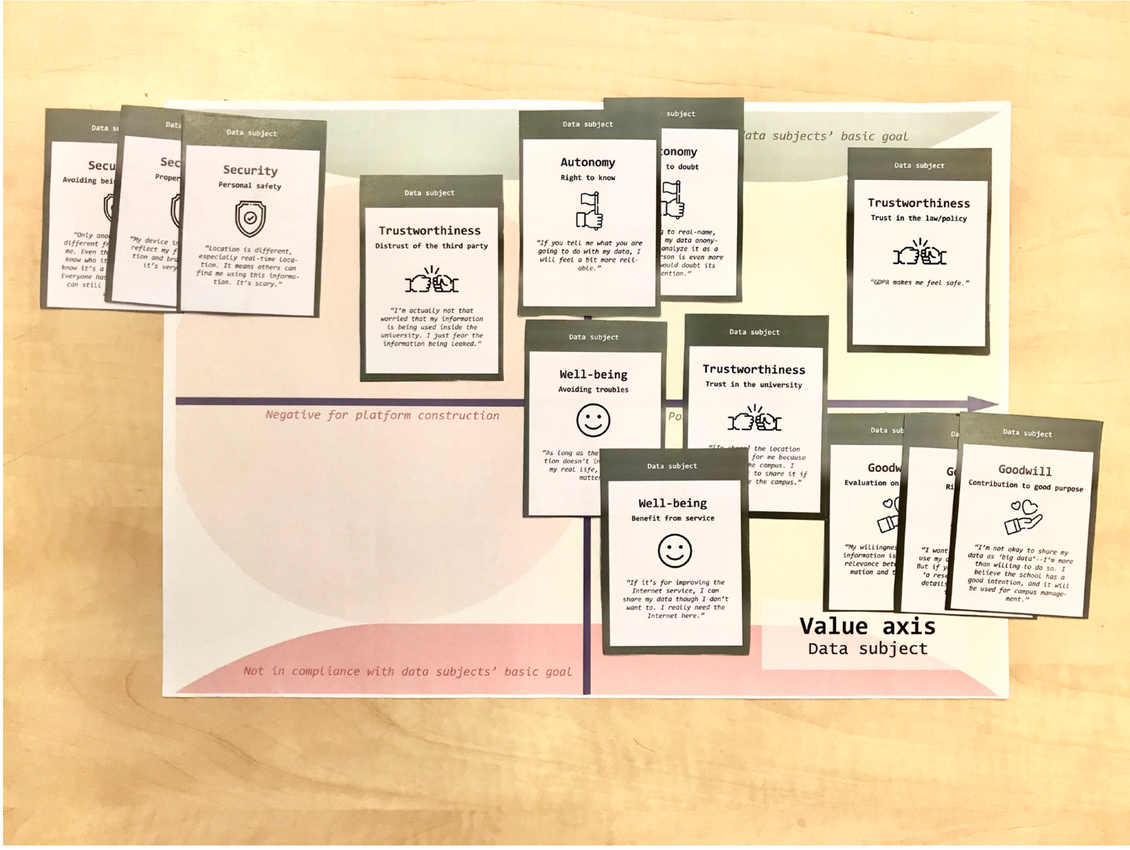
liance with data subjects' basic goal
Positive for platform construction
Value axis
Data subject

3) Value dilemma scenario

When dealing with all sorts of values together, it's inevitable that there are value conflicts. The question is, when value conflict emerges, which value is of greater importance? To help the audience to answer the question, a value dilemma map is created.

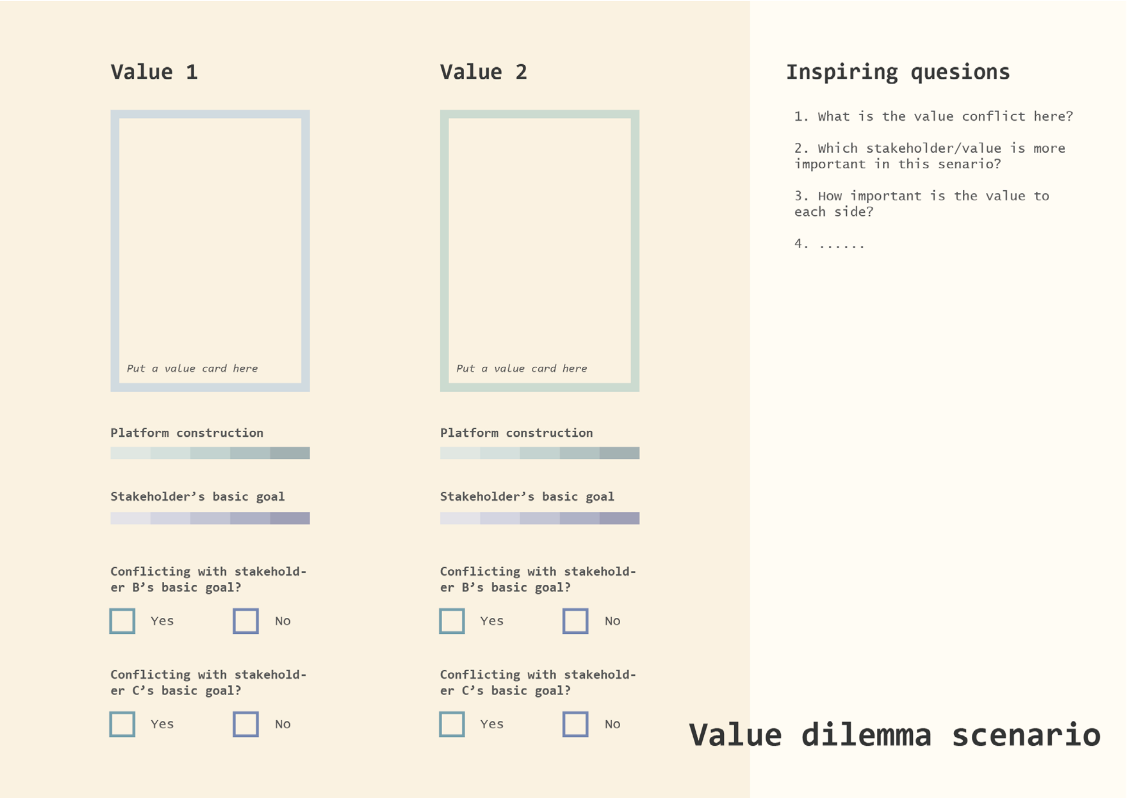
The audience can pick up two value cards that have conflict with each other and put them into the two boxes on the left. Below there are 4 questions that help to 'grade' the two values. The first question is how much does the value supports the platform construction. Based on the mapping result of the value axis, the audience can grade it from 1 to 5, 5 means it's very positive for platform construction, while 1 means it's negative for platform construction. The second question comes from the vertical axis of the value axis. If the value completely complies with the data subject's main goal, it will get a 5. Otherwise it will be graded as 1. The third and fourth question is to compare the value with other two stakeholder's main goal. If it also supports other two's main goal, then this value is of more importance. On the right side of the map, inspiring questions are raised to help the audience think more about this value conflict. Figure 5-10 shows a value dilemma scenario canvas.

