Towards Improving Transparency, Intervenability, and Consent in HCI

Farzaneh Karegar

Faculty of Health, Science and Technology
Computer Science

LICENTIATE THESIS | Karlstad University Studies | 2018:9
Towards Improving
Transparency, Intervenability, and Consent in HCI

Farzaneh Karegar
Towards Improving Transparency, Intervenability, and Consent in HCI

Farzaneh Karegar

LICENTIATE THESIS

Karlstad University Studies  |  2018:9

urn:nbn:se:kau:diva-66109

ISSN 1403-8099


ISBN 978-91-7063-933-3 (pdf)

© The author

Distribution:
Karlstad University
Faculty of Health, Science and Technology
Department of Mathematics and Computer Science
SE-651 88 Karlstad, Sweden
+46 54 700 10 00

Print: Universitetstryckeriet, Karlstad 2018

WWW.KAU.SE
Towards Improving Transparency, Intervenability, and Consent in HCI

Farzaneh Karegar
Department of Mathematics and Computer Science
Karlstad University

Abstract

Transparency of personal data processing is enforced by most Western privacy laws, including the new General Data Protection Regulation (GDPR) which will be effective from May 2018. The GDPR specifies that personal data shall be processed lawfully, fairly, and in a transparent manner. It strengthens people’s rights for both ex-ante and ex-post transparency and intervenability. Equally important is the strict legal requirements for informed consent established by the GDPR.

On the other hand, the legal privacy principles have Human-Computer Interaction (HCI) implications. People should comprehend the principles, be aware of when the principles may be used, and be able to use them. Transparent information about personal data processing should be concise, intelligible, and provided in an easily accessible form, pursuant to the GDPR. Nonetheless, the answer to the question about how HCI implications can be addressed depends on the attempts to decrease the gap between legal and user-centric transparency, intervenability, and consent. Enhancing individuals’ control in a usable way helps people to be aware of the flow of their personal information, control their data, make informed decisions, and finally preserve their privacy.

The objective of this thesis is to propose usable tools and solutions, to enhance people’s control and enforce legal privacy principles, especially transparency, intervenability, and informed consent. To achieve the goal of the thesis, different ways to improve ex-ante transparency and informed consent are investigated by designing and testing new solutions to make effective consent forms. Moreover, ex-post transparency and intervenability are improved by designing a transparency enhancing tool and investigating users’ perceptions of data portability and transparency in the tool. The results of this thesis contribute to the body of knowledge by mapping legal privacy principles to HCI solutions, unveiling HCI problems and answers when aiming for legal compliance, and proposing effective designs to obtain informed consent.

Keywords: GDPR, Informed Consent, Intervenability, Transparency, Usable Privacy
Acknowledgements

It is the most exciting part of doing a Ph.D., writing the acknowledgement. No scientific method, university rule or even a bad-tempered reviewer can affect the way you write. I am on my own to reveal my feelings about the half-way journey I traversed.

Foremost, in my opinion, education and learning have been the saviour of mankind through the history. Not all countries in the world provide their inhabitants with free education unconditionally, including mine. I have been privileged but it does not mean that I am not concerned about all the lovely, innocent girls and boys who are unjustly deprived of making a future for themselves. I hope soon I can significantly contribute to making their lives fairer and happier. Then it will be my greatest accomplishment.

I am grateful for Karlstad, a gloomy city in the mid of January 2015 which soon disguised to a picturesque, joyful one thanks to my beloved friend Susanne, the very first face I saw in Karlstad, who warmly welcomed me and made me feel home. Stories of my life in Karlstad moved me more than Tehran and changed me forever. I became stronger, more patient, more ambitious, and older while wiser and more hard working I suppose.

It introduced me the most wonderful, skillful, and supportive supervisor, Simone, I could ever have on the earth, even Mars or moon who trusted me and if it were not for her I would still be wandering around city thinking why I was really here. It gave me the awesome opportunity to work with a nice, knowledgeable co-supervisor, John Sören, and to work a year full of valuable learnings for me with Melanie Volkamer. It provided me with a nice working environment whose colleagues always cheered me up with their encouragement, helpful conversation, and calm, smiley faces. Thank you for all that you have done for me.

My gratitude and my feelings for my awesome, protective parents, my beloved, nice, supportive, kind Kyoomars, my sister and her cute, lovely Sobhan, my amazing brother-in-law, Ali, and my kind parents-in-law cannot be squeezed in words, written in a limited space in my non-mother tongue. I would simply like to say  "دوستتان دارم یا تمام وجود" and keep the rest for the moment we meet and hug.

I would also like to thank the European Commission for their support. This research has been partly funded by the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 653454.

Karlstad University, February 12, 2018

Farzaneh Karegar
List of Appended Papers

This thesis is based on the work presented in the following papers:


Some of the appended papers have been subject to few minor editorial changes.

Comments on my Participation

Paper I Simone Fischer-Hübner is the main author of the paper. I was a team member with Julio Angulo and Tobias Pulls for implementing the Genom-Synglig version of the Data Track and worked with Tobias Pulls for the latest stand-alone version of the tool. I conducted pre-user tests for the latest stand-alone version of the Data Track tool and contributed by reporting the results of that for this paper.

