Key Tension Points and Design Guidelines for GDPR Compliance: Designing for a News Service Application

MELIS BURT KUT
Key Tension Points and Design Guidelines for GDPR Compliance: Designing for a News Service Application

Author
Melis Burt Kut, melissab@kth.se

Media Technology, Master of Science in Computer Science and Engineering

Supervisor: Karey Helms, KTH, School of Electrical Engineering and Computer Science, Department of Media Technology and Interaction Design.

Examiner: Roberto Bresin, KTH, School of Electrical Engineering and Computer Science, Department of Media Technology and Interaction Design.

Date of submission: 2018-06-13
ABSTRACT
Digitization poses a threat to the fundamental rights of individuals’ personal sphere. This is due to deficiency within the current bylaws to protect data subjects’ privacy and the lack of social codes for handling privacy in the virtual space. Colossal amount of implicit data processing, takes away data subject’s control over their personal data. In order to protect data subjects from this treacherous relationship, between stakeholders and data subjects, the European Union has issued the new General Data Protection Regulation that was enforced in May 2018. Companies operating within EU thereby face substantive legislative reform in data protection. However, there are no current guidelines for how to acclimatize to the new regulation of processing personal data, especially for subsidiary companies. This study therefore addresses this gap by detailing the design process of attaining GDPR compliance for a subsidiary news service application. From this process, nine key tension points were identified and reformulated into five design guidelines more broadly applicable to design for privacy. In addition, two boundary objects and a transparency-layer strategy were formulated.

SAMMANFATTNING
Key Tension Points and Design Guidelines for GDPR Compliance: Designing for a News Service Application

Melis Burt Kut
KTH Royal Institute of Technology
Stockholm, Sweden
melissab@kth.se

ABSTRACT
Digitization poses a threat to the fundamental rights of individuals’ personal sphere. This is due to deficiency within the current bylaws to protect data subjects’ privacy and the lack of social codes for handling privacy in the virtual space. Colossal amount of implicit data processing, takes away data subject’s control over their personal data. In order to protect data subjects from this treacherous relationship, between stakeholders and data subjects, the European Union has issued the new General Data Protection Regulation that was enforced in May 2018. Companies operating within EU thereby face substantive legislative reform in data protection. However, there are no current guidelines for how to acclimatize to the new regulation of processing personal data, especially for subsidiary companies. This study therefore addresses this gap by detailing the design process of attaining GDPR compliance for a subsidiary news service application. From this process, nine key tension points were identified and reformulated into five design guidelines more broadly applicable to design for privacy. In addition, two boundary objects and a transparency-layer strategy were formulated.

Author Keywords
GDPR; Integrity through design; Privacy; Transparency; Right to Information; Designing with data

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI).

INTRODUCTION
Every year, the amount of data we produce doubles [15]. Within only a minute, hundreds of thousands of Google searches and Facebook posts are produced, which are subsequently processed and stored within a database, each data breadcrumb put together creating a coherent whole of who we are, what we feel and think [15]. Data can be shared intentionally, when we for example intentionally post on social media platforms, and unintentionally, such as traces of personal data and metadata that are collected actively and passively [5]. Big data and advanced Information Technology (IT) can store and process several exabyte of data [33]. This could be a threat to individual’s capabilities to protect their personal sphere of life and their control of privacy. Privacy online is important, in order to allow data subjects to be able to determine for themselves what personal information to share and to whom. Even though the rise of IT has enabled and still enables excessive possibilities, privacy ‘has become ‘a casualty of progress driven by IT’” [31], creating a personal integrity conundrum. As new technologies arise with increasing power, the clarity and agreement on privacy fades [33], resulting in a treacherous relationship.

The current European Union Protection Directive sets the fundamental rights to privacy in Europe [20]. Though established in 1995, it is not as self-evident as it might have been over 20 years ago, making the directive outdated and inapplicable to today’s technically developed society. In an attempt to modernize the current Data Protection Directive 95/46/EC, the European Parliament, European Commission and European Council have proposed a new directive called the General Data Protection Regulation (GDPR) [34]. Thereby, as of May 25th, 2018, in attempt to standardize data privacy laws across Europe and protect the privacy of citizens, GDPR will be going into effect [34].

The following three regulations within the GDPR will allow users to take back the control over their data. The new regulation gives data subjects the Right to be Informed, Right to Access and the Right to be Forgotten [34]. That is, individuals have the right to have knowledge and access of what specific personal data is being processed, for what purpose and where this data is being stored, respectively, the right to ask for erasure [35]. Corporates are required to inform data subjects of their rights. This information is a transparency requirement under the GDPR and is required to be “concise, transparent, intelligible” and “easily accessible” using a “clear and plain language” [35,36].