With the value axis and value dilemma scenario, the audience can compare between different stakeholder's values, and find out which value is of more importance to the platform developing when value conflict happens.



◄

Figure 5-9: Example of using value axis of data subject



Value 1

Put a value card here

Platform construction

Stakeholder's basic goal

Conflicting with stakeholder B's basic goal?

□ Yes □ No

Conflicting with stakeholder C's basic goal?

□ Yes □ No

Value 2

Put a value card here

Platform construction

Stakeholder's basic goal

Conflicting with stakeholder B's basic goal?

□ Yes □ No

Conflicting with stakeholder C's basic goal?

□ Yes □ No

Inspiring quesions

1. What is the value conflict here?

2. Which stakeholder/value is more important in this senario?

3. How important is the value to each side?

4. ......

Value dilemma scenario

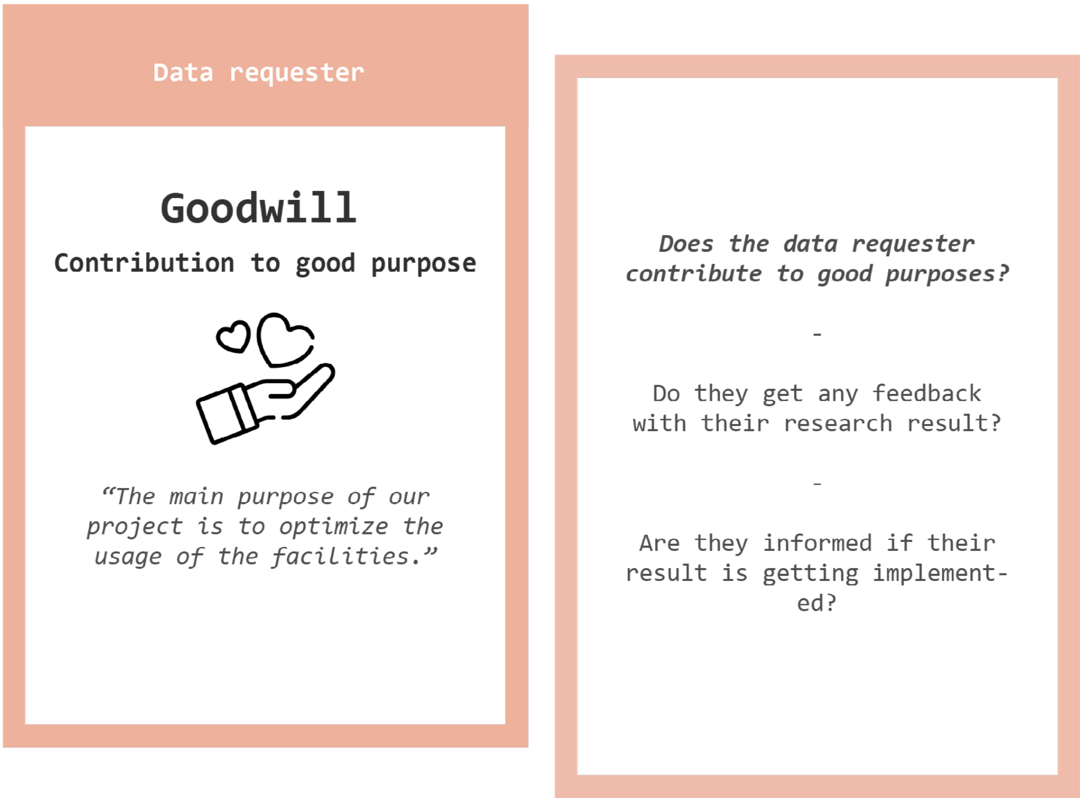
◄

Figure 5-10: Value dilemma scenario canvas

MVP 2

Based on the feedback of MVP 1, the goal of MVP 2 is no longer to solve the value conflicts by choosing one value over another, but to inspire the audience to transform the values into tangible features. The form of the card set is kept, and the other content including the context and stakeholder introduction are also embedded into the card set.

The most improvement in MVP 2 is to add an inspiration part which helps the designers and developers to transform the vague values into something functional. A list of triggering questions are added to each of the value card, asking if the value is fulfilled in such a way. Take data requester's value Goodwill (sub-value Contribution to Good Purpose) as an example (see figure 5-11). To help the developers measure if the researchers contribute to a good purpose, he can ask himself the last two questions on the back of the card: Do they get any feedback with their research results? Are they informed if their result is getting implemented? These two questions will trigger tangible functionalities to the platform (e.g. a feedback chanel in each of their project information page), and they can use the cards to reflect on their design, to see if they miss any values, or there are more features to add.



Data requester

Goodwill

Contribution to good purpose

"The main purpose of our project is to optimize the usage of the facilities."

Does the data requester contribute to good purposes?

-

Do they get any feedback with their research result?

-

Are they informed if their result is getting implemented?