Paper II The idea of this paper, as well as the design of the user study, originated from a discussion with my supervisor, Simone Fischer-Hübner. I conducted all the interviews. Simone Fischer-Hübner assisted me in four of the ten interviews and improved and elaborated my text in the introduction, background, and related work sections. I was the sole author of other sections
and I received valuable comments from Simone Fischer-Hübner and Tobias Pulls who implemented the back-end of the tool and helped me in front-end development especially for map view.

**Paper III** Daniel Lindegren, our assistant in the CREDENTIAL project, and I prepared the test design for all three tests in the paper and the relevant test objectives together with my supervisors, Simone Fischer-Hübner and John Sören Pettersson. I wrote the main body of the test documentation for a project deliverable from which much information is derived for the paper, both concerning the first test and for the overall design. I also wrote the related work section. My co-supervisor, John Sören Pettersson, guided the assistants conducting the second test, and also outlined the paper and reduced the text from the deliverable to make a comprehensive and comparative text fits within the limits of a conference paper. My supervisor, Simone Fischer-Hübner, and I contributed to the structure and correctness of the text.

**Paper IV** I was the main author of the paper and I conducted all the 80 user studies. A discussion with my co-supervisor, Melanie Volkamer, led to the idea of this paper. Melanie Volkamer gave me feedback throughout the whole process and advised me in structuring the user study and writing. Nina Gerber conducted most of the statistical analysis based on what I described as hypotheses and variables defined to check them. Nina Gerber also contributed in writing the result section. My supervisor, Simone Fischer-Hübner, gave me valuable feedback and comments for all sections and contributed significantly in writing the introduction, analysis of legal requirements and conclusion sections.

**Other Publications**


I summarized and reported the outcome of the focus groups held at the IFIP Summer School (21-26 August, Karlstad, Sweden, 2016) to discuss research challenges concerning the trade-off between privacy, efficiency, and usability, end-user trust, and adoption factors identified for the CREDENTIAL project. I also wrote a section on why metadata privacy matters and discussed how metadata in the context of identity providers and data sharing might reveal more information, providing some examples.

Besides editing, I was the main author of two sections: i) task analysis of use cases in all three eGovernment, eHealth, and eBusiness domains, and ii) description of user interfaces (V1) and user evaluations in the CREDENTIAL project.
Contents

List of Appended Papers vii
Acronyms xv

INTRODUCTORY SUMMARY 1

1 Introduction 3
  1.1 Objective ........................................... 4
  1.2 Structure ........................................... 5

2 Background 5
  2.1 General Data Protection Regulation ..................... 5
  2.1.1 Ex-ante Transparency and Informed Consent .......... 6
  2.1.2 Ex-post Transparency and intervenability ............. 7
  2.2 Transparency Enhancing Tools/Technologies (TETs) ........ 8
  2.3 Single Sign-On Identity Management Systems ............ 9
  2.4 Research Project: CREDENTIAL .......................... 11

3 Research Questions 12

4 Research Methods 13
  4.1 Eliciting and Mapping Legal Requirements ............. 13
  4.2 User Experience Evaluation .............................. 14
    4.2.1 Usability Tests and Measurements .................. 14
    4.2.2 Expert Cognitive Walkthrough ..................... 17
    4.2.3 Questionnaires and Surveys ....................... 17
    4.2.4 Interviews ....................................... 18

5 Contributions 19

6 Related Work 20
  6.1 Previous Work on Ex-ante Transparency and Informed Consent 21
  6.2 Previous Work on Ex-post Transparency ................ 23

7 Summary of Appended Papers 26

8 Conclusion and Future Work 29

PAPER I:
Transparency, Privacy and Trust - Technology for Tracking and Controlling my Data Disclosures: Does This Work? 41

1 Introduction 43
# Table of Contents

1 Introduction 81
2 Background to the Present Study: the CREDENTIAL Project 82
3 Previous Studies 83
4 Evaluation Goals 85
5 User Test Design 86
   5.1 Recruitment of Participants 86
   5.2 Test Procedure 87
   5.3 Description of Interactive Tasks 88
6 Results 89
   6.1 Ease of Use and User Experience Metrics 89
   6.2 Expressed Preference for Data Selection 90
   6.3 Recall of Consents to Data Disclosure 90
      6.3.1 Mental Model and Preference for Authentication via an App 92
7 Discussion 93
8 Conclusions: A Way Forward for Identity Access Management GUI Design 95

**PAPER IV:**

Helping John to Make Informed Decisions on Using Social Login 101

1 Introduction 103
2 Requirements 105
   2.1 Literature Review, CW and Derived Requirements 106
   2.2 Legal Requirements 108
3 Proposed Solution 109
4 Methodology and Study Design 111
   4.1 Ethics, Recruitment & Demographics 113
   4.2 Study Design 114
5 Evaluation 116
   5.1 Tutorial 116
   5.2 UI 117
6 Further Findings and Discussion 119
**Acronyms**