Alongside these actions, to protect data subjects, corporations face multiple challenges. Apart from that breaches can be met with high fines, [37] there is also the larger implication of losing users’ trust for a company. This requires a corporate responsibility to protect users instead of exploiting them. To meet policy requirements, companies
need to adopt new courses of actions. These new regulations will compel IT-systems within companies to make comprehensive modifications [9]. Corporates require having better understanding of their data processes, as this can be a complex myriad of data and collaborations. Currently, there are no practical mechanisms or frameworks for organizations to support data subjects to inquire a detailed view of what personal data is withheld and why [12]. There is also little debate, within the human-computer interaction (HCI) field, on the subject of how users will react to the new legislation and manage privacy [1].

These challenges will not only affect the parent companies, but also its subsidiary, as they are inextricably linked, creating an intricate unravelment to be GDPR compliant. If a subsidiary runs afoul of the GDPR, the parent company will also be affected. Thereby, the findings from this study will be useful for other similar subsidiary companies. This study will use a news service that is part of a bigger ecosystem, with the business-model 'content targeted advertising'. Within a bigger company ecosystem such companies face the challenge of informing their users without having to compromise their business model, nor the user experience. Thus this study is also of interest for interaction design and user experience researchers challenged with designing for and around GDPR legislations.

Research Question
The main purpose of this research is to understand how current data management and systems work within a subsidiary news service in Sweden, in order to further apprehend local processing activities and thereby be able to answer the following research question: What key tension points should be acknowledged when designing privacy policies for a subsidiary news service in Sweden, in order to meet requirements set by the General Data Protection Regulation?

This paper defines a key tension point (KTP) as a segment part of a bigger problem. Highlighting different KTPs allows to fragmentize a concept of a specific problem into smaller interpretations of that problem. By reconceptualizing the individual ideas used to form a coherent whole of the problem, allows more insights to be gained.

In the remainder of this paper, the meaning of personal data and privacy will be presented, as well as an explanation of why GDPR is necessary. Next, Research through Design as a method and UX Lean process is described. Then in the results, the contributions will reveal nine tension points and jointly end with a discussion on each tension point.

BACKGROUND
This section will first define what is meant by data. Then a closer review of personal data and privacy will be provided, as this work lies in the intersection of GDPR and corporations.

A data entity is “a piece of information” that is abstract and intangible [29]. Data entities on their own are merely trivial numbers, letters, characters and symbols, with no meaning, or in the case of visualizations, they are simply pixels or “pretty pictures, abstract art” [29]. Metadata, is “data about data”, which includes, among other things, the context of the individual datum and its origin. Without metadata, data would be meaningless [29]. “But data is not information.”, Drucker writes [8]. “Information is data endowed with relevance and purpose.” Data takes on meaning when it is processed, interpreted and provided with a context. It is only then data becomes information [38]. In the words of Rosenstein, “data are the kernels of what eventually may become knowledge, but require increasing levels of understanding as they are first transformed into information. Once information progresses further and is put to use, it then becomes knowledge” [24]. Disclosing this continuum from data to information to knowledge, enables one to grasp how opaque processing of trivial mundane data entities put together, can lead to an infringement of personal integrity and individual freedom.

Personal Data
Defining personal data can be considerably troublesome, where any information or data that is either directly linked to, or can be linked to, an individual person, is regarded as personal data, making the definition vague and ambiguous [33]. Manders-Huits and Van den Hoven coined the term “identity-relevant information” as a definition of personal information [7]. They further distinguish between referential and attributive data, i.e. data that directly refers to a specific person, respectively, data that describes a characteristics of a person [22]. Even though attributive data does not directly refer to one specific person, the aggregation of attributive data makes it possible to identify the person [22]. For this reason, the GDPR defines personal data as; “any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier” [39]. An example of an identifier for a news service application could be an identification number of the data subject, topics that the data subject follows or the data subject’s personal settings.

Violation of Personal Data
Bellotti expresses concerns for how connectivity and sharing of personal data poses an unethical threat to privacy due to the risks that emerge from the processing of personal computing [1]. Users have little understanding for how their personal data is dealt with; where this data is stored and
shared. The lack of control puts the user into a vulnerable position, where personal data can easily be leaked or taken advantage of, without the user’s awareness. According to the Data Loss Archive and Database (DLDOS), “more than 300 million personal records were exposed from more than 900 reported breaches in the US” [21].

More recently, (March, 2018), the whistleblower, Chris Wiley, exposed how the data analytics firm Cambridge Analytica, harvested private information from more than 50 million individuals without their consent or awareness. Private social media activity of users were exploited from their Facebook profiles, in order to build a system that could profile and target individual US voters, with personalized political advertisements [4]. Wiley expresses his concerns of how it is possible to change an individual’s perception of something by harvesting personal data, analyzing that data to build a psychological profile and defining what kind of messaging the user is susceptible to, in order to target the individual in ways that the user cannot see or understand, resulting in a distorted perception of reality [4].

Users of a news service application could hold of similar concerns, as a news service application is a similar space that holds of the possibilities to affect users opinions by decreasing information diversity and thereby creating a so called filter bubble [2].