▲

Figure 5-11: A data requester's value card in MVP 2

5.4 Evaluation of the concept

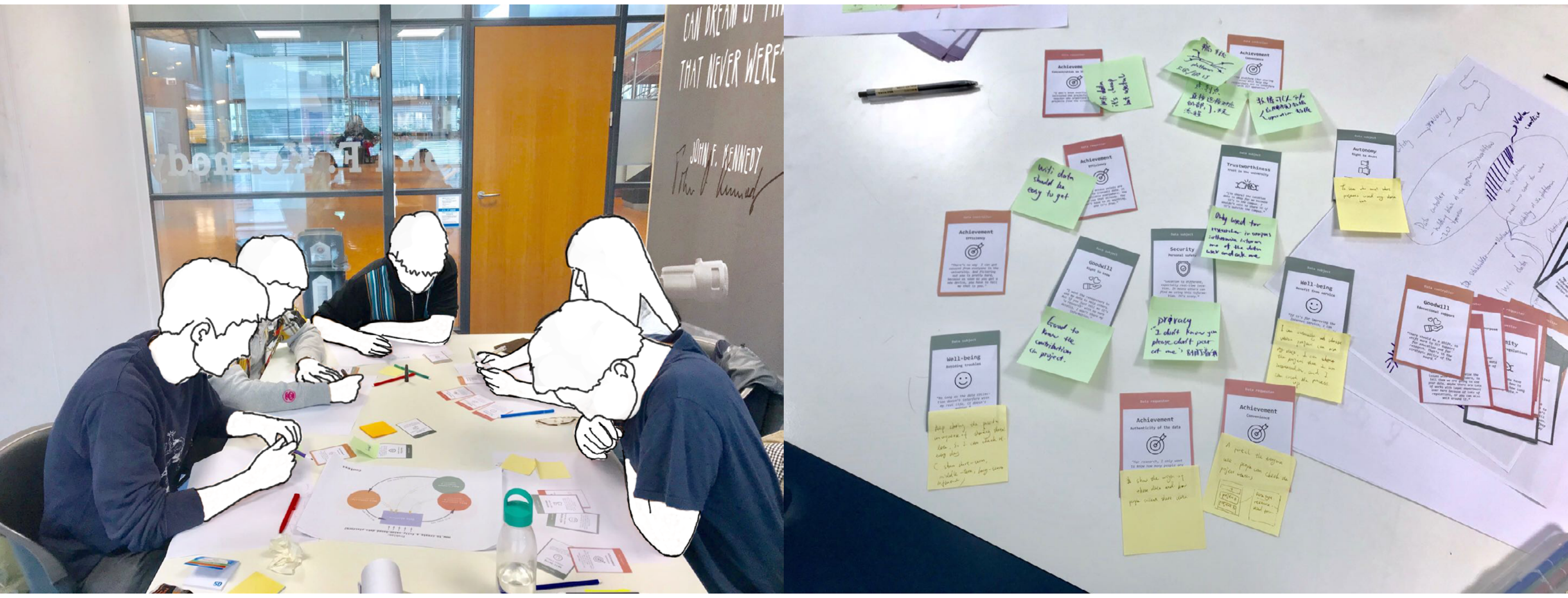
During the two rounds of MVP testing, evaluation of the MVPs are conducted. The evaluation session includes:

1. A group of 5 industrial design students using the MVP 1 cardset brainstorming what kind of feature could be implemented to the platform, inspired by the value themselves;

2. Four computer science students reviewing the content of the cardset, giving technical suggestions of improving the contents;

3. Two expert interviews, one with Andy Zaidman, a professor of software engineering in the Faculty of EEMCS. The cardset was shown to him, asking if the form and content fits into a normal software engineer's workflow. Another interview was back to Lolke Boonstra, since he has the vision of how will the platform be developed in the future. The cardset was shown to him as well, asking if the usage of the cardset fits with his vision of designing the functionality of the platform, and how will the cardset be used in any of the platform exploration research.

After the evaluation, the goal of the toolkit, the scenario of usage become clear, and the content of the cardset is also improved.



▲

Figure 5-12: Brainstorm session for evaluation

Goal of the toolkit

After the evaluation, a more clear goal of the toolkit is settled (see figure 5-13). The toolkit should consist of three parts, the first part shall introduce the context clearly and quickly, to help the audience get to the point as soon as possible. The second part should explain the values in a simple and systematic manner, instead of throwing everything together to the audience and confusing them. The third part should be an inspiration, that it helps to translate value into tangible features.

Scenario of usage

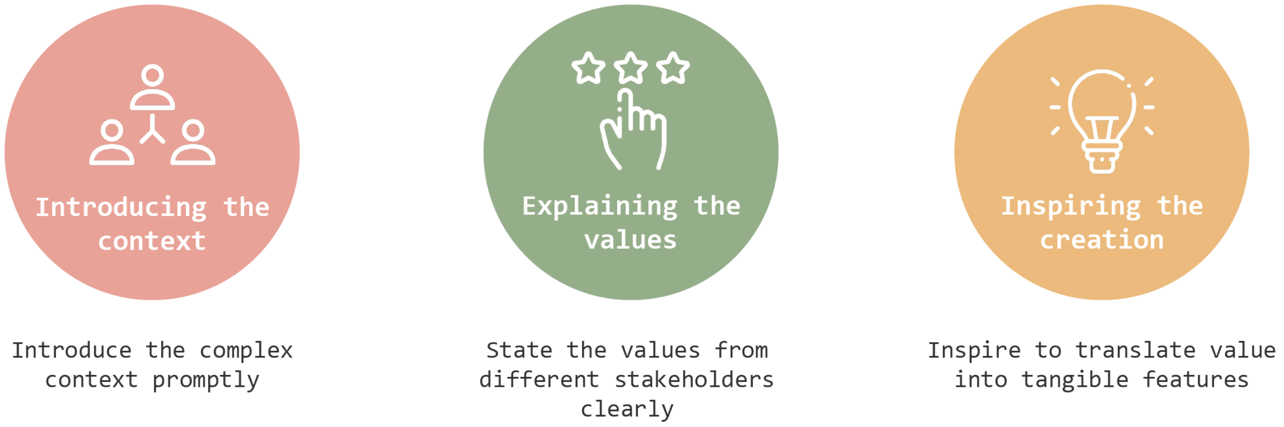
Even though the software engineers are more in favor of a digital form of a toolkit, they all mentioned that the form of a physical card set is a good choice during a group discussion. Considering the vision of the co-creation between different stakeholders, the using scenario of the card set becomes clear: it should be used during a group discussion or brainstorm session, where the questions on the back of the cards help them the diverge and come up with new ideas.

However, the idea of designing a self-explanatory card set is not very wise. Since there are too much background information in the context, it's too time and effort consuming for the users to read through everything before they start using the cards. Since the scenario of the usage will be a group discussion, it makes sense if there is a facilitator for this session, and there is an instruction for the facilitator to use the cardset.

Content of the cards

The setting of the inspiration cards is good, because it kind of gives a hint when people cannot come up with their own ideas. However, now there's too little inspiration cards, that some values are still missing its example of implementation.

Some of the questions are still too vague that it raises even more questions. It is expected that the questions could be asked in a more technical way, so the engineers are not confused.


