Designing a GDPR-Compliant and Usable Privacy Dashboard

Philip Raschke, Axel Küpper, Olha Drozd, Sabrina Kirrane

Designing a GDPR-compliant and Usable Privacy Dashboard

Philip Raschke, Axel Küpper, Olha Drozd, and Sabrina Kirrane

Service-centric Networking, Telekom Innovation Laboratories,
Technical University Berlin, Germany
Vienna University of Economics and Business, Vienna, Austria
{philip.raschke,axel.kuepper}@tu-berlin.de
{olha.drozd,sabrina.kirrane}@wu.ac.at

Abstract. The role of personal data gained significance across all business domains in past decades. Despite strict legal restrictions that processing personal data is subject to, users tend to respond to the extensive collection of data by service providers with distrust. Legal battles between data subjects and processors emphasized the need of adaptations by the current law to face today's challenges. The European Union has taken action by introducing the General Data Protection Regulation (GDPR), which was adopted in April 2016 and will inure in May 2018. The GDPR extends existing data privacy rights of EU citizens and simultaneously puts pressure on controllers and processors by defining high penalties in case of non-compliance. Uncertainties remain to which extent controllers and processors need to adjust their existing technologies in order to conform to the new law. This work designs, implements, and evaluates a privacy dashboard for data subjects intending to enable and ease the execution of data privacy rights granted by the GDPR.

Keywords: Data privacy, Privacy dashboard, General Data Protection Regulation, Usability, Transparency-enhancing tools, Privacy-enhancing tools

1 Introduction

In the age of digitalization, the data privacy of an individual can be severely violated by technology. Cases like Google Spain v AEPD and Mario Costeja González[1] highlight the extent of harm technology can do to an individual person by simply providing inaccurate (in this case outdated) information about the *data subject*. Its controversy had to be eventually decided by the European Court of Justice (ECJ), the highest court of the EU. While the case was solved with a verdict in favor of individuals' data privacy, doubts remained, which were fueled by the revelations of Edward Snowden in 2013, also called the *Snowden*

[1] ECLI:EU:C:2014:317. http://curia.europa.eu/juris/documents.jsf?num=c-131/12, last accessed: 07/04/2017.

Effect[2], and underlined by the invalidation of the *Safe Harbor Privacy Principles* by the ECJ[3] in 2015. The EU addresses these concerns with the General Data Protection Regulation (GDPR)[4], which comes into force in May 2018. The GDPR replaces the Data Protection Directive[5] of 1995 by extending the data privacy rights of data subjects in the EU with the goal to adapt to modern data privacy challenges.

A major change of the GDPR, among others, is the explicit requirement of *transparency* when processing personal information.[6] In the recitals of the GDPR the lawmakers explain that "[t]he principle of transparency requires that any information and communication relating to the processing of those personal data be easily accessible and easy to understand, and that clear and plain language be used".[7] Taking it literally, this would mean data subjects should be able to obtain *any information* they want, including the time a *controller* (i.e. a legal entity that processes personal information) accessed their personal data, from which source, to which *processors* (i.e. legal entities that process personal information on behalf of the controller) it has been forwarded, which data has been derived from it, and so on. However, in times of *Big Data* and *Cloud Computing*, providing this information can be very complex, considering the sheer amount of data a controller might process of a single data subject. Moreover, the processing often involves external third parties, since controllers might use the infrastructure of one or multiple service providers.

The personal data in question is mostly processed digitally, thus it is accessed and assessed by technical means. Granting the privacy rights of the GDPR should be realized by the same means. For this reason, we propose a privacy dashboard, which aims to offer and manage these data privacy rights. To tackle the complexity of the task and achieve a user-friendly result, a usability engineering methodology is applied.

The remainder of the paper is structured as follows. Section 2 discusses requirements for the privacy dashboard imposed by the GDPR. In Section 3, we give an overview of related work in the field of transparency-enhancing tools (TETs) as which privacy dashboards are classified. Section 4 presents the methodology, which is adapted to design the privacy dashboard. In Section 5, we analyze the potential users of the dashboard and the tasks they are supposed to fulfill with it. Based on the analysis, a design is derived that is presented and discussed in Section 6. The development of a prototype and its evaluation

[2] What is Snowden effect? - Definition from WhatIs.com. `http://whatis.techtarget.com/definition/Snowden-effect`, last accessed: 07/17/2017.

[3] ECLI:EU:C:2015:650. `http://curia.europa.eu/juris/documents.jsf?num=c-362/14#`, last accessed: 07/17/2017.

[4] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 4.5.2016, p. 1-88 [hereinafter GDPR]

[5] Council Directive 95/46, 1995 O.J. (L 281) 31 (EC) [hereinafter Directive 95/46].

[6] GDPR art. 5(1)(a)

[7] GDPR Recital 39

are presented in Section 7 and 8 respectively. Finally, we conclude our work in Section 9.