Privacy
Privacy can be segmented into the following four categories; information privacy, social privacy, psychological privacy and physical privacy [14, 3, 6, 26]. This study will focus on Information privacy. Information privacy can be defined as “the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others” [28]. Hughes argues in line with the abovementioned; “Privacy is the power to selectively reveal oneself to the world” and further defines privacy by distinguishing it from secrecy; “Privacy is not secrecy. A private matter is something one doesn’t want the whole world to know, but a secret matter is something one doesn’t want anybody to know” [18].

Privacy in the Virtual Space
In a shared physical space we act accordingly to well established set of social protocols that include normal cues of a clear distinction between the public and private space. Similarly, when dealing with privacy in the physical space, we refer to these protocols, in order to act appropriately. However, the virtual space lacks such well-recognized social protocols as there are no established unwritten social codes between the public and private space [17]. This, in turn, can result in chaos of invasion and violation of the personal sphere. Furthermore, the definition of privacy in the virtual space is undefined. Privacy is usually interpreted as to be synonymous with data protection, which is partially misleading as privacy does not entail to specifically protect data but rather the personal sphere that is represented by the data and their “relations associated with a person”, referred to as the data shadow of a person’s privacy [10].

The Privacy Paradox
The right to privacy is a general human right. It is the right for a person to keep a domain around them and choose which parts in this domain that can be accessed by others [30]. However, the balance between privacy and disclosure can entail a sacrificing process involving a delicate and problematic balance between for example social liberty versus alienation. To consider absolute privacy would imply alienation from society and other human beings, according to Gavison [11]. However, the opposite would mean a total loss of privacy [11]. Therefore, each individual is, according to Westin, “continually engaged in a personal adjustment process in which he balances the desire for privacy with the desire for disclosure and communication of himself to others, in light of the environmental conditions and social norms set by the society in which he lives” [28]. The incongruity between “online self-disclosure behavior” and privacy concerns is known as the privacy paradox [14]. According to the privacy calculus theory, [14] the intention to disclose personal information is based on a risk-benefit analysis, where the individual's final behavior is determined by the privacy trade-off between, resulting costs and benefits [14]. Further developing this theory could lead to the argument that privacy can be seen as a commodity that can be exchanged for expected benefits [14].

Cognitive Deficiency Theory
In order to fairly engage in a risk-benefit analysis, the individual needs to know what the privacy trade-offs are between the resulting costs and benefits. Legal documents such as privacy policies are usually verbose; creating information overload, and difficult to understand; containing overly legalistic, technical and specialist language and terminologies [25]. Privacy policies are therefore usually unread, which results in users not being aware of how their personal data is processed [23]. Users are thereby left with an impaired judgment, [25] making it close to impossible to equally measure their privacy trade-offs. Due to the lack of knowledge of how to take precautions regarding personal information online, the cognitive deficiency theory argues that privacy behavior does not reflect the individual’s privacy concerns [14].

GDPR purports to help enhance transparency in order to allow data subjects to easily assess the level of data...
protection of their personal data and to be able to make decisions based on a fair and transparent, risk-benefit analysis.

**METHOD**

* A general overview of the main methods used will be provided below.

**Research through Design**

This research paper employs a qualitative approach by way of Research through Design (RtD). RtD is a method that allows the design process in itself to drive exploration of both problems and solutions. In this way, new knowledge is acquired via the act of making and embedded within resulting design processes and artifacts [32,16]. By means of the construction of artifacts, such as prototypes, RtD allows researchers to explore new materials (*in this study being GDPR*). Thereby the researcher can codify their own understandings and frame problems, allowing the created artifacts to become embodiments of the possible future, instead of the present and past [19]. This paper, therefore adopts RtD, in preference to other methods.

First an understanding for data streams was acquired in order to formulate a design space. Thereafter, a lean UX process was adopted to determine the pace within the method. Thereby, an understanding for how to address users with PAs of their personal data was identified. The resulting KTPs emerged through the entire design process.

**Data Inventory**

In order to get a better understanding of implicit data streams and what local PAs that the news service application necessitates informing their users of, a data inventory was conducted. This was done by a mapping exercise within excel, which served as a basis for further iteration, that was done through several meetings with a legal advisory. Furthermore, a category mapping of the PAs was carried out, based on the official PA document, in order to serve basis for the first of two Design Studios.

**Interview with a back-end programmer**

Alongside documenting the PAs, a semi-structured interview was carried out with one of the back-end programmers. The interview also included examination of existing code and architecture. This was done to be able to further understand how current data management and systems work and thereby visualize implicit data life cycles.

**Lean UX Process**

A Lean UX process was adopted, which was inspired from the book; *Lean UX, Designing Great Products with Agile Teams* by Jeff Gothelf and Josh Seiden [13]. The advantage of adopting a Lean UX cycle is the possibility to remove the “I’m not sure if this is a good idea” and the potential feature design debate away from the design process [13]. This is accomplished by working in rapid, iterative cycles where data generated during the process can be used in each iteration. Thereby, Lean UX allows feedback to be obtained as early as possible, in the process, in order to acquire quick results and hence, make quick decisions.