Introducing the context
Explaining the values
Inspiring the creation
Introduce the complex context promptly
State the values from different stakeholders clearly
Inspire to translate value into tangible features

▲

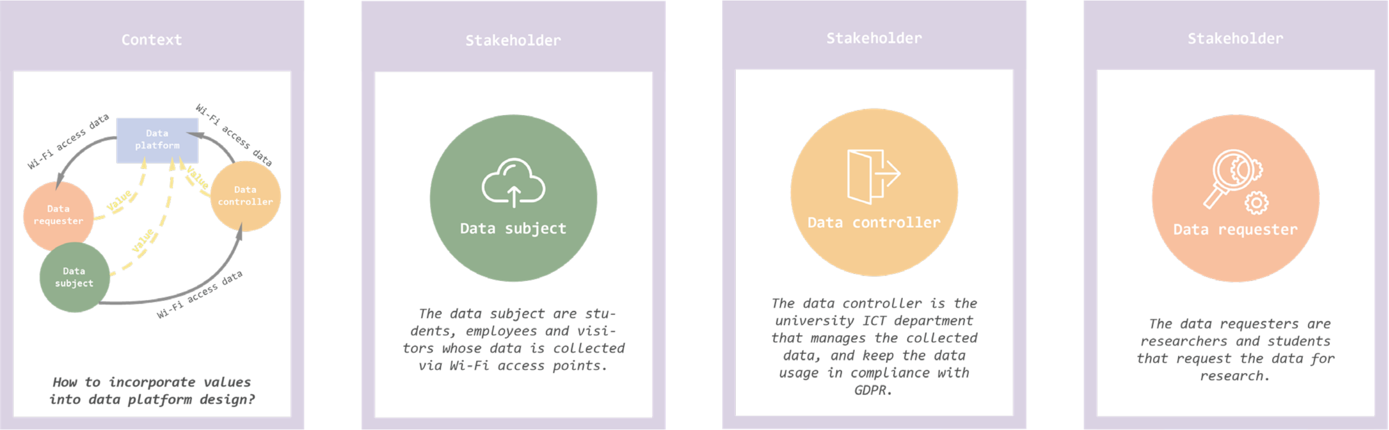
Figure 5-13: Goal of the final toolkit

5.5 Final design

Based on the feedback of the MVP 2, a final design of the toolkit was complete. To achieve the goal of the toolkit, there are three kinds of cards in this cardset.

The first kind of cards are designed to achieve the first goal, introducing the complex context. The context, the stakeholders and the relationships between them are shown on the 'Context Cards', so the audience can read through it and have an impression of the overview.

The second kind of cards are designed to achieve the second goal, explaining the values of different stakeholders. So the main part of the cardset is designed as 'Value Cards'. The form of the front side is still the same with MVP 2, while some of the quotes are rephrased, in order to avoid ambiguity. There are 4 questions on the back side, triggering the audience thinking of tangible functions. The first question is a 'yes or no' question, asking if the value on the front side is met or not. The second and the third question are closely related to the first question, transforming the first question into something measurable. The fourth question is a 'how' question, asking how to fulfill the value on the card. After answering the first three questions, the fourth question is quite open to answer, and the answer to it might be a feature to the platform.
If it's still hard to think about the answer to the fourth question, there is the third kind of cardset 'Inspiration Cards'. inspirations on this card sets comes from the brainstorm session and interviews before, giving an example of a technique solution to fulfill the value. The audience can either think about the solution on the inspiration card, or diverge from the cards to come up with other functionalities.



How to incorporate values into data platform design?

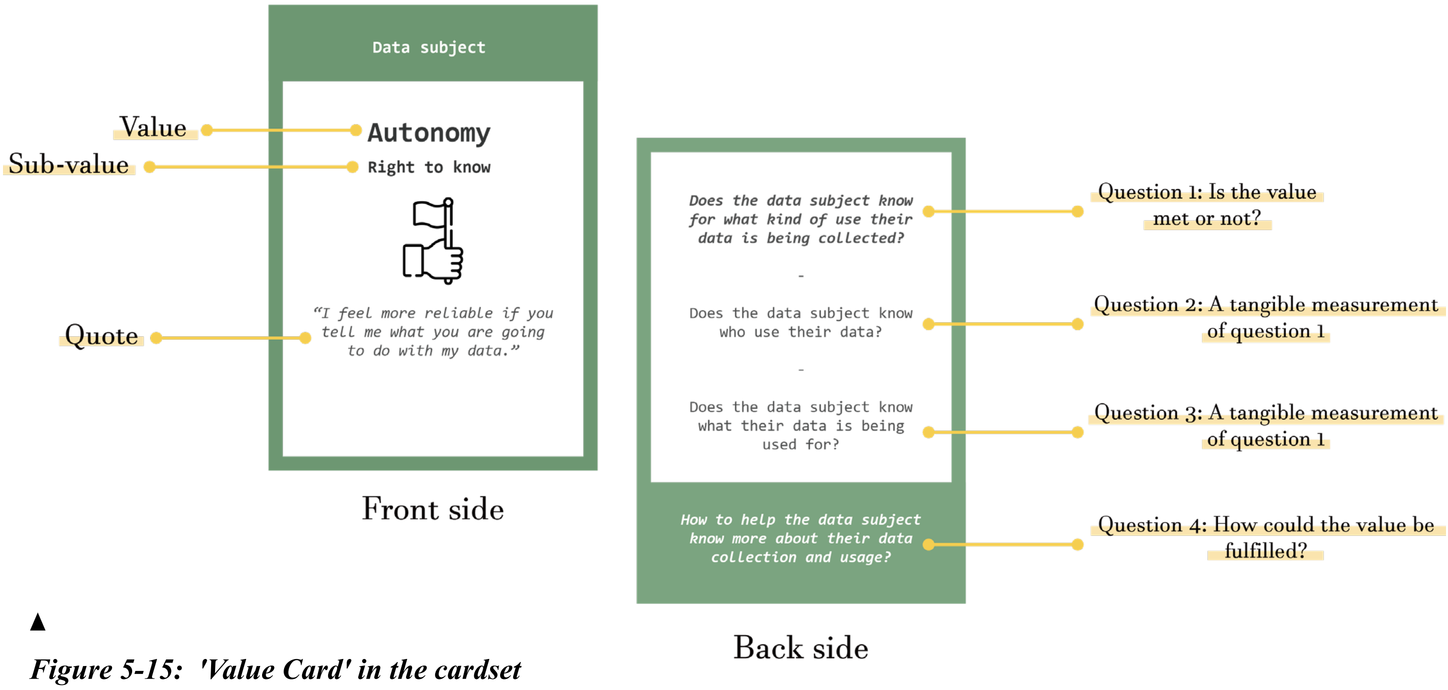
The data subject are students, employees and visitors whose data is collected via Wi-Fi access points.