2 GDPR

The GDPR will be law in 28 countries, but more will be affected by it due to its territorial scope. Controllers from abroad will be subject to it if they offer goods or services to European data subjects or monitor behavior, which happened in the Union.[8] The GDPR consists of 99 articles and 173 recitals. It is a comprehensive regulation covering multiple scenarios in which personal data is processed. This can be seen in Article 6 of the GDPR, which defines conditions for lawful processing of personal data. Given informed consent by the data subject[9] is only one out of a number of bases, including processing personal data to fulfill legal obligations[10] or for tasks carried out in the public interest[11]. To narrow the scope, we only focus on processing of personal data based on consent given by the data subject.

To access, review, and manage personal data in a digital format, technological means are necessary. Thus, compliance with the GDPR requires technology to adapt to it. Furthermore, new means must be introduced to grant and use the data privacy rights of the GDPR. Bier et al. [2] draw the same conclusion.

As stated above, the explicit requirement of transparency is one of the major changes of the GDPR compared to its predecessor, the Data Protection Directive of 1995. It required personal information to be "processed fairly and lawfully"[12], which is extended by the GDPR by adding the expression "*and in a transparent manner*"[13] to it. As mentioned in the previous section, the recitals attempt to narrow the transparency principle down, however, it remains debatable which information has to be provided to the data subject to meet the transparency requirement. The data subject can be provided with an overwhelming amount of meta information that is measured whenever personal data is processed. The meta data could give answers to the questions: When was the data collected? From which device was it obtained? To whom was it forwarded? What is the physical location of the processing servers? A first step towards transparency is to grant the *right of access*[14]. Siljee's [14] *Personal Data Table* fulfills all requirements to realize the execution of this right. The Personal Data Table should be extended by an element to depict data flows to involved processors.

Articles 16 and 17 of the GDPR grant data subjects the right to request rectification[15] and erasure[16] of data *without undue delay*. Moreover, the controller

[8] GDPR art. 3(2)
[9] GDPR art. 6(1)(a)
[10] GDPR art. 6(1)(c)
[11] GDPR art. 6(1)(e)
[12] Directive 95/46 art. 6(1)(a)
[13] GDPR art. 5(1)(a)
[14] GDPR art. 15(1)
[15] GDPR art. 16
[16] GDPR art. 17(1)

is obliged to respond to these requests within one month. This time period is extendable by two additional months with regard to the complexity of the task and the number of requests.[17] For our design of the dashboard, this means the Personal Data Table must offer the possibility for each data item to request rectification or erasure of the corresponding information.

The Data Protection Directive required consent to be given *unambiguously*[18], while the GDPR now requires the informed consent to be given for *one or more specific purposes*[19]. The recitals advise that if data is used for multiple purposes, consent shall be given for each purpose separately.[20] Furthermore, the data subject shall have the right to withdraw consent at any time and as easy as it was to give consent.[21] The dashboard must include a possibility to review consents given, the purposes they were given for, and a functionality to withdraw them at any time.

The dashboard is supposed to work as interface between data subject and controller. Requests for rectification, erasure or withdrawal of consent cannot be expected to be responded to immediately. Thus, a message section to obtain status information about pending requests is reasonable. The controller may approach the data subject via the dashboard to ask for consent of processing personal data for additional purposes. This way the privacy dashboard may be extended by ex ante capabilities, while being mainly designed as ex post TET.

3 Related work

Since decades there are numerous and manifold tools that address data privacy issues. Hedbom [5] provides a classification of TETs in 2008. The criteria to classify the tools include the possibilities of control and verification, the target audience and the scope of the tool, the information it presents, technologies it uses, and its trust and security requirements. Hedbom discusses his classification by applying it to examples. For this reason, the Transparent Accountable Data Mining (TAMI) system [16], the Privacy Bird[22], the PRIME project [4], the approach to obtain privacy evidence in case of privacy violations by Sackmann et al. [12], and Amazon's book recommendations service [17] are presented and explained.

Based on his work, Janic et al. [6] further develop the classification and extend its definitions of TETs by identifying and discussing 13 tools. According to them, tools like the Mozilla Privacy Icons[23] and Privacy Bird fall under tools that address the complexity of privacy policies of websites. The PrimeLife

[17] GDPR art. 12(3)
[18] Directive 95/46 art. 7(a)
[19] GDPR art. 6(1)(a)
[20] GDPR Recital 32
[21] GDPR art. 7(3)
[22] Privacy Bird. `http://www.privacybird.org`, last accessed: 07/20/2017.
[23] Privacy Icons. `https://disconnect.me/icons`, last accessed: 07/20/2017.

Privacy Dashboard[24] and the Google Dashboard[25] are ex post TETs, which provide information on collected and stored data by service providers. Lightbeam[26] and Netograph[27] visualize user tracking that is realized via third party cookies. The tool Web of Trust[28] ranks websites according to their trustworthiness, which bases on a reputation system. Janic et al. classify Me & My Shadow[29], Firesheep[30], Panopticlick[31] and Creepy[32] as tools that aim to raise privacy awareness by informing the user about techniques commonly used to violate their data privacy. The tool Privacy Bucket[33] and the Online Interactive Privacy Feature Tool by Kani et al. [8] have been released after the paper of Janic et al. was published, but fit in the previous described category.