**A4Cloud** Accountability for Cloud. 9, 23, 25

**CSS** Cascading Style Sheets. 17

**CW** Cognitive Walkthrough. 14, 17

**DPD** Data Protection Directive. 3, 6

**GDPR** General Data Protection Regulation. 3–8, 12, 13, 25, 26, 29

**GUI** Graphical User Interface. 4, 9

**HCI** Human-Computer Interaction. 4, 6, 13, 15, 19, 29

**HTML** HyperText Markup Language. 17

**IdP** Identity Provider. 10, 11, 17, 20, 23

**IUIPC** Internet Users’ Information Privacy Concerns. 18

**P3P** Platform for Privacy Preferences. 9

**PETs** Privacy Enhancing Tools. 8

**PISA** Privacy Incorporated Software Agents. 13

**PRIME** Privacy and Identity Management in Europe. 25

**PrimeLife** Privacy and Identity Management in Europe for Life. 9, 25

**SP** Service Provider. 3, 8–11, 20, 26

**SSO** Single Sign-On. 9, 10, 12, 17, 20

**SUS** System Usability Scale. 15, 16, 18

**TET** Transparency Enhancing Tool. 4, 5, 7–9, 12, 13, 19, 23, 25, 26

**UI** User Interface. 4, 13, 15–17, 25, 29, 30

**USEMP** User Empowerment for Enhanced Online Presence Management. 9
1 Introduction

Humans have always had a need for privacy and cherished it for their lives as Ovid, a Roman poet born in 43 BC, said: “Bene vixit, bene qui latuit.”¹ But the concept of privacy has evolved from the protection of one’s body and home into the direction of privacy of personal information. Nowadays, it is generally accepted that everybody has a need for privacy, although the way it is appreciated varies, depending on the culture and the individual’s characteristics. In 1890, the American lawyers Samuel Warren and Louis Brandeis described the right to privacy in an influential article [95]: it is the “right to be let alone”. Later, in 1967, Alan Westin defined privacy in terms of self-determination in his Privacy and Freedom [96]: “privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others”. This definition which described privacy as individuals’ control over their personal spheres was a new development.

In the digital world, an information asymmetry and power imbalance, however, exists between the entities who process personal data, such as Service Providers (SPs) and individuals who may be affected by automated decision-making, or failures in processing their personal data. Further, the distributed nature of the Internet makes it hard for people to keep track of where the information about them is stored, to whom it is dispensed, and for what purposes it is used [68]. Controlling the proper use of data, people should be able to get information on the recipients, how their personal data are processed, by whom, for what purposes, and possibly from which sources the data originated. Offering transparency for achieving this type of control plays a significant role in the protection of individuals’ privacy and since it is a social trust factor [4] it can also ameliorate people’s trust in SPs.

Transparent information has to be provided before, during, and after the data processing in order to ensure transparency. Thus, the concept of transparency comprises both ex-ante and ex-post transparency [44]. Ex-ante transparency entails guiding people’s decision-making process before the data disclosure and informing them about the potential implications of their choices. Ex-post transparency, on the other hand, provides people with relevant information about their data after disclosure including the information about which data are actually disclosed, processed, and stored by whom; and if the data processing is in accordance with the negotiated or stated policies. Transparency of personal data processing is also enforced by most Western privacy laws, including the EU Data Protection Directive (DPD) 95/46/EC [87] and the new General Data Protection Regulation (GDPR) [88] which will replace the DPD in May 2018. In GDPR, pursuant to Art. 5 (1) personal data shall be processed lawfully, fairly, and in a transparent manner. The GDPR grants enhanced rights to individuals for both ex-ante and ex-post transparency, intervenability, and consent, such as the right to access (Art. 15), the right to rectification and erasure (Art. 16, 17), the right to data portability (Art. 20),

¹To live well is to live concealed.
the right to object (Art. 21), and the right concerning automated individual decision-making (Art. 22). Generally speaking, transparency has the purpose of informing people about which of their data are received by whom, under which conditions, and what are the implications of data disclosure. Hence, it is of paramount importance that people obtain the correct mental models regarding the flow of their personal data. Benefiting from new technologies and avoiding information overload, one way to exert transparency and intervenability rights pursuant to GDPR is by using Transparency Enhancing Tools/Technologies (TETs). TETs are technologies, methods or tools which enhance transparency and can provide their users with more control over their personal data. TETs can be considered as tools providing insights into how personal data are collected, stored, processed, and disclosed in an accurate and comprehensible way.

The legal privacy principles, such as transparency principle, have Human-Computer Interaction (HCI) implications as they describe “mental processes and behaviour of the end-user that must be supported in order to adhere to the principle” [69]. Particularly, the transparency principle requires people to comprehend the transparency and control options, be aware of when they can be used, and be able to use them. Privacy laws including the GDPR define what should be provided legally for transparency. Also, the GDPR requires services to provide individuals with information about their personal data processing in a concise, intelligible, and easily accessible form, using clear and plain language. However, the answer to the questions about how fair processing can be achieved by meaningful transparency and how HCI implications should be addressed, for example, by the design of a usable Graphical User Interface (GUI), depends on the attempts to decrease the gap between legal and user-centric transparency, intervenability, and informed consent. We can argue, based on the research done by Pollach [72], Böhme and Köpssel [16], and Adjerid et al. [3] as some examples, that providing users with more transparency and control, without considering the disconnection between legal transparency and intervenability and the current practice in which users agree to almost every consent request, do not have the effect desired by lawmakers. Tsormpatzoudi et al. [90] emphasize the importance of involving end-users as stakeholders in the ‘Privacy by Design’ process, as the end-users should ultimately profit from ‘Privacy by Design’. Also, Cavoukian stresses that the ‘Privacy by Design’ principle ‘Respect for Privacy’ extends to the need for UIs to be “human-centred, user-centric, and user-friendly, so that informed privacy decision may be reliably exercised” [21]. Thus, besides aiming for legal compliance, different ways to provide usable and user-centric transparency, intervenability, and informed consent are required.