The data controller is the university ICT department that manages the collected data, and keep the data usage in compliance with GDPR.

The data requesters are researchers and students that request the data for research.

▲

Figure 5-14: 'Context Card' in the cardset



"I feel more reliable if you tell me what you are going to do with my data."

Does the data subject know for what kind of use their data is being collected?

-

Does the data subject know who use their data?

-

Does the data subject know their data is being used for?

How to help the data subject know more about their data collection and usage?

Question 1: Is the value met or not?

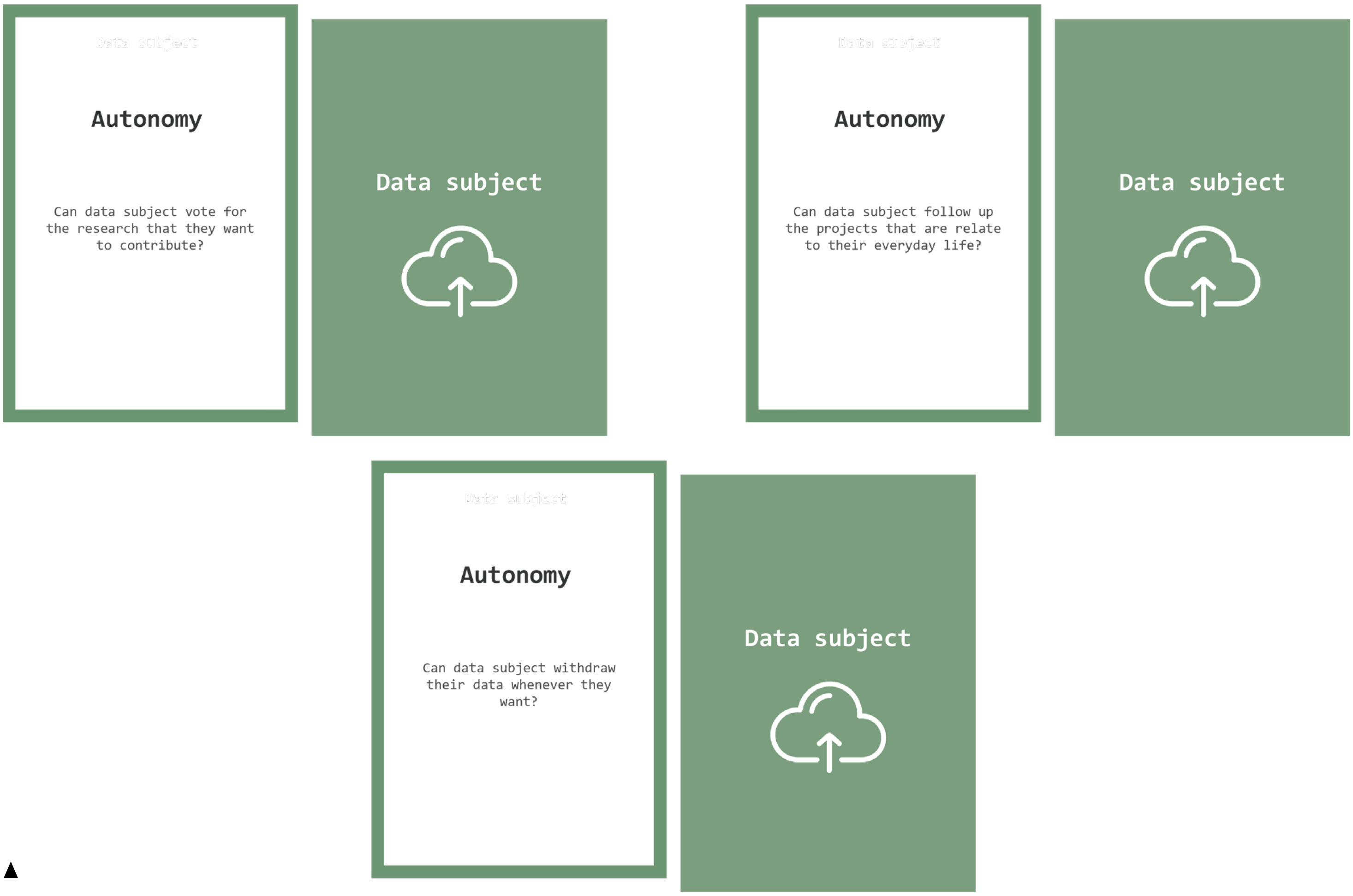
Question 2: A tangible measurement of question 1

Question 3: A tangible measurement of question 1

Question 4: How could the value be fulfilled?

▲

Figure 5-15: 'Value Card' in the cardset



Can data subject vote for the research that they want to contribute?

Can data subject follow up the projects that are relate to their everyday life?

Can data subject withdraw their data whenever they want?

▲

Figure 5-16: 'Inspiration Card' in the cardset

5.6 Scenario of the toolkit

Based on the feedback of the MVP 2, a This toolkit is designed to be used by the designers and developers, when in the very early stage of the platform development where they need to decide what kind of functionalities will be there in the platform. According to Lolke, there will be co-creation sessions between the ICT technical engineer and researchers, together deciding what kind of platform it will be like. This toolkit is designed to be used in such a context, where a group of people discussing and brainstorming the functionalities of the platform.

The users will firstly read the ‘Context Cards’, understanding the data flow and other background information. Then they will pick several ‘Value Cards’ and read them, thinking about how to incorporate such values into the platform. The questions on the back is expected to give them inspiration. If they are stuck with any value, they can turn to the ‘inspiration cards’. These cards will provide them with even more tangible features or related questions, so they can diverge their ideas based on these existing features.

To read through the cards in three steps will help them absorb, understand, inspire and diverge, and hopefully transform all those values into features, to measure or to improve them.