To the best of our knowledge, the most recent privacy dashboards under development are GenomSynlig, which was merged into the Data Track project[34] by Angulo et al. [1] published in 2015, and the tool PrivacyInsight by Bier et al. [2] presented in 2016. While Data Track visualizes data disclosure in a so-called *trace view* and thus realizes the transparency principle of the GDPR, PrivacyInsight aims to address the GDPR as whole including the transparency principle, right to rectification and erasure, and the withdrawal of consent. Bier et al. identify legal and usability requirements for a privacy dashboard. In total they present 13 constraints, eight that are legal and five that are usability requirements. A brief summary of the legal prerequisites is given below, while the usability requirements are left out due to page limitations.

R1 The right to access must not be formally or technically constrained.
R2 A privacy dashboard must be accessible by every data subject.
R3 Access to all data must be provided.
R4 All data must be downloadable in machine-readable format.
R5 Data flows to all processors and internal data flows must be visualized.
R6 All sources of personal data must be named.
R7 For all processing steps a purpose must be given.
R8 Means to request rectification, erasure, or restriction must be provided.

[24] PrimeLife Dashboard. `http://primelife.ercim.eu/results/opensource/76-dashboard`, last accessed: 07/20/2017.
[25] Google Dashboard. `https://myaccount.google.com/dashboard`, last accessed: 07/20/2017.
[26] Lightbeam for Firefox - Mozilla. `https://www.mozilla.org/en-US/lightbeam`, last accessed: 07/20/2017.
[27] netograph. `http://netograph.com`, last accessed: 07/20/2017
[28] WOT (Web of Trust). `https://www.mywot.com`, last accessed: 07/20/2017.
[29] Me and my Shadow. `https://myshadow.org`, last accessed: 07/20/2017
[30] Firesheep - codebutler. `http://codebutler.com/firesheep`, last accessed: 07/20/2017.
[31] Panopticlick. `https://panopticlick.eff.org`, last accessed: 07/20/2017
[32] Creepy by ilektrojohn. `http://www.geocreepy.com`, last accessed: 07/20/2017
[33] mfredrik/Privacy-Bucket Wiki. `https://github.com/mfredrik/Privacy-Bucket/wiki`, last accessed: 07/20/2017.
[34] pylls/datatrack: A tool that visualizes your data disclosures. `https://github.com/pylls/datatrack`, last accessed: 07/20/2017.

The requirement R2 includes in particular design strategies that enable access for data subjects with disabilities like visually impaired people. The privacy dashboard must implement accessibility interfaces like the WAI-ARIA[35] standard by the World Wide Web Consortium. The requirements R3, R5, R6, R7, and R8 impose a usability challenge with respect to the sheer amount of data taken into consideration. Internal and external data flows, as demanded by R5, can be complex to be visualized depending on the number of internal entities and external processors. Designing these data flows as graph in a comprehensible manner can be challenging. However, the information it depicts is fundamental in order to enable transparency. To support the data subject and to improve the intelligibility of this graph, it is reasonable to categorize and label personal data. A data subject might not be able to review each data flow to all processors in detail, but is interested in certain data categories.

4 Methodology

For the design and implementation of the dashboard, we adapt Nielsen's *Usability Engineering Lifecycle* [11]. It is considered fundamental in the field of usability engineering. In addition, it suits the design of systems well which address inexperienced users that desire to solve complex tasks [15]. For the following summary of the Usability Engineering Lifecycle Möller's notation [10] is used.

The development process starts with the *Analysis* phase, which examines the users, the tasks to be solved with the system, and the context of use. In the *Design* phase, the system is designed iteratively, however there may be parallel design versions, which are tested separately. In the *Prototyping* phase, the system is partly implemented. In this phase a differentiation is made between *horizontal*, *vertical*, or *scenario-based* prototypes. Horizontal prototypes present all functional capabilities of the system to the user, but do not provide the actual functionality. Vertical prototypes implement a certain feature of the system in depth, but do not include and present all planed functionalities to the user. The presentation but not full implementation of a certain feature is called scenario-based prototype.

The resulting prototype is evaluated in the *Expert Evaluation* phase by so-called usability experts in contrast to the *Empirical Testing* phase, which involves real users of the system, who are invited to test the tool under laboratory conditions. In the context of software engineering, this means a specific environment is set up including a predefined and tested device, a certain network connection, specific input tools, and so on. Various user studies can be conducted in both phases to either measure the overall quality of the system, or to identify flaws in the design. One of them is the cognitive walkthrough, which was first introduced by Lewis et al. [9] in 1990. After this phase, the next iteration starts, beginning with the Design phase. If the system is eventually deployed, feedback from real

[35] WAI-ARIA (Web Accessibility Initiative). `https://www.w3.org/WAI/intro/aria.php`, last accessed: 07/25/2017.

users in real-life scenarios can be collected and evaluated to further improve the system.