1.1 Objective

The objective of this thesis is to propose usable tools and solutions to enhance people’s control and enforce legal privacy principles, especially transparency, intervenability, and informed consent. To achieve the goal of the thesis, it
is studied that how GDPR legal privacy principles, including transparency, intervenability, and consent principles, can be enforced with the help of a user-side TET and different design approaches, particularly, in the context of identity providers. Consequently, this thesis aims to help people to exercise their rights pursuant to GDPR, gain control over and comprehend the flow of their personal data online, make informed decisions, and finally preserve their privacy.

1.2 Structure

This thesis presents an introductory summary and a collection of four papers in the area of improved transparency, intervenability, and informed consent which are either authored or co-authored by the author of this thesis. The remainder of the introductory summary is structured as follows: Section 2 provides the fundamental background for the work presented in this thesis. Section 3 outlines the research questions of this thesis and Section 4 discusses the research methodologies applied. The main contributions of this thesis are presented in Section 5 followed by the related work in Section 6. Section 7 provides a summary of appended papers. Concluding remarks and a brief discussion on future work in Section 8 ends the introductory summary.

2 Background

Forming the foundations for the terms and concepts used in this thesis, this section provides the necessary background to understand the topics under discussion. However, the background information in this section is not intended to be exhaustive or elaborate, but rather to provide the sufficient definitions for readers.

2.1 General Data Protection Regulation

GDPR [88] which will be effective from May 2018 is designed to coordinate data privacy laws across European countries, protect EU citizens’ data privacy, and transform the way organizations approach data privacy. Two terms from GDPR which are used repeatedly in this section are clarified here in the following definitions: i) Data Subject is an identified or identifiable natural person to whom data are related, and ii) Controller is the entity that determines the purposes and means of processing personal data.

In GDPR, transparency is an explicit requirement and a core principle of data protection: pursuant to Art. 5 (1) personal data must be processed fairly, lawfully, and in a transparent manner. Art. 12 (1) GDPR also obliges controllers to provide transparency: all information and communications concerning processing of personal data should be provided to data subjects “in a concise, transparent, intelligible and easily accessible form, using clear and plain language” [88]. This thesis places an emphasis on legal provisions for
transparency, intervenability, and consent that have HCI implications, and thus need to be addressed.

In the following, legal principles of transparency, informed consent, and intervenability pursuant to different recitals and articles in GDPR for which some solutions are proposed in this thesis are elaborated.

2.1.1 Ex-ante Transparency and Informed Consent

Ex-ante transparency is a precondition for data subjects to be in control and for rendering a consent (pursuant to Art. 7 and 9). Consent enables data subjects to authorise controllers to process their personal data. In GDPR, the definition of consent in Article 4 (11) expands the old definition of consent provided in the DPD and has stricter requirements regarding achieving informed consent. However, there is a burgeoning doubt regarding the effectiveness of informed consent in the context of personal data processing [3,16,83] and GDPR does not discuss what makes consent effective. In other words, although GDPR clarifies what should legally be provided to data subjects, which consequently can help them to make conscious and autonomous choices, it does not argue (and it is out of the scope of GDPR) to what extent data subjects are capable of providing informed consent, using legally compliant services. Actually, it is the role of ex-ante transparency tools and methods to address these issues, and help people to make informed consent.

Consent should be given by a clear affirmative action (Art. 4 (11) GDPR). According to Recital 32 GDPR, the affirmative action could include ticking a box, choosing technical settings or other statements which clearly indicate the data subject’s acceptance of the proposed processing of his or her personal data. Thus, implicit and opt-out consent and particularly silence, pre-ticked boxes or inactivity are presumed inadequate to confer consent. Given by a clear affirmative action, consent should also be informed, unambiguous, freely given, and specific. According to Opinion 15/2011 of Article 29 Working Party on the Definition of Consent [7] and Guidelines of Article 29 Working Party on Consent under Regulation 2016/679 [8], to be specific, the scope and the consequences of the data processing in the consent should be clear and precise. To be freely given, there should be no risk of compulsion, deception, intimidation or significant negative consequences if users do not consent. To be unambiguous, doubt is removed from the procedure of giving consent by individuals. In other words, there should be no ambiguity about the data subject’s intentions to give consent.