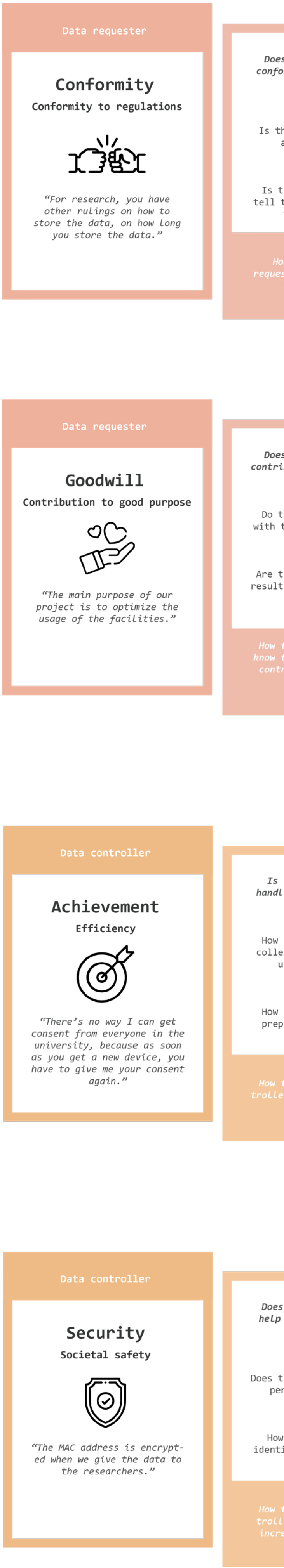

Data requester
Conformity
Conformity to regulations
“For research, you have other rulings on how to store the data, on how long you store the data.”
Data requester
Goodwill
Contribution to good purpose
“The main purpose of our project is to optimize the usage of the facilities.”
Data controller
Achievement
Efficiency
“There’s no way I can get consent from everyone in the university, because as soon as you get a new device, you have to give me your consent again.”
Data controller
Security
Societal safety
“The MAC address is encrypted when we give the data to the researchers.”

Figure 5-17: An overview of part of the value cards

s the data requester
rm to the regulations
well?

-

here any instructions
about the ruling?

-

here any feedback to
them if the conform to
the regulations?

w to help the data
ster better conform to
the regulations?

Data requester

Achievement

Concentration on the goal

“The teacher who organized the projects find the clients and data, I don’t have to think about it.”

Can the data requester concentrate on the research?

-

Is there anyone else help them to get the data?

-

How long does it take for them to get the data?

How to help the data requester concentrate on their goals instead of taking much effort to get the data?

Data controller

Achievement

Maintaining Internet service

“The ICT Department is collecting the data to provide Internet service. They moniter the number of users to provide enough accesss points.”

Does the platform help the data controller maintaing the Internet?

-

Does it have more flexible data storage limitation that helps maintainance?

-

Does it save time from preparing data sets for data requesters?

How to help the data controller maintain the Internet better?

s the data requester
bute to good purposes?

-

hey get any feedback
their research result?

-

hey informed if their
is getting implemented?

to let data requester
that their research is
ributing to good purpose?

Data requester

Well-being

Avoiding troubles

“We can tell the data subjects we are using their data and then do lots of works with legal department, or we can also walk around it becaue it’s grey area.”

Can data requester avoid troubles brought by the regulation?

-

Is there a legal way to walk around the problem?

-

How much procedure does it save if they avoid the problem?

Is there any ways to help data requester avoid troubles brought by the regulations lawfully?

Data controller

Goodwill

Educational support

“There is going to be a shift for ICT, from operation to ICT support for education and research”

Is the data controller doing more educational support?

-

How many data requester does the platform help to get research support?

-

How research are supported with this platform?

How to help the data controller support more for education and research?

the data controller
ing data efficiently?

-

long does it take to
ct data that could be
sed for research?

-

long does it take to
are the datasets for
data requesters?

to help the data con-
r handle the data more
efficiently?

Data subject

Well-being

Benefit from service

“If it’s for improving the Internet service, I can share my data becaue I really need it.”

Does the data subject benefit from any service by providing data?

-

Does their data directly improves the Internet maintainance?

-

Does their data contribute to other research that result in better campus service?

How to let the data subject know that they are benefiting from the service?

Data subject

Autonomy

Right to know

“I feel more reliable if you tell me what you are going to do with my data.”

Does the data subject know for what kind of use their data is being collected?

-

Does the data subject know who use their data?

-

Does the data subject know what their data is being used for?

How to help the data subject know more about their data collection and usage?

the data controller
to increase societal
safety

-

heir data contains any
rsonal information?

-

difficult it is to
ify a real person from
their data set?

to help the data con-
er handle the data to
ease societal safety?

Data subject

Security

Property safety

“My device information might reflect my financial situation and brand preference, I don’t want to share it.”

Does the data collection affect their property safety?

-

Does the data reflect their financial situation?

-

Does the data reflect their brand preference?

How to keep the data subject’s financial situation private?

Data subject

Security

Personal safety

“Real-time location means others can find me using this information. It’s scary.”

Does the data collection affect their personal safety?

-

Is it possible for a data subject to be identified?

-

Is it possible that others can find the data subject?

How to keep the data subject secure while using the data?

Chapter

6

Conclusion

This chapter discusses and concludes the whole project, and has a personal reflection

6.1 Discussion and conclusion

The goal of this research project is to build a basis for a value-based data platform in a multi-stakeholder context. The project is triggered by the enforcement of GDPR, which brings difficulties for data controllers and data requesters to handle data, which also triggers the development of the data platform, which will in the future manage the data in a legal way. But the aim of the project goes beyond GDPR, that apart from laws and regulations, ethics is also a vital factor that influences the design of the data platform.

The research starts with reviewing literature about laws and ethics. It would be easier to aim at a final data platform that is only GDPR-compliant, that means to develop a platform that only store lawful data for research purposes. But design has always been without a doubt an inherent moral job to do, that means to incorporate ethics into the platform is the designer's responsibility.

The research then explores the awareness of the three stakeholder entities towards personal data privacy, their current ways of dealing with data, and elicit their values from the insights. During the process, the current problems of data management, the interaction of the stakeholders and the hopes of a data platform are also found out.

During the whole process, the overview of the data platform keeps becoming more and more clear. At first, it's a simple platform that stores the data from the ICT Department and transmit data to the researchers, and it seems that the data requesters are the only users who benefit from. Then from the review of GDPR, it has research as a legal ground to base on, so it doesn't have to stick to issues like informed consent and complete anonymization. After the literature review of the design ethics, data subject becomes an indirect stakeholder of the data platform, that their values should be considered even though they are not the target users. Finally with the research of different stakeholders, the goal of the platform also becomes more complex: to help the data controller manage the data without interference and lawfully; to help the data requester get the data more efficiently and get instructions on how to use the data lawfully; and to help the data subjects have more control over their data.