5 Analysis

Users of the privacy dashboard are potentially all natural persons in the EU. According to the statistics provider Eurostat of the European Commission, over 500 million humans lived in the Union in 2016.[36] These millions of people live in 28 countries, speak 24 official languages and almost the same amount of migrant languages, while using three different writing systems.[37] In 2016, 15.6% of the European population were younger than 14 years, 11.1% of them were between the age of 15-24, 34.1% between 25 and 49, 20.1% between 50-64, 13.8% between 65-79, and 5.4% older than 80 years.[38] These numbers highlight the challenge a uniform interface for this user base will be, however, it is further reasonable to investigate the user base's affiliation with information and communication technology. In 2016, about 71% of all individuals in the EU and 92% between the age of 16 to 24 accessed the Internet on a daily basis.[39] Moreover, 8 out of 10 users use a mobile device to access the Internet.[40] In 2012, 80% of individuals between the age of 16 and 24 used the mobile Internet to participate in social networks.[41]

Consequently, it can be inferred that a technological mean like a privacy dashboard reaches the majority of the user base, since it is rather familiar with technology and with the Internet. Web applications, which are optimized for mobile devices, suit well as platform. The privacy dashboard is intended to be used to execute data privacy rights granted by the GDPR. These rights are identified as the following tasks the tool should be used for:

T1 Execute the right of access
T2 Obtain information about involved processors
T3 Request rectification or erasure of data
T4 Consent review and withdrawal

[36] Eurostat - Population. `http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tps00001&plugin=1`, last accessed: 07/18/2017.
[37] Europeans and their Languages. `http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_386_en.pdf`, last accessed: 07/25/2017.
[38] Eurostat - Population by age group. `http://ec.europa.eu/eurostat/tgm/refreshTableAction.do?tab=table&plugin=1&pcode=tps00010&language=en`, last accessed: 07/25/2017.
[39] Eurostat - Internet use and activities. `http://ec.europa.eu/eurostat/web/products-datasets/-/isoc_bde15cua`, last accessed: 07/25/2017.
[40] Eurostat - Internet use by individuals. `http://ec.europa.eu/eurostat/documents/2995521/7771139/9-20122016-BP-EN.pdf/f023d81a-dce2-4959-93e3-8cc7082b6edd`, last accessed: 07/25/2017.
[41] Eurostat - Purpose of mobile internet use. `http://ec.europa.eu/eurostat/web/products-datasets/-/isoc\_cimobi\_purp`, last accessed: 07/25/2017.

The Analysis phase also includes the investigation on how the identified tasks would be or are solved without the tool. To the best of our knowledge, there is no dedicated tool to exercise any of these data privacy rights. Consequently, the execution of these rights heavily depends on the context of the controller. If the controller processes personal information digitally and offers the data subject a user interface, then the right to access, rectify, and erase data can be expressed or realized via this user interface. However, to inform about involved processors or to review and withdraw previously given informed consent, data subjects have to revert to written correspondence with the controller or to long privacy policies that nobody reads [3], but may give all required information on how data is forwarded to external third parties or the formal procedure to withdraw consent. It often remains uncertain how and whether controllers respond to these written requests of data subjects. In cases of severe privacy violations with social or economic damage, legal actions need to be taken.[42]

6 Design

This section discusses two possible architectures to deploy and operate the privacy dashboard and presents a first design approach, which serves as a basis for the development of the prototype.

6.1 Architecture

We ideally envision one privacy dashboard to manage all privacy rights with regard to all controllers a data subject is concerned with. As Figure 1 shows, Approach 1 requires each controller to deploy and operate their own instance of the tool, which the data subjects can access individually, while Approach 2 allows data subjects to access one instance of the dashboard to manage all controllers they deal with.

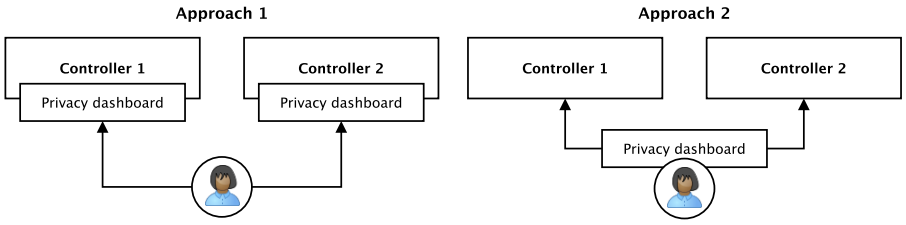

Fig. 1. Architectural alternatives for the deployment of the privacy dashboard. Either as single point to manage all controllers, or as data privacy management tool for every controller separately.

[42] ECLI:EU:C:2014:317 `http://curia.europa.eu/juris/documents.jsf?num=c-131/12`, last accessed: 07/25/2017.

A controller-operated instance of the privacy dashboard is easier to integrate into the data processing infrastructure of the controller. Consequently, no conversion of the personal data in question is necessary to adapt to an interface of an external third party. The controller would be able to modify and extend the privacy dashboard, for instance, to implement the visualization of customized or proprietary data formats. Security vulnerabilities are avoided, since the personal data in its entirety does not leave the boundaries of the controller, but queried chunks of it are transmitted to the data subject. The proximity of the privacy dashboard to the infrastructure of the controller eases the immediate and automated application of requests to rectify or erase inaccurate personal data. Requests made by the data subject could directly trigger internal processes providing all necessary parameters to take instant action. If the controller uses authentication mechanisms to authenticate data subjects in order to provide a service, the same technique can be used by the privacy dashboard to authenticate a data subject before delivering personal data.