Finally, to be informed, GDPR obliges controllers to provide specific information to data subjects. There are a number of references to the Articles of GDPR and the Recitals that adds some insights about the informed term: pursuant to Art. 13 (1) GDPR and as stressed in Recital 42, when personal data are collected from data subjects, e.g. in the authorisation dialogue, individuals should be made aware at least about what data will be collected and used, the identity of the controller and the intended purposes of the processing of data. Furthermore, according Art. 13 (2) GDPR, the controller shall
provide the data subject with some further information to ensure fair and transparent processing. Such policy information includes, but is not limited to, information of recipients/categories of recipients, the period for which the personal data will be stored and information of the existence of rights for data subjects to withdraw consent at any time\(^2\), access and rectify data, and data portability, for example. Guidelines of Article 29 Working Party on Consent under Regulation 2016/679 [8] also list the minimum content requirements for consent to be informed as some crucial elements which are necessary to make a choice.

Apart from the legal principles for consent to be valid and informed, some researchers working on consent in the context of information privacy adopted the theory of informed consent [29] and considered how it can be applied in an online environment \([33, 34, 65]\). The most remarkable research study is that of Friedman’s et al. [33] model of informed consent in the context of online interactions. The model is aligned with the definition of informed consent in the GDPR. Friedman’s et al. model is based on six components: i) Disclosure, ii) Comprehension, iii) Voluntariness, iv) Competence, v) Agreement, and vi) Minimal Distraction. Disclosure and Comprehension assure consciousness and Voluntariness, Competence, and Agreement constitute the consent. In addition, the activities for informing users and helping them to give their consent should happen with Minimal Distraction.

2.1.2 Ex-post Transparency and intervenability

Besides some obligations about providing information prior to obtaining data from individuals, e.g. ex-ante transparency, controllers are also obliged to provide ex-post transparency. Ex post transparency is also a prerequisite for intervenability, one of the privacy protection goals. Intervenability “aims at the possibility for parties involved in any privacy-relevant data processing to interfere with the ongoing or planned data processing”, as defined by Hansen [40]. Data subjects should be aware of the existence of intervenability rights and how to exercise them. For data subjects, the right to rectification (Art. 16), the right to erasure (Art. 17), the right to restriction (Art. 18) as well as the right to object (Art. 21), the right to withdraw a consent (Art. 7 (3)), and the right concerning automated decision-making including profiling (Art. 22) are parts of the intervenability. In this thesis, an ex-post TET is designed which helps users to exercise their rights to access and data portability, which both can be used for enhancing transparency and intervenability.

The Right to Access (Art. 15) comprises the right to have information about the data being processed, data processing purposes, and data recipients or categories of recipients. The right to access extends the information to be provided by the controller to also include information about the data retention period, the right to lodge a complaint with the supervisory authority, and safeguards taken to transfer data to a third country. In addition, data

\(^2\) Being informed of the rights to withdraw consent at any time prior to giving consent is also mentioned in Art. 7 (3).
subjects shall be informed about the existence of automatic decision-making, including profiling, and at least in those cases, the logic involved and the consequences of data processing. The data subjects shall also have the right to obtain a copy of the personal data undergoing processing in a commonly used electronic format if they have made their requests by electronic means. This right is a prerequisite for exercising some other data subjects’ rights, because data subjects first need to be aware of their collected personal data and the corresponding processing to be able to revoke their consents, correct, delete or restrict the processing.

The Right to Data Portability (Art. 20) is the right to receive personal data concerning the data subject which are provided by the data subject to a controller and the right to transmit the data to another controller (or to have them transmitted directly from one controller to another one where it is feasible technically). Not being specific about the format, the received data should be structured, commonly used, and machine-readable to ensure interoperability. The objective is to prevent data subjects from being locked into privacy-unfriendly SPs by providing the opportunity to easily change the providers and transfer their data. The data that must be included in the file are the data provided by the data subject intentionally and directly, for example, submitted via online forms or any observed data which may, for example, include a person’s search history, traffic data, and location data. Such observed data are also actually provided by the data subject. Inferred and derived data which are created by the controller may be included depending on the context. Inferred and derived data will not be considered as provided by the data subject and thus not within the scope. However, there is no clear description of what can be considered as derived data in GDPR. It is worth mentioning that anonymous data and data which do not relate to the individual making the request will not be included in the file.

2.2 Transparency Enhancing Tools/Technologies (TETs)

The relation between Internet users and SPs who collect users’ information is typified by high information asymmetry [82]. Put differently, the proliferation of data and services online makes it hard and demanding for users to determine what data about them are actually collected, how the data are used by SPs and which extra information can be inferred from the collected data.

Targeting at reducing the information asymmetry, TETs increase the transparency for users in terms of more information, knowledge, and control over their data. Hansen [39] defined TETs as “tools which can provide to the individual concerned clear visibility of aspects relevant to these data and the individual’s privacy”. Therefore, while Privacy Enhancing Tools (PETs) aim at data-minimization TETs aim at providing users with insight into data handling behaviours. Needless to say, besides offering a good level of transparency, security and privacy of TETs should also be avouched to ensure that TETs are not used against end-users by their providers [76].