**UNDERSTANDING THE DESIGN SPACE**

The following provides the processes of how the design space was approached, along with the results of these actions.

**Processing Activities**

Each controller must maintain a record of Processing Activities. A Processing Activity (PA) is any activity that handles an individual’s data [40]. The record shall contain the points listed below.

**Data Inventory**

As data processing is a complex myriad, with a lot of dependencies, it is of interest for the company to document product PAs, in order to understand data streams and management. In this way, companies can get an understanding for what actions are needed to be taken, in order to be GDPR compliant. The results from the PA exercise played a significant role for creating the designs (i.e. the Minimal Viable Products; MVPs), as it is infeasible to create a privacy policy without an understanding of what and how to inform data subjects of the processing of their data. For each PA, it’s legal basis was defined, in order to know when and how to ask for consent or provide opt-outs. The data inventory resulted in an excel file where each PA was listed. For each PA, the following points were specifically clarified.

- **On what platform is the PA taking place?** (Ex. Apps, Web)
- **Using what assets? Where it is being stored and whom is it is being shared with?** (Ex. Google Analytics, Mixpanel)
- **Its purpose category?** (Ex. Customer Care, Product Development, Necessary Functionality)
- **What precise personal data is being collected?** (Ex. Push Tokens, Behavioral Data, Email)
- **How is the data being collected?** (Ex. Entered by user or tracking)
- **The Legal basis of the processing?** (Ex. Consent, Contractual Obligation, Legitimate Interest)

This further resulted with a additional understanding for two key issues that the company needed to address, which required a more technical approach within the user interface of the privacy policy. These two main issues held of;
Control of Privacy and Data Take-out/The Right to be Forgotten.

Key Tension Point 1: Control of Privacy
Data subjects have the right to have full control over their personal data. This is a key tension point (KTP) that all companies need to be aware of, in order to be GDPR compliant. As the news service is a subsidiary, i.e. part of a bigger conglomerate, the news service helps other products within the conglomerate by collecting user behavior and thereby allowing targeting advertisement on other product platforms that are part of the conglomerate. This resulted in the following questions of how to treat control options from an end-user perspective; 1) if data subjects were to opt-out from an activity on the news service application, would that imply that they would also be opting-out from the conglomerate? Also, 2) would it be possible to choose to opt-out from the conglomerate’s data collection system but not the subsidiary’s? Similar observations were made in regards to giving consent; if data subjects gave their consent for a data processing on the news service application, would this mean that they automatically gave their consent to the whole ecosystem to collect their data?

As the conglomerate also necessitates to be GDPR compliant, they are also required to provide a privacy control page. That being said, the challenge of how the privacy control page of the subsidiary would interact with the privacy control page of the conglomerate was also distinguished as an issue.

Key Tension Point 2: Data Take-out and The Right to be Forgotten
The data subject has the right to request for erasure of personal data concerning them, from the controller. The data subject also has the right to request for a data take-out. As, logged in users and non-logged in users both can use the news service application, a crucial issue that was observed was, how the stored personal data from a non-logged in data subject would affect the database, if the data subject were to log in? Furthermore, what would the implications be if the non-logged in data subject would log in, or vice versa, and request for a data take-out or data erasure?

Data Life Cycle of the PA; Registering as a New User
Implicit data can be seen as data that is not intentionally provided but derived from analysis of explicit data. Thereby, explicit data could yield implicit data, allowing assumptions to be made based on non-explicitly declared information. Explicit interaction demands direct attention. When the peripheral information, from the engagement with explicit interaction, behaves in the background, i.e. implicit interaction, the data subject receives a response, in the form of an explicit interaction. The choreography between implicit and explicit interactions are intangible, thereby making it difficult for a data subject to grasp the data life cycle, i.e. the process of how data yielded from an explicit interaction is stored via implicit interactions.

Understanding this process allows one to understand PAs in a more complex form. A simple example could be; registering as a new user. The following is an example of the processing of a data life cycle, in the scenario of registering a user. When a user signs up to create an account, the user is engaging with explicit interactions. Explicit personal data is provided by the data subject, such as email, age and gender. This explicit data is then stored explicitly in settings. However, in order to be able to store this data explicitly, implicit interactions are required, where first a conglomerate id is created, this id is a user account id, created in order to distinguish the user within the conglomerate database. The user id is then sent to the news service application and thereby registered within the settings database. See Figure 1.

![Figure 1. Implicit and explicit interactions of registering as a new user within a news service application](image)

Mapping Processing Activities
The Processing Activities were categorized into four categories. These categories helped to serve as a boundary object for the multidisciplinary team during the first design studio. A boundary object is an object which is adaptable enough to translate between viewpoints of participants from different social worlds, yet robust enough to preserve a common identity across the heterogenous workspace. [27] During the design studio, each team member tackled the problem with different concerns, motivated by their own domain, which they inhabited, where the boundary object facilitated the communication between the domains, adopting different identities according to the domain that it inhabited. The four different categories held of: Product Development, Technical Maintenance, Personalization and Advertisement. These four categories were also used as a boundary object to address the data subjects of the different PAs.
CONSTRUCTING THE DESIGN SPACE

Design processes and their contributions of how to continue in the next stage will be presented in this section.