While the data subject might benefit from a single end point to address all privacy concerns to, Approach 2 also implies a series of challenges. This approach is more challenging from an architectural perspective, since personal data from all controllers needs to be aggregated and served by a dedicated component. This would either require the standardization of a common data format or an agreement on an existing one. Interestingly, the right to data portability[43] granted by the GDPR may force controllers to develop or agree upon a common data format to exchange personal data. Still a transformation of the personal data is necessary, to adapt to the visualization logic of the external-operated privacy dashboard. A single machine that stores personal data of one or more individuals from multiple controllers is a security and privacy risk itself. Therefore, programmable interfaces should be defined by each controller to allow querying certain chunks of data. These interfaces require an authentication mechanism to ensure that personal data is transmitted to the right data subject. In this architecture distributed authentication techniques have to be used to solve the task. Consequently, the dashboard is ideally executed on the data subject's device, so no third party has to be involved, however, this comes along with hardware requirements that could violate Requirement R1 (*The right to access must not be formally or technically constrained.*) of Section 3.

In general, the adoption of the privacy dashboard by all controllers appears as a more likely approach, if it saves controllers the development of an individual privacy dashboard from scratch. Again, the assumption is made here that compliance with the GDPR implies the introduction of a privacy dashboard (see Bier et al. [2] R2).

6.2 Data taxonomy

The GDPR's explicit requirement of personal data to be processed transparently highlights the significance of the right to access. In order to execute T1 (as

[43] GDPR art. 20

defined in Section 5), all personal data has to be presented to the user. This data is displayed ex post like Siljee's [14] Personal Data Table. This enables answering the question: Which data collected the controller in question about me? The most challenging aspect of this task is to realize the visualization of huge amounts of diverse data. Consequently, the first approach to reduce the complexity of the data is to drill down the amount by limiting the presented data based on a time criteria such as data of the last month, week, or day. Simultaneously, by introducing this limit the dashboard needs to offer a functionality to select a time range the data subject wants to consider and review. This way, the data subject is able to ask more precisely the above mentioned question for a specific time range.

Despite this limitation, it might be that the sheer amount of data still overwhelms the data subject. Thus, it is reasonable to categorize the data and display the different categories. Since the context of use is data privacy, it is consequent to categorize the data according to a data taxonomy that addresses data privacy. Fortunately, Schneier [13] developed such a data taxonomy for social networks. A brief description of the categories is given below:

Service data is any kind of data that is required in order to provide the service in question (name, address, payment information).
Disclosed data is any data that the data subject intentionally provides on the own profile page or in their posts.
Entrusted data is any data that the data subject intentionally provides on other users' profile pages or in their posts.
Incidental data is any kind of data provided by other users of the service about the data subject (a photo showing the data subject posted by a friend).
Behavioral data is any kind of data the service provider observes about the data subject while he or she uses the service (browsing behavior).
Derived data is any kind of data derived from any other category or data source (profiles for marketing, location tracks, possible preferences).

To apply the data taxonomy to all kinds of controllers and not just to online social networks, we propose a generalization of Schneier's taxonomy. For this reason, we categorize disclosed and entrusted data into the category *Intentional data*, since both types of data are provided by the data subject intentionally. Furthermore, comprehensible labels for the categories are defined below:

Service data - Service data
Intentional data - Data I provided
Incidental data - Data of me provided by others
Behavioral data - Data of my behavior
Derived data - Inferred data about me

These categories can be applied to all kinds of controllers, although not each controller processes all categories of data. In our design for each category a view is offered with an individual Personal Data Table and the time limitation functionalities described above. In case that one or more categories are not applicable

to the domain of the controller, a simple information can be given that no data for this category is available. This might also confirm expectations of the data subject with regard to data collection practices of certain controllers. By applying the data taxonomy and offering separated views for each data category, the dashboard allows the data subject to easily find out whether a controller collects behavioral data of him or her or whether another user disclosed information about him or her.

7 Prototype

A prototype was developed with the JavaScript framework *React*[44] and the library *Material-UI*[45] to comply with Google's design standard *Material Design*[46]. The prototype has been made publicly available online[47]. With respect to the chosen methodology, a horizontal prototype has been developed that implements and presents all features to the user, however, provides reduced or no actual functionality. In practice, this means the scenario of our prototype is completely artificial.

We therefore define an online social network provider as our made-up controller that processes personal data of its users similar to popular services like *Facebook* or *Twitter*. All data presented in the dashboard is fake and does not belong to a natural person. However, to simulate a person's personal profile as accurate as possible with regard to the amount of data, we adapt an existing model from a study of the advertising agency Jung von Matt[48]. Furthermore, requests to rectification, erasure, or withdrawal of consent are not processed by a controller's backend. The filtering of data according to its processing context, data type, or time of its processing is implemented.