Different classifications of TETs based on parameters that can be com-
pared to choose the right TET in a specific context are available [41, 43, 49, 67, 98, 99]. For example, based on their execution environments TETs can be categorised into client-side, server-side and trusted third-party TETs [98, 99]. Server-side TETs like Google Dashboard allow authenticated users to receive information about collected, processed or forwarded data to which the services have access. In a third-party TET like the DataBait tool presented by Popescu et al. [73] as a part of USEMP project, the user trusts a third party to have the user’s data for providing transparency functions. DataBait derives numerical value indicators from individual data items and visualises these values using a web-based GUI. Client-side TETs like Mozilla’s Lightbeam3 make users’ data transparent which are stored locally on users’ devices under their control. From the usability point of view, client-side TETs may be more demanding to set up and their security is dependent on the security of users’ devices. However, they are theoretically the more privacy-friendly solution, as the users retain control over their data.

TETs can further be categorised based on the relation between the time at which they provide users with transparency information, and the time at which personal data are collected and processed by controllers [98]. Consequently, TETs can be classified as: i) ex-ante TETs like the Platform for Privacy Preferences (P3P) [23] and the PrimeLife Policy Language [6] which provide information to an end-user prior to data disclosure to a service provider, ii) ex-post TETs like A4Cloud Data Track [5, 14], Axiom’s AboutTheData portal4, Datacoup5, and DataSelfie6 which provide data to the user after personal data disclosure to a service provider, and iii) real-time TETs like some browser extensions including Mozilla’s Lightbeam, Ghostery7, and Privacy Badger8 which provide transparency during data collection and processing, for example, by providing users with a real-time visualisation of companies that follow them on the Internet.

2.3 Single Sign-On Identity Management Systems

Users are increasingly acquiring identities online. A study conducted by Florêncio and Herley showed that a typical user has 25 different identities, each of which has different credentials [31]. Encountering difficulties in managing these identities, users may select some countermeasures like password reuse which endanger their security. An attacker who is able to compromise one service provider can also compromise others protected by the same password which accordingly reduces overall security to that of the weakest site [25].

Single Sign-On (SSO) systems have been proposed and are being progressively used to implement authentication for modern SPs, also known as rely-

---

3https://www.mozilla.org/lightbeam/
4https://www.aboutthedata.com/
5https://datacoup.com/
6https://dataseffic.it/
8https://www.eff.org/privacybadger
ing parties, and applications. In SSO systems, authenticating to an Identity Provider (IdP) confirms the user’s identity for multiple SPs which in turn absolves SPs of the need to authenticate users themselves. Using SSO systems, SPs rely on authenticated identities to make authorisation decisions and it relieves users of remembering many sets of credentials.

SSO solutions provided by social networks, i.e. social logins, are widely deployed nowadays. Facebook, Google, and Twitter are three top English-speaking SPs that act also as IdPs [92]. In spite of differences in web SSO protocols, the flow of data from the user’s perspective is identical. A high-level process of the social login focused on OAuth protocol, which is adopted by notable IdPs, e.g. Facebook, is depicted in Figure 1, from the user’s perspective. The process is as follows: i) the user selects an IdP, ii) the SP redirects the user to the IdP to authenticate, iii) the user enters her credentials and signs into the IdP (this step is skipped if the user is already signed in), iv) on successful authentication, the IdP presents an authorisation dialogue, v) the user can see (and edit if possible) what information will be shared with the SP, and finally vi) the IdP redirects the user back to the SP and gives the SP an access token with which it can communicate with the IdP and retrieve information until the expiration date of the token. The user is now signed up for the SP.

Having social networking capabilities makes widely-deployed IdPs uniquely qualified to provide rich personal information about users. However, the convenience that using an IdP offers has a downside which is the privacy cost: i) IdPs may share some personal information with SPs that the SPs otherwise would not have known. This information is provided directly by users to the IdPs. ii) IdPs are not only able to see the personal information in plain-text but they also learn from metadata, for example, to which services and when their customers communicate.

Countering the first privacy concern, during the sign-up process social logins like Facebook show the users the types of personal information which will be shared with an SP. The types of personal information to be shared are typically displayed in an authorisation dialogue (consent form) lacking minimum transparency information for a consent to be informed such as clear data processing purposes in the first layer (see Section 2.1.1). Additionally, current authorisation dialogues of social logins often have unclear opt-out, instead of clear opt-in, choices which makes it difficult for users to conceive, notice and control the personal information they share. Previous studies show that users are not aware of the information they consented to share and the extent of access the SPs receive for their personal information [12,86]. Thus, users do not make informed consent since they should be fully informed as to what they are consenting when they use social login methods.

Solving the second privacy problem which is related to metadata, researchers suggest some cryptographic solutions summarized in [54]. However, the given techniques are not entirely suitable for real-world usage [54].

To what extent IdPs succeed at conveying the personal information that is shared with SPs and to the extent they give users the ability to make informed consent, and how it can be improved, are the questions that still need to be
answered.

Figure 1: Flow of data from the user’s perspective using social logins

2.4 Research Project: CREDENTIAL

The work done as a part of this thesis has been conducted within the scope of the research project CREDENTIAL. CREDENTIAL is an EU-funded Horizon 2020 project for developing, testing, and presenting cloud-based services to store, manage, and share personal information including sensitive digital identity information more securely than existing technology. Three different pilots in the eGovernment, eHealth, and eBusiness domains are developed within the project to demonstrate how the CREDENTIAL technology can be deployed in various contexts [45]. The central component of the tools developed within the project is the CREDENTIAL Wallet which offers a set of security and application services. The services in CREDENTIAL provide mechanisms such as authentication and authorisation and the services employ novel cryptographic technologies such as proxy re-encryption and malleable signatures.