As an end result, a cardset is designed to communicate the values and inspire the designers and developers of the data platform. The cardset has two kinds of cards: value cards and inspiration cards. The value cards shows the values of different stakeholders, while the inspiration cards gives examples of how to fulfill those values.

The design of the cardset is not only about communication, since the value itself is a vague concept. So another aim of the cardset is to help the developers measure whether the value on the card is met, and how to transform the values into tangible functionalities. The card set inspires the audience by asking them inspiring questions about how could the values be fulfilled. Besides only raising questions, the cardset also provides possible features as examples to fulfill the values. The audience could already use this inspiration cards to develop such a feature, or they could also use it to come up with more ideas.

To evaluate if the cardset really helps to complete an ethical design, we can go back to Aral Balkan and Laura Kalbag's Ethical Design model. First the platform will be GDPR-compliant, it keeps the data private and secure, yet accessible and open, so it fits the first layer Human Rights. Then the features will be developed considering its main target user, data requesters' needs, that is to be functional, convenient, and reliable so they don't need to worry whether they can use the data or not, that fits the second layer Human Effort. Finally it take all the stakeholders' values into consideration, it tries to do good for all the stakeholders, so it fits the third layer Human Experience. Thus we could say, the design of the cardset helps to develop a value-based platform, which is an ethical design.

In conclusion, the research project defines the responsibility of a designer to complete ethical design, explore the values of the stakeholders in the context, and help to incorporate those values into the future data platform with the design of a cardset as toolkit for its audience.

There are still some limitations of this project. First, the research with the data subjects is still limited. Only students took part in the research, and their insights cannot represent the employees' insights considering their knowledge and experience. Second, the content of the cardset still need to be polished with professional technical people. The features given as examples are still limited due to my field of study, and the questions that inspire certain functionalities should also be further discussed with professional computer scientists. Finally, the design of the cardset hasn't been evaluated with its real users in the envisioned scenario, which is to be used in a multi-stakeholder co-creation session.

6.2 Personal reflection

There are lots of thinking alongside this research project. The first good thing about the project is the great atmosphere of research in the TU Delft. Sometimes when I read related literature, I found out that the authors are the teachers in this university and I can directly send emails to them to ask questions. Their works also fit perfectly with this research setting, and I can feel the happiness of doing research in such a top notch university.

Since I have little research knowledge, I also learnt a lot through this project. For example, to use a design theory (VSD) throughout the whole project, from constructing questionnaires and interview questions, to analyse the result and apply it into design. Also there are lots of small things to mention when learning how to do research: how to structure a real research report, how to do a literature review, and how to analyse the results and incorporate it into a design.

But there are more things that I have to reflect which leads to the incompleteness of the project. The first thing is bad time management, which took me longer to reach the mid-term meeting, and finally result in not finishing the design of the toolkit. The second thing is lack of proactiveness, which results in lack of communication with my chair and mentor, and lead to a bad result in the last phase of the project, that I found out my focus on the GDPR is completely out of the direction. These two things could be avoided with more detailed planning and sticking to the schedule, but I failed to do so.

I also find my lack of experience in managing such a big project on my own. Because I spend more time struggling with the report, sometimes I feel like I'm doing two different projects at the same time: designing a toolkit and writing a research report. The report need me to converge all the findings and write them down in a clear manner, while designing the toolkit need me to diverge to different possibilities. There are a lot of times I found myself doing these two things in a parallel way, which in fact they should complete one another.

Also to understand the context took me more effort than I expected. As mentioned in the discussion, the picture of the data platform keeps changing alongside the research. It's hard to grasp all the elements in the context, for they are all vague to me: a regulation, ethics, values, a future platform, and a toolkit as an inspiration. I'm really happy that finally I did not only try to understand them (a little bit), but also tried to transform my explorations and my understandings to something more feasible to other people.

The final cardset is not a complete design, for it still needs to be polished with its content, to be finally presented to its target audience, that are the developers and researchers. I would want to keep working on it to complete it after graduation.

To conclude, the whole project is a meaningful learning experience to me, for conducting research in an unfamiliar field, and apply my learnings and skills in the last two years. But there are so many things for me as a designer to keep improving, so many drawbacks to avoid in my future projects.

References

1. Bourgeois, Jacky & Kortuem, Gerd & Kawsar, Fahim. (2018). Trusted and GDPR-Compliant Research with the Internet of Things. 10.1145/3277593.3277604.

2. B. Bonné, A. Barzan, P. Quax and W. Lamotte, "WiFiPi: Involuntary tracking of visitors at mass events," 2013 IEEE 14th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM), Madrid, 2013, pp. 1-6.
doi: 10.1109/WoWMoM.2013.6583443

3. Antonin Danalet, Loïc Tinguely, Matthieu de Lapparent, Michel Bierlaire, Location choice with longitudinal WiFi data, Journal of Choice Modelling, Volume 18, 2016, Pages 1-17, ISSN 1755-5345, https://doi.org/10.1016/j.jocm.2016.04.003.

4. Alessandro E.C. Redondi, Matteo Cesana, Building up knowledge through passive WiFi probes, Computer Communications, Volume 117, 2018, Pages 1-12, ISSN 0140-3664, https://doi.org/10.1016/j.comcom.2017.12.012.

5. M. Zhu et al., "A measurement study of a campus Wi-Fi network with mixed handheld and non-handheld traffic," 2015 IEEE 28th Canadian Conference on Electrical and Computer Engineering (CCECE), Halifax, NS, 2015, pp. 848-853. doi: 10.1109/CCECE.2015.7129385

6. Kalogianni, Eftychia & Sileryte, Rusne & Lam, Marco & Zhou, Kaixuan & Ham, Martijn & Verbree, Edward & van der SPEK, Stefan. (2015). Passive WiFi Monitoring of the Rhythm of the campus.

7. Dutch DPA investigates WiFi tracking in and around shops. (n.d.). Retrieved from https://autoriteitpersoonsgegevens.nl/en/news/dutch-dpa-investigates-wifi-tracking-and-around-shops

8. Batya Friedman, Peter H. Kahn, Jr., Jennifer Hagman, Rachel L. Severson, and Brian Gill. The watcher and the watched: Social judgments about privacy in a public place. Human–Computer Interaction, 21(2):235– 272, May 2006. URL http://www.tandfonline.com/doi/abs/10.1207/ s15327051hci2102_3.