As it can be seen in Figure 2, we designed a three-column layout for the dashboard. We define general functionalities like reviewing given consent, displaying the privacy policy, and obtaining information about involved third parties, which are presented in the left column. Also in the left column and under the general functionalities, filter options are provided allowing the user to display personal data processed in a specific context, of a certain data type, and in a defined time range. The meaning of each processing category and each data type is visually supported by an icon, which is used in other components of the dashboard as well. In the center of the layout, the queried personal data is listed vertically in chronological order beginning with oldest entry. Each entry is furnished with an

[44] React - A JavaScript library for building user interfaces. `https://reactjs.org/`, last accessed: 11/13/2017.

[45] Material-UI. `http://www.material-ui.com/`, last accessed: 11/13/2017.

[46] Material Design. `https://material.io/`, last accessed: 11/13/2017.

[47] Privacy dashboard — IFIP Summer School 2017. `http://philip-raschke.github.io/GDPR-privacy-dashboard`, last accessed: 01/19/2018

[48] Jung von Matt study on typical German Facebook profile. `https://de.linkedin.com/pulse/das-h\%C3\%A4ufigste-facebook-profil-deutschlands-raphael-brinkert`, last accessed: 11/13/2017.

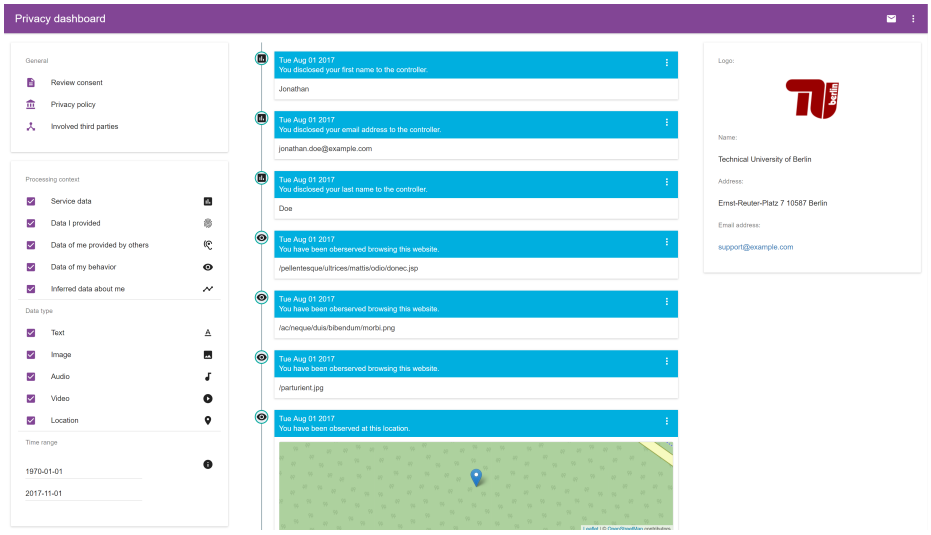

Fig. 2. The layout of the developed prototype. General functionalities and filter options are presented on the left-hand side. The queried data is in the center sorted chronologically beginning with the oldest entry. General information about the controller are presented on the right-hand side.

icon that gives information on its processing context. Under the actual date of when the processing took place, a short descriptive text about it is presented in the header of an entry, which is displayed above the actual personal data. On the right-hand side, general information about the controller are given, such as name, physical address, and email address to directly contact the controller.

In order to use the dashboard to execute task T2, a graph is displayed that shows the user data flows between controllers and involved processors (see Figure 3). In real-life scenarios often many processors are involved in the processing of personal data. There can be multiple controllers as well (so-called joint controllers[49]). Depending on the number of involved processors in the processing of the data subject's personal data, the complete graph can be shown as whole or processors can be clustered into groups according to their business domain for instance. Edges are annotated with data categories giving information on which data is exchanged. The arrows denote the direction of the data flow to clarify whether parties are just provided with data or if parties are actively exchanging data with each other. For the implementation of this graph the JavaScript library *vis.js*[50] has been used. Angulo et al. [1] propose a similar but more detailed approach with the trace view. To reduce complexity, data categories instead of specific data items are used in our approach.

[49] GDPR art. 26

[50] vis.js - A dynamic, browser based visualization library. `http://visjs.org/`, last accessed: 11/13/2017.

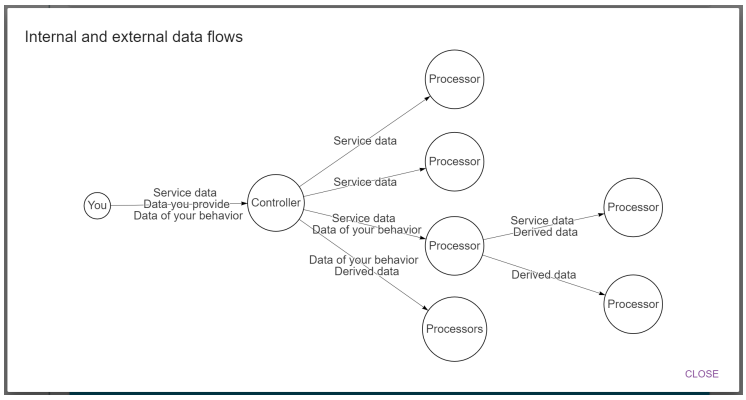

Fig. 3. A graph visualizing internal and external data flows between controller and processors. Edges are labeled with data categories indicating which data is exchanged with whom.