Figure 2 shows a basic high-level view of the architecture of CREDENTIAL involving: i) the Wallet, a cloud-based service, which acts as an IDP and a data access manager, ii) a CREDENTIAL account holder (i.e. user) who owns the data stored in the Wallet and receives its services via an app on her device, and finally iii) a data receiver (e.g. an SP or another CREDENTIAL account holder) who gets access to the data stored in the CREDENTIAL Wallet or authentication assertions issued by the Wallet [45]. Upon a data request, e.g. from an SP, a re-encryption key derived from the user’s private key and the public key of the SP is created at the user side, and the user defines an access policy specifying which data may be disclosed. The SP will thus receive encrypted data compliant with the disclosure rules and is able to decrypt the data with its private key.

The user authenticates to her CREDENTIAL Wallet account via the CREDENTIAL Wallet app to manage her encrypted personal data stored in the Wallet. In this thesis, the user interfaces of the CREDENTIAL Wallet app are prototyped for authentication and authorisation processes. The ob-
jective of prototyping for user interfaces of the CREDENTIAL Wallet app is twofold: i) to evaluate people’s ability to use such an app, and ii) to answer general questions regarding people’s understanding of consent-giving to data-disclosure and the appreciation of more privacy-friendly SSO solutions than what is currently available.

![Diagram](image)

**Figure 2:** General architecture of the CREDENTIAL technology

## 3 Research Questions

This thesis is based on a general question regarding usable transparency, intervenability, and informed consent which is outlined in this section.

*How can legal privacy principles of the GDPR for transparency, informed consent, and intervenability be enforced in a usable way?*

This question is discussed in all appended papers. The two first papers discuss the solutions to enforce ex-post transparency using a TET called Data Track and how ex-post transparency can be employed to help people to exercise their rights. Two later papers, on the other hand, investigate how to enforce ex-ante transparency by improving consent forms (authorisation dialogues) which can help people to make more informed decisions. First of all, in Paper I, the rights of data subjects for transparency and intervenability pursuant to the GDPR are discussed and it is explained how these rights can be enforced using TETs with a special focus on different versions of a TET called Data Track. Targeting the newly introduced right of data portability, Paper II introduces a new version of the Data Track tool which visualizes exports of personal data and sheds a light on what people think about data portability and how it can be exercised using TETs. In Paper II, users’ per-
ception of ex-post transparency functions in the latest stand-alone version of the Data Track tool is also discussed. Paper III investigates users’ understanding of authorisation and authentication processes in the context of identity providers accessible via mobile apps. Paper III proposes an HCI solution to help users to pay more attention to the actual data they disclose by improving ex-ante transparency and consequently improving the quality of consent-giving. Finally, Paper IV focuses on two privacy decisions users should make in the context of identity providers. To facilitate users to consciously select a method to sign up, Paper IV proposes a tutorial. The tutorial aims to improve users’ knowledge about advantages and disadvantages of the methods to sign up. Further, Paper IV proposes solutions to design legally compliant authorisation dialogues (consent forms) which improve ex-ante transparency and significantly help people to give informed consent.

4 Research Methods

The methods employed in this thesis consider users through the development cycle of the proposed design solutions which to some extent follow a human-centred design [48]. In Section 4.1, the methods used to elicit and map privacy legal requirements to HCI solutions are described. Section 4.2 presents the user experience evaluation methods exploited to tackle the research question of this thesis and to evaluate the proposed design solutions. A brief discussion on the motivation for the choice of employed methods is also presented.

4.1 Eliciting and Mapping Legal Requirements

Legal privacy principles have HCI implications that are related to mental processes and human behaviours and can, consequently, be satisfied by HCI techniques. Legal privacy principles that shall be enforced by UIs of the sign-up process for social logins and legal privacy principles that can be enforced by TETs like the Data Track tool are elicited by reviewing the GDPR [88]. Mapping of legal privacy principles to HCI principles and proposed design solutions are partly based on the PISA project [69]. However, we used GDPR and extended the work conducted by Patrick and Kenny [69]. The core concepts in the HCI requirements mapping legal privacy principles can be grouped into four categories as explained in Patrick and Kenny’s work [69]: i) Comprehension which entails the requirements allowing a user to understand all the control options and rights related to her data, ii) Consciousness whose HCI requirements demands users’ awareness or their attention to controlling options and all rights related to their data, iii) Control which refers to the ability of users to exercise their rights after being aware of them and comprehend what to do, and iv) Consent which entails achieving the informed consent from users who must agree to privacy policies or terms and conditions, must fully comprehend to what they agree and the consequences that agreement may have.
4.2 User Experience Evaluation

There are three main categories of user experience evaluation methods: i) inspection, ii) inquiry, and iii) usability testing. An inspection such as Cognitive Walkthrough (CW) was proposed by Lewis et al. [60] in 1990 and has evolved ever since which links the interface walkthrough to a cognitive model [61]. The CW method focuses on the basic principles of usability and on the cognitive activities of users. Unlike CW, which is conducted without end-users, usability testing involves an evaluator observing participants interacting with an interface (i.e. completing tasks) to determine usability and users’ problems.