Iteration 1

Design Studio 1

A design studio creates a space for cross-functional teams to come together and visualize potential solutions to a design problem. By breaking down organizational silos and thereby allowing all fellow teammates to focus on the same challenge and express themselves on the specific matter, gives the possibility to tackle a problem from many viewpoints [13]. For this reason, a design studio was held with the whole team, consisting of developers, product managers, designers and other competencies. The design studio consisted of the following five steps: 1) to ensure that all participants are tackling the same problem; a presentation was prepared, 2) individual idea generation; all participants drew as many ideas as they could come up with in ten minutes, 3) idea presentation and constructive feedback, 4) step two was repeated, but where iterations were made on previous ideas, 5) converge on one idea as basis for creating and testing the Minimal Viable Product (MVP). However, the design studio resulted with two final designs, instead of one, since the team could not agree, due to that two very interesting themes emerged.

Minimal Viable Product (MVP)

A MVP is the smallest possible product that can be built to test a hypothesis; where it is used as a tool for learning and is therefore useful when wanting to test different approaches [13]. Using a MVP makes it possible to focus on validating the proposed solutions by testing it on actual users and receiving feedback [13]. The two resulting designs (paper sketches), from the design studio, were drawn digitally, in order to make the sketches more clear.

MVP 1

The first MVP consisted of a descriptive copy with a personal tone, to inform data subjects of the processing of their personal data. The information was presented as a message from the news service team, where three different employee representatives, (editor, UX and developer) presented why data was important for them to make the product better. For example the developer approached the users with the headline “I wish to deliver a quick and well-functioning app”.

MVP 2

The second MVP, (see figure 2), was an example of how a detailed data flow chart could look like. This data flow chart visualized how each data entity could be used and for what reason. Stripping away the abstract ambiguous content down to a detailed data flow chart, made it possible to defamiliarize the meaning of data and thereby create space for critical reflection. In this way, it was also possible to get a better understanding of the fears and values related to personal data.

Figure 2. Minimal Viable Product 2; Data Flow Chart

Privacy Segment Personas

Through interacting with stakeholders, a resource emerged, which contained privacy segment data, identified by the conglomerate. This was synthesized into a framework of three different privacy segments that made it more accessible. The following three privacy segments were used in order to categorize participants accordingly to personas; A, B and C.

Persona A: Highly paranoid about data and digital footprint. Believes to know more than the average person about privacy and security and considers people to not be protective enough with their data.

Persona B: Aware of data risks and takes some precautions but not too worried about it. Feels ok with sharing data but is sensitive to misuse of data such as spam.

Persona C: Careless of data tracking, feels that they have nothing to hide and are willing to trade data for better services, sees it as a duty to improve experiences. Likes to be entertained and believes in personal growth by using the internet to explore interests and gather perspectives.

Synthesizing these privacy segments was a way to verify that all entities were communicating the same thing. Thereby, this also served as a boundary object.

Case study 1

Four current users of the news service application were recruited. The two MVPs were tested on these four users. Unfortunately, the group only consisted of males between the ages of 44 - 72. All males had a background of previous higher educations and high job positions. First, semi-structured interviews were carried out in order to categorize the participants into privacy segments accordingly to the three personas A, B and C. Thereafter,
Further Defining Privacy Segments based on MVP 1 and 2

Participants expressed themselves differently based on their fundamental concerns for privacy. The privacy segments were verified and further observations were added to the privacy segments from the user studies.

**Persona A**
Based on the presented MVPs, the user was very skeptical to the data processing and the information that was provided regarding the data privacy. The user thought that only a fraction of what personal data was actually stored and managed was being presented. For MVP 1 the user expressed irritation for not receiving concrete facts and instead concealing information with deceiving text. For MVP 2, skepticism still remained and the user doubted if we were being honest about all data processing or only showing a fraction of the data processing: “I understand everything! But be honest, is this all the data you collect?”

**Persona B**
Users felt more confident regarding data processing when we informed the user what we were doing with their data. One user expressed: “I think it's good! That's what you do not know on other sites, you know that data is being collected but not what exact data.” Another also explained: “I feel safer. This answers a lot of questions. I like when it's not so mysterious.”

Furthermore, a concern for MVP 2 was expressed. Even though the user liked the logic and preferred visual images, the user was concerned that others would not be as familiar with such visualizations of data. The user also expressed concern for the used terminology: “…maybe device ID is something that my mother would not understand.”

**Persona C**
Users felt very confident in the management of their data. They thought it was good to use the data, in order to improve the product. Users also became curious about their own data and what was registered. To the question, “What are your thoughts about companies collecting data about you?”, one user replied: “I think it's good. Then you can analyze what are the customer needs and in that way know what you can do to meet them.” Another replied: “Good to make use of the data that is available!”