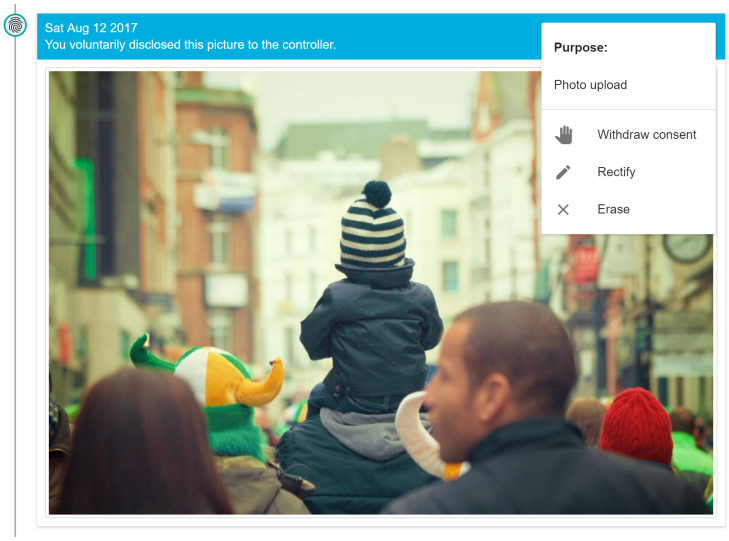

Fig. 4. For each specific data item the user is given information on the purpose of its processing, where applicable the possibility to withdraw consent, and the possibility to request rectification or erasure of the data.

Task T3 requires the privacy dashboard to offer a possibility to request rectification or erasure of the data item in question (see Figure 4). Additionally, for each data item information on the purpose of its collection and processing is given (see Figure 4). Multiple purposes can be listed here, if data is processed for more than one purpose. With the help of this component the data subject can answer the question: For what reason does the controller collect and process this data? A redirection to a separate section allows the user to review given consent and the possibility to withdraw it (see Figure 5). Since consent is supposed to be bound to a specific purpose, there is a label and a short description text to give more details about the purpose in question. With a simple interaction, like a click, it is possible to withdraw consent as easily as it was to give it.[51]

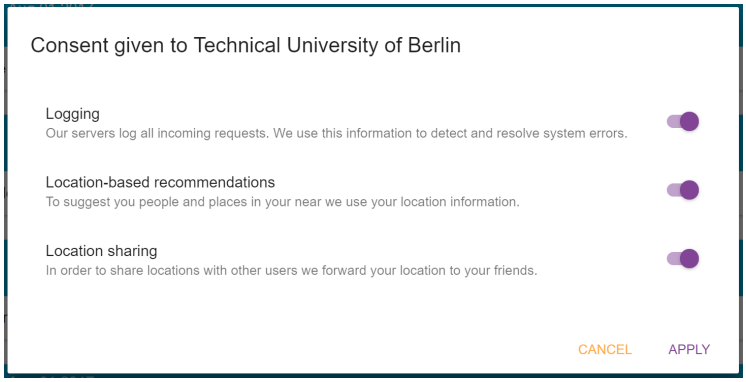

Fig. 5. A list of purposes for which consent has been given by the data subject. For each purpose a label and a short descriptive text is given. Consent can be withdrawn by simply clicking the toggle on the right.

8 Evaluation

To evaluate the design approach presented in this paper, an expert evaluation has been carried out according to Nielsen's Usability Engineering Lifecycle. The usability of the data categories is in focus of this evaluation. Möller [10] proposes a formative analysis consisting of a so-called *Thinking Aloud* test [7] with three to five participants to identify design flaws in a system. In the test, participants are asked to solve one or more specific tasks by interacting with the system while thinking aloud. An analysis of the participants' thoughts and remarks is conducted subsequently. In an expert evaluation so-called usability experts instead of real users are used, since the system might be in a too early stage to present it to external users. For this reason, three fellow researchers were given the following task consisting of multiple questions that have definitive answers.

[51] GDPR art. 7(3)

"European law gives you the right to request from any entity that processes your personal data access to it. Imagine you requested access to your personal data from a company and you're confronted with the tool in front of you. Please answer the following questions:"

- *Which data did you have to provide when creating an account for this service?*
- *Did you provide any voice recordings to the service?*
- *Have you disclosed your location voluntarily?*
- *Has anyone provided the controller with photos of you?*
- *Does this service provider track your location?*
- *Has the service provider knowledge about your gender?*
- *Does the service provider know your income?*
- *Does the service provider know which websites you visit?*