Inquiry methods entail gathering subjective input (e.g. preferences) from participants, typically through interviews [58], surveys [26], and questionnaires. These inquiry methods are often combined with usability tests [59].

In this thesis, an expert CW as an inspection method, usability tests and user studies combined with inquiry methods such as surveys, questionnaires, and semi-structured interviews are conducted. Utilising all of these methods, we collected data from users, some were qualitative and some were quantitative. We analysed quantitative data using quantitative evaluations based on analytics, i.e. statistical analysis. Qualitative data from the interviews and open-ended questions in questionnaires were analysed using qualitative evaluations including the identification of themes, coding, clustering similar data, and reducing data to meaningful and important points. Following in this section, more details about the methods used are provided and the motivations for the choice of the methods along with their strengths and limitations are elaborated.

4.2.1 Usability Tests and Measurements

Usability tests are conducted to determine if a system is suitable for the planned audience and for its predesignated purpose. Typically, the tests aim to uncover any difficulties encountered by users as they go through a set of predetermined tasks [27], which should be carefully selected to reflect realistic scenarios. When conducting usability tests, questionnaires, interviews, observations, potential system logs, and screen recordings are useful in gathering users’ opinions, understanding, attitudes, and feedback about the system. Careful considerations should be made, for example, for participants’ recruitment process, the material used, the standardisation of the test plan, the readiness of the evaluated prototype, and the way to ask the questions and describe the goal of the study in order not to bias the participants, especially in usable privacy and security studies. To evaluate the prototypes presented in Papers III and IV, usability evaluations are conducted following the guidelines in [78]. A typical usability evaluation session, in this thesis, usually consist of introducing the study with a cover story, obtaining participants’ consent, asking them to complete a pre-test questionnaire before tasks, asking participants to complete a set of tasks role-playing a persona using a prototyped UI, handing them a post-test questionnaire after the task, debriefing, and compensating.
the participants. A test moderator observed users doing their assigned tasks and took notes. Usability evaluations are selected as individuals’ feedback, attitudes, and perception regarding the proposed interfaces and concepts in question are required. Moreover, the usability evaluation is relatively simple and straightforward and it can be combined with other methods to complement it, such as questionnaires. The inclusion of other methods are also effective in inferring some conclusions from the participants’ verbal opinions and the observations made by test moderators. For elaborated descriptions of the user studies conducted in this thesis please refer to Papers I-IV.

Measuring Usability. Before conducting the usability tests, it is required to determine how to measure the usability. Usability is defined as “extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” in ISO 9241 210:2010 [48]. Effectiveness is defined as “accuracy and completeness with which users achieve specified goals”. Efficiency is defined as “resources expended in relation to the accuracy and completeness with which users achieve goals” and satisfaction is “freedom from discomfort and positive attitudes towards the use of the product” [47]. The major limitation of this standard as a quality model is that it is too abstract [84].

Hornbæk [46] summarized usability factors measured in HCI studies published in core related journals and proceedings. In literature, as shown by Hornbæk [46], prevailing measures of effectiveness are binary task completion, accuracy, recall, completeness, quality of outcome, and expert’s assessment. In the user studies conducted in this thesis, effectiveness is measured as the accuracy of doing the tasks and the recall of information. In other words, accuracy of doing the tasks measures if users have difficulties or any severe interruptions while doing the tasks and recall refers to measures of how much information users can remember after having used the UIs.

Efficiency is mainly measured by time (e.g. task completion time), input rate, mental efforts, usage patterns, communication efforts, and learning [46]. Fifty-seven percent of the HCI studies included in Hornbæk’s survey [46] measured time as task completion time. Task completion time is also used to measure the efficiency in usability tests of this thesis.

Finally, satisfaction is primarily measured by standard questionnaires, preferences, satisfaction with the interface, and users’ attitudes and perceptions. Although some studies prefer standard questionnaires for measuring satisfaction or build directly upon previous work for questions on overall users’ satisfaction, many researchers use their own satisfaction measures [46].

System Usability Scale (SUS) is originally invented by John Brooke [18] to evaluate practically any kind of system. The SUS questionnaire focuses on providing lightweight, ten five-point Likert scale questions, and subjective feedback from users [18]. Among the other validated questionnaires, SUS has been selected and used in hundreds of usability studies [10] and there are also a large number of reviews and evaluations of its effectiveness, strengths, and weaknesses [11,91]. In addition, Ruoti et al. [79] study the usability of seven
web authentication systems in which they utilise the SUS as a standard metric for empirical analysis. Ruoti et al. report that SUS produces reliable and replicable results and they recommend to formally evaluate the new authentication systems for usability using the SUS scale. Consequently, in the Paper III and Paper IV, the SUS questionnaire is used in the user studies to measure the overall usability and satisfaction in the context of identity providers.

**Cover stories.** Cover stories are used in this thesis to introduce the studies in which participants are asked to take part. The goal is to avoid participants from being primed for the actual objective of the study. Therefore, the real purpose is obfuscated ethically, both during the recruitment phase, and during interactions with