9. Friedman, P.H. Kahn and A. Borning. Value Sensitive Design: Theory and Methods. University of Washington, 2002

10. Friedman, Batya; Kahn, Peter H.; Borning, Alan; Huldtgren, Alina (2013), Doorn, Neelke; Schuurbiers, Daan; van de Poel, Ibo; Gorman, Michael E. (eds.), "Value Sensitive Design and Information Systems", Early engagement and new

technologies: Opening up the laboratory, Philosophy of Engineering and Technology, Springer Netherlands, pp. 55–95, doi:10.1007/978-94-007-7844-3_4, ISBN 9789400778443, retrieved 2019-09-03

11. https://2017.ind.ie/ethical-design/

12. Knobel, Cory P., Bowker, Geoffrey C.. 2011. “Values in Design.” Communications of the ACM 54 (7): 26–28. https://cacm.acm.org/magazines/2011/7/109899-values-in-design/fulltext

13. Verbeek, 2006. Materializing Morality, Design Ethics and Technological Mediation. https://journals-sagepub-com.tudelft.idm.oclc.org/doi/pdf/10.1177/0162243905285847

14. Akrich, M. 1992. The de-scription of technological objects. In Shaping technology/building society, ed. W. E. Bijker and J. Law, 205-24. Cambridge: MIT Press.

15. Latour, B. 1992. Where are the missing masses? The sociology of a few mundane artifacts. In Shaping technology/building society, ed. W. E. Bijker and J. Law, 225-58. Cambridge: MIT Press

16. Heidegger, M. 1927. Sein und Zeit. Tübingen, Germany: Max Niemeyer Verlag.

17. Ihde, D. 1990. Technology and the lifeworld. Bloomington: Indiana University Press.

18. C.Y. Baldwin, K.B. Clark. Design Rules: The Power of Modularity, vol. 1, The MIT Press, Cambridge, MA (2000)

19. Johnson, D. G. & T. M. Power, 2005. “Computer systems and responsibility: A normative look at technological complexity,” Ethics and Information Technology, 7: 99–107.

20. Noorman, M. (2018, February 16). Computing and Moral Responsibility. Retrieved from https://plato.stanford.edu/entries/computing-responsibility/#RetConMorRes

21. Burg, S. van der, & Gorp, A. van. (1970, January 1). Simone van der Burg & Anke van Gorp, Understanding moral responsibility in the design of trailers. Retrieved from https://philpapers.org/rec/VANUMR

22. Langdon Winner, Do Artifacts Have Politics?, 109 Daedalus 121, 123 (1980); see also Caro, supra note 1, at 318.

23. https://www.merriam-webster.com/

24. https://2017.ind.ie/ethical-design/

25. Friedman, B., Kahn, P. H., & Borning, A. (2009, January 26). Value Sensitive Design and Information Systems. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470281819.ch4

26. Oluwatosin, A. (2017, July 15). Difference Between Ethics and Values.

Retrieved from https://keydifferences.com/difference-between-ethics-and-values.html

27. Alexei Czeskis, Ivayla Dermendjieva, Hussein Yapit, Alan Borning, Batya Friedman, Brian Gill, and Tadayoshi Kohno. Parenting from the pocket: Value tensions and technical directions for secure and private parent-teen mobile safety. In Proceedings of the Sixth Symposium on Usable Privacy and Security, SOUPS ’10, pages 15:1–15:15, New York, NY, 2010. ACM. URL http://doi.acm.org/10.1145/1837110.1837130.

28. Joerges, B. (1999). Do politics have artefacts? Social Studies of Science, 29(3), 411-431.

29. https://www.apple.com/lae/researchkit/

30. Ouellette, J. (2016, July 26). Apple's Health Experiment Is Riddled With Privacy Problems [UPDATED]. Retrieved from https://gizmodo.com/apple-s-health-experiment-is-riddled-with-privacy-probl-1783878924

31. Mostert, M., Bredenoord, A. L., Biesaart, M. C., & van Delden, J. J. (2016). Big Data in medical research and EU data protection law: challenges to the consent or anonymise approach. European journal of human genetics : EJHG, 24(7), 956–960. doi:10.1038/ejhg.2015.239

32. Apple's ResearchKit frees medical research. Nat Biotechnol 2015; 33: 322.

33. Costa FF: Social networks, web-based tools and diseases: implications for biomedical research. Drug Discov Today 2013; 18: 272–281.

34. Rodriguez LL, Brooks LD, Greenberg JH, Green ED: Research ethics. The complexities of genomic identifiability. Science 2013; 339: 275–276.
Sethi N: The promotion of data sharing in pharmacoepidemiology. Eur J Health Law 2014; 21: 271–296.

35. Greely HT: The uneasy ethical and legal underpinnings of large-scale genomic biobanks. Annu Rev Genomics Hum Genet 2007; 8: 343–364.

36. O'Brien SJ: Stewardship of human biospecimens, DNA, genotype, and clinical data in the GWAS era. Annu Rev Genomics Hum Genet 2009; 10: 193–209.

37. McCormick TH, Lee H, Cesare N, et al. . Using Twitter for demographic and social science research: tools for data collection and processing. Sociol Methods Res . 2015;46(3):390–421.

38. J.Blumenstock, G.Cadamuro, R.On, Predicting poverty and wealth from mobile phone metadata. Science 350, 1073–1076 (2015). doi:10.1126/science.aac4420pmid:26612950

39. Tang PC, Lansky D. The missing link: bridging the patient-provider health information gap. Health Aff (Millwood) 2005 Sep;24(5):1290–5. doi: 10.1377/hlthaff.24.5.1290.

40. Bastian Greshake Tzovaras, Kevin

Arvai, Mairi Dulaney, Vero Estrada-Galiñanes, Beau Gunderson, Tim Head, Dana Lewis, Oded Nov, Orit Shaer, Jason Bobe, Mad Price Ball: Open Humans: A platform for participant-centered research and personal data exploration. bioRxiv 469189; doi: https://doi.org/10.1101/469189

41. Schwartz, Shalom H.; Cieciuch, Jan; Vecchione, Michele; et al. (October 2012). "Refining the theory of basic individual values". Journal of Personality and Social Psychology. 103 (4): 663–688. doi:10.1037/a0029393