All participants struggled to answer the questions at the beginning, but managed to improve quickly answering the last questions rather fast and confidently. All participants answered the first question using the chronological order instead of using the respective data category assuming the data provided first is the data required for the registration. This is a clear indicator that the data category *Service data* is redundant and can be categorized as intentional data ("Data I provided"). The so-called *AppBar* at the top also contributed to confusion. The participants understood the privacy dashboard as a service itself, therefore tried to answer the first question with regard to required information in order to use the privacy dashboard itself. The participants found that the filter options were not visible enough and should be placed more prominent, considering that they are an essential part in the task solving process. Another concern of the participants is the technical feasibility of the data categories. This applies to incidental data ("Data of me provided by others") and derived data ("Inferred data about me") in particular. Generally, the scenario of the privacy dashboard is important. The participants were interested whether the system is operated by the controller or as a separate service, and if it can be used offline or if an Internet connection is required. The evaluation reveals that refining the data categories is necessary in order to improve the usability of the dashboard. However, it also shows that the developed prototype can be used by data subjects to answer questions relating to their data privacy.

9 Conclusion

This work presents the design and implementation of a privacy dashboard, which addresses the requirements of the GDPR and enables the data subject to execute data privacy rights with the tool. To substantiate the dashboard's design, its potential users and the tasks they are supposed to fulfill with it were analyzed and discussed. A prototype has been developed and evaluated. The results of the evaluation indicate that our design approach is worth pursuing and reasonable, yet needs further improvements and user tests. The redefinition of the

data categories and their technical feasibility will be researched in future work. Furthermore, architectures for the deployment of the privacy dashboard need more investigation. Comprehensive user studies are necessary to refine the current design of the dashboard and to develop alternative approaches.

Acknowledgments

Supported by the European Union's Horizon 2020 research and innovation programme under grant 731601. The authors would like to thank Dirk Thatmann, Iwailo Denisow, and Sebastian Zickau for their valuable feedback to the prototype.

References

1. Angulo, J., Fischer-Hübner, S., Pulls, T., Wästlund, E.: Usable Transparency with the Data Track - A tool for visualizing data disclosures. Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA 15. 18031808 (2015).
2. Biere, C., Kühne, K., Beyerer, J.: PrivacyInsight: The Next Generation Privacy Dashboard. Lecture Notes in Computer Science. 9857, 1226 (2016).
3. Borgesius, F.Z.: Informed consent: We can do better to defend privacy. IEEE Security and Privacy. 13, 103107 (2015).
4. Hansen, M., Borcea-Pfitzmann, K., Pfitzmann, A.: PRIME - A European Project for Privacy-Enhancing Identity Management. it - Information Technology. 47, 352359 (2005).
5. Hedbom, H.: A Survey on Transparency Tools for Enhancing Privacy. The Future of Identity in the Information Society - IFIP Advances in Information and Communication Technology Volume 298. 4th IFIP WG 9.2, 9.6/11.6, 11.7/FIDIS International Summer School, Brno, Czech Republic, September 1-7, 2008, Revised Selected Paper. 6782 (2009).
6. Janic, M., Wijbenga, J.P., Veugen, T.: Transparency enhancing tools (TETs): An overview. Workshop on Socio-Technical Aspects in Security and Trust, STAST. 1825 (2013).
7. Jaspers, M.W.M., Steen, T., Bos, C. Van Den, Geenen, M.: The think aloud method: A guide to user interface design. International Journal of Medical Informatics. 73, 781795 (2004).
8. Kani-Zabihi, E., Helmhout, M.: Increasing service users privacy awareness by introducing on-line interactive privacy features. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 131148 (2012).
9. Lewis, C., Polson, P.G., Wharton, C., Rieman, J.: Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. In: Proceedings of the SIGCHI conference on Human factors in computing systems Empowering people - CHI 90. pp. 235242. ACM Press, New York, New York, USA (1990).
10. Möller, S.: Quality Engineering. Springer Berlin Heidelberg, Berlin, Heidelberg (2003).
11. Nielsen, J.: Usability engineering. Elsevier, (1994).

12. Sackmann, S., Strüker, J., Accorsi, R.: Personalization in privacy-aware highly dynamic systems. Communications of the ACM. 49, 32 (2006).
13. Schneier, B.: A Taxonomy of Social Networking Data. IEEE Security & Privacy Magazine. 8, 8888 (2010).
14. Siljee, J.: Privacy transparency patterns. In: Proceedings of the 20th European Conference on Pattern Languages of Programs - EuroPLoP 15. pp. 111. ACM Press, New York, New York, USA (2015).
15. Thatmann, D., Raschke, P., Küpper, A.: Please, No More GUIs!: A User Study, Prototype Development and Evaluation on the Integration of Attribute-Based Encryption in a Hospital Environment. In: Proceedings - International Computer Software and Applications Conference. pp. 496502. IEEE (2016).
16. Weitzner, D.J., Abelson, H., Hanson, C., Hendler, J., Mcguinness, D.L., Jay, G., Waterman, K.K., Berners-lee, T., Kagal, L., Sussman, G.J.: Transparent Accountable Data Mining: New Strategies for Privacy Protection. 112 (2006).
17. Zwick, D., Dholakia, N.: Whose Identity Is It Anyway? Consumer Representation in the Age of Database Marketing. Journal of Macromarketing. 24, 3143 (2004).