The main takeaway from this was that, users from the different privacy segments responded differently to the two presented MVPs. This confirmed the necessity for a transparency-layered approach, in order to be able to reach out to all privacy segment personas. This thereby, resulted with a transparency-layered approach in the second iteration.

**Iteration 2**

**Design Studio 2**
The second design studio was necessary in order to find out how to present the transparency-layers. It was held in the exact same manner as design studio 1. However, the participants only consisted of the UX-team. Also, the iterated ideas were based on the findings from the previous user study. This design studio resulted with the team to converge on one single idea, to further test.

**Transparency Layers**
Based on the findings from the privacy segments and central corporate decisions, a transparency-layer guide was adopted. The layer guide presented by the parent company was analyzed and thereby adopted accordingly, to suit the news service. The transparency-layer guide suggested four layers of information. However, three layers of information was instead adopted to the final MVP with a focus on layer 2, as layer 2 was the main layer that all privacy segments would reach. The first layer was to be presented when users entered the application and thereby shown before starting any tracking or collection of data. The second layer held of general information with examples of how and what data could be tracked and collected. This layer also provided a data control panel, where opt-outs, data take-outs and data deletion was offered. The third layer was intended to give a more detailed explanation of what exact data is processed and for what specific reason. In this way, the different privacy segments can be addressed accordingly to their needs of how much they want to find out.

**MVP 3**
MVP 3 was created with the help of the wireframing tool Balsamiq and focused on layer 2, in the form of a clickable prototype with a descriptive copy and a personal tone. Links were included to allow the user to find out more, if so wished. Specific examples of what data is used for product development were given, as an example. Also the conglomerate was explained and a data control page with data opt-out, take-out and data erasure, was provided. MVP 3 allowed for a transparency-layered approach on different privacy segments to be tested.

**Case Study 2**
Six current users of the news service application were recruited to the case study. The users consisted of four females and two males, between the ages of 25 - 34. General privacy semi-structured interview questions were asked, to be able to categorize users into privacy segments. Thereafter the users were shown the clickable low-fi
prototype and given a set of specified tasks to complete. As the participants completed the assigned tasks, they were asked to think aloud, i.e. Think-aloud task based user testing. Thereafter, open-ended questions were asked regarding the assignments and about topics that had relevance to the MVP. From this, the remaining KTPs were discovered.

RESULTS AND DISCUSSION
The above mentioned activities resulted in nine tension points, where two, i.e. KTP 1 and 2, are mentioned in the ‘Understanding the Design Space’ section and are a result from the data inventory. KTP 1 and 2 require a more technical approach. KTP 3-9 are mentioned below together with a discussion for each.

Key Tension Point 3: Data collection Awareness
Participants that already were aware of data collection in general were not as scared by the presentation of the MVPs. Participants found it positive that companies were being honest about the collection and tracking of data and thereby more open to it, as they expressed that this is something that almost all companies already do but do not explicitly admit; “Everyone does it so it’s just good to be open with it and say we do this, and this is what you get for it.” Another participant expressed; “I do not think it’s a surprise to me. It’s hard to be private and protected all the time, the only way would be to stay away from the internet altogether, but then you can of course influence how much data you give away.” Participants that were aware of data collection were also curious to find out more about what data was being collected.

However, participants that did not have prior knowledge of data collection in general, were troubled by the MVPs, especially MVP 2. Participants found it scary that data was collected and when they saw explicitly what data could be collected of them; “It’s a bit new to me that I’m caught up in this, I have never known that I’ve ever been a contributing factor... I feel unsafe giving away my data, it feels like this is a violation of my privacy and my personal integrity.”

This KTP shows that not only are the different privacy segments something that should be considered when designing privacy policies but also awareness. However, it can also be argued that, when GDPR comes into effect, there will be even more awareness of data collection, which might create less of a surprise.

Key Tension Point 4: Data Deletion
Another feature, that might be new to the data subject, is the possibility to ask for a data erasure. Participants did not believe that their data would be erased when they pressed delete data. This was due to that they believed it to be too easy to delete valuable data. “I don’t believe that my data is being deleted. It seems so simple and it feels like it is stored and stored again. It is hard to believe that all valuable information can be deleted that easily.” This exemplifies the awareness of data being valuable and demonstrates participants to value their data highly and their awareness of their datas’ significance for companies. Thereby, this could be interpreted as participants showing less trust for the company, as data subjects’ express concerns for being deceived. This also further identifies users’ perception of data and their different mental models of what data is. Users might not only see data as one single entity but a database, which constructs who they are and thereby allowing companies to take advantage of this database, to sell them products, by using targeted advertisement based on their data.

Another explanation for not believing their data to be deleted could be due to the perception of time. The perception of time might have an affect on the appreciation for the completed task. When the time progress of a performed action exceeds the user’s expected time evaluation, the user may not appreciate the completed task or believe that the action happened at all. Thereby, the reduction of the perceived waiting time for the data deletion, i.e. perceived as a big database, together with the awareness of datas’ significance, might result in distrust. This results with the takeaway of adding waiting time as well as an extra step in between of where the data deletion is placed. As the data deletion option was placed in layer 2, this would suggest to place it in layer 3, as well as some steps in between, such as asking the user if they are sure they want to delete their data. Participants also thought that it would be good to be able to choose specific data to delete and not be forced to either keep or delete all data. This highlights the technical implementations that require to be examined and thereafter, further explained to the user.

In regards to identifying users’ definition of what data is, KTP 5 further establishes the infrastructure for users’ understandings and perceptions of data.

Key Tension Point 5: Active and Passive Data
A difference between data generated knowingly; actively, and data generated from a click or user behavior; passively, was identified among participants. For example, one participant explained how posting something on Facebook was according to them perceived as more valuable data than a click on an ad or article. What the participant did not produce in his or her own words or pictures, felt acceptable to share. Active data was thereby defined as more valuable and hence more threatening to an individual’s personal integrity than passive data.
This tension point, explains how data subjects refer to data and value data. Defining data from the end-user perspective helps to understand the end-user’s core values and interpretations. By identifying what data means for them and which data is of significance, they are thereby expressing the boundaries of where their definition of personal data, moves between. As a data subject, you do not see the data when you click on something, but you see it when you for example upload a picture. In this way, active data is explicit and tangible, making this data valuable, whereas passive data is insignificant as it is intangible data. However, making the implicit data explicit, by explicitly visualizing implicit interactions such as from a data take-out or data control page, all of a sudden, can make implicit data important, as now it is something that the user can see and comprehend. This explains why tension point 6 and 7 are especially significant.

Key Tension Point 6: Purpose for Collection of Data
Participants found that collecting data was acceptable as long as its purpose was stated for the collection of the specific data. For example one participant did not understand why gender was data that was necessary to collect; “I’ve always wondered why it’s important to provide gender? That data does not feel relevant. I do not understand what you are going to do with that info.” Another participant similarly explains how it feels understandable if a map application would ask for location as necessary data to be provided but if the news service application, for example, would require this, it would be “uncomfortable”. Thereby, it could be concluded that by providing the reason for why data is needed, the user will be more susceptible and thereby be more prone to allow the sharing of that data. Giving purpose for collecting data before the collection of data, allows the data subject to understand why this data entity is required to be collected and thereby could lead to a more transparent and trusting relationship between data subjects and stakeholders. However, a simple explanation can sometimes not be satisfactory enough, as data subjects still might feel the relationship to be treacherous if their data is used without value given back. This brings us to the next KTP: return of value.

Key Tension Point 7: Return of Value
An incentive to share data is created when value is returned. Participants were positive to sharing their data if they would receive something in return for doing this. To the question, “What are your thoughts about companies collecting data about you?”, one participant replied: “I think it’s positive as long as I get something out of it.” To the same question, the following participant, compares their general knowledge of collecting data to their existing knowledge of Google, who also collects and tracks data. The participant expresses anxiousness about collecting data but then shows openness for it, as long as some value is given back; “I think this is the same thing as Google is doing, keeping track of my history and my preferences, and in that way knowing how to customize the content... I get a little ‘big brother’ feeling, but if it helps to make a better experience for me then I can live with it.”

The discussion of receiving value in return for giving away ones data, demonstrates the participants’ engagement within a risk-benefit analysis based on a privacy trade-off assessment. This demonstrates how participants think of data, where they almost see it as a physical metaphor i.e. a commodity. Not receiving value, exposes their concerns of themselves being the product, where they implicitly might ask; am I the product? This results with the takeaway that corporates need to give back value to their data subjects, when asking for data. To be able to do this, corporates necessitate to know what their data subjects value and how they might assess costs and benefits within privacy trade-offs. In regards to values, two key value mismatches were identified, i.e. key tension point 8 and 9.

Key Tension Point 8: Advertisement and Personalization
Participants found the control panel to be positive since they could easily get an overview and be able to simply opt-out. Since participants felt that they trusted the news service brand, they were willing to give consent to the collection of their data. However, users felt a little less at ease giving their consent to advertisement. Participants tried to specifically find out more about advertisement by pressing on the link to read more. One participant explicitly expressed: “But here I would have liked to read more about how my data is used for ads. If it’s only based on news I’m reading and if it’s linked externally with, for example, cookies.”

Some participants also expressed their interest for personalization and would have wanted to know more. However, when reading about personalization, participants interpreted that news would be shown based on what they read. For this reason, it is possible to conclude that participants expressed a fear of resulting in a state of intellectual isolation, i.e. a filter bubble, which would be a result from past click-behavior and search history. Hence, data subjects would only be shown information that would potentially agree with their viewpoints. This thereby would result with isolation within one's own cultural and ideological bubbles. As this was not something that the news service did, this KTP was very important to clarify, since this was a key value that data subjects showed great anxiousness for. The main takeaway was thereby, not only focusing on what data processing to inform users of but the value of informing users of such that corporates do not do,
such as explaining that information is not provided accordingly to the users click- and search-behavior.

This also demonstrates what participants valued from the different categories consisting of, Product Development, Technical Maintenance, Personalization and Advertisement. Expressing interest and fear for the two categories; Advertisement and Personalization, reveals the value mismatch for how data subjects think that they receive a “value”, which they do not value. This creates a value mismatch.

Key Tension Point 9: Stakeholder
After reading more about the conglomerate, participants changed their willingness to give their consent, especially in regards to ads, as this already was a concern from the beginning. Participants expressed disappointment, since they had felt that they had been misinformed and trusted the news service with their data but this did not mean that included the whole company. One participant explained how the previous information felt misleading, when finding out about the conglomerate; "It was a little sneaky I think. That I say yes to share data with the news service application and then find out that it's not just the news service application but it's the whole group! Must get that information before! You must be able to choose which products you want to share data with and which you don't want to!"

Some users were surprised that the news service was a part of a bigger corporate. Participants also wondered why the whole conglomerate needed their data and what specific data was being sent to the conglomerate. Users also explained how they might not have had any problems sharing data with the news service because they thought that it was a great product but what if there would be other products that they might have been less fond of, within the conglomerate. Then, they would have not wanted to share their data, which in turn could affect the news service.

When considering an important personal matter and introducing a new unexpected variable to the equation, results in reflecting on the wrong thing, which can affect the experience negatively. However, if one were to familiarize users of the parent company relationship, someplace else, this would make it possible to desensitize the users. The main takeaways thereby held of making sure to inform users about the ecosystem of the news service application and its relation to the conglomerate before presenting the privacy control page, as well as, specifically clarifying what exact data is being shared with the whole conglomerate, how and for what reason.

Future Research

As the provided data in the MVPs were fictitious, the final MVP should be retested with real personal data of the participants. This is something that could have affected the study. Further attention should also be placed on providing a full-layered approach, when conducting case studies with different privacy segment users. Future research, in coherence with this study, should also look into how to further address the presented KTPs, identified in this study. Given that GDPR is a topic that is currently of great relevance, a lot of new research will not only be accessible before the law comes into effect, but also after, where an example would be to study the users’ acceptance of the new bylaw, after data subjects have been familiarized with the new legislation.

CONCLUSION
This paper aims to suggest guidelines for how to approach the new challenge of GDPR from a user experience perspective. This was done by analyzing relevant proceeding processes, which suggests that, even though a data inventory is a requirement of the GDPR, it can also be used as a boundary object, when further analysed and presented to the rest of the team and users. The boundary object is a contribution of, in an effective way, enabling different units, such as researcher, stakeholders and users, to understand each other and thereby be able to extract tension points. This is a way of working that can consequently help units to lead to a solution. Further, identifying different privacy segments, can also serve as a boundary object. It was also concluded that, addressing different privacy segments with a layered approach could help to provide the information accordingly to the different segments. Presenting the key information immediately and then specifying more detailed information elsewhere, allows users to choose for themselves, how much information they would like to access. This paper, thereby, suggests the following; 1) Data Inventory Matrix, 2) Privacy Persona Mapping, as boundary objects, and 3) Transparency-Layering, to use as a strategy, within the design process of exploring the design space and establishing a common ground, between all involved parties.

Understanding the data subjects’ values and understandings, not only benefits GDPR compliance, but also lays the grounds for a similar language of how to speak of data processing. The KTPs presented in this paper helps to lay those grounds. From a design perspective, the nine KTPs can be reformulated into five actionable design guidelines:

1. Create a Privacy Common Ground: Map and identify values and integrity infringements. For example, KTP 4, 5 and 8, identifies how users define data and which PAs risk to violate users’ privacy.
2. Identify Privacy Trade-off Rewards: For all PAs, hold an exercise where each PA’s purpose is explained and its respective reward, for a user. This guideline is based on KTP 6 and 7, (explaining the purpose respectively the benefits).
3. Desensitize and Familiarize Users Appropriately: Potential costs should be prior explained in an appropriate context other than the privacy policy, to familiarize users with subsidiary structures. KTP 9 further suggests how to deal with some potential privacy problems by suggesting to desensitize users to avoid false assumptions and distrust.
4. Privacy Design for Data Subjects: Design for all users, when designing a privacy policy. This could include users that are aware/not-aware of data collection (for example KTP 3), existing users, old users returning to your product and new users. The transparency-layer also helps with this.
5. Map User Journey for a Control Panel: Identify and define the inextricable technical relationships of data managements provided within the control panel. For example between subsidiary and its conglomerate (KTP 1) or non-logged in vs. logged in, user journeys (KTP 2). Identify implications of the different user journeys and define their possible design solutions.

These findings are general in their context of a news service but in a broader aspect, can be useful for others whom aim to be GDPR compliant, in all companies. The guidelines are also advantageous for how to design for privacy, in general. Therefore, these findings are not only applicable to other similar subsidiary companies, but also for other companies, designers or entities/individuals working with GDPR and design.

REFERENCES


37. EUGDPR. Retrieved March 15, 2018 from https://www.eugdpr.org/key-changes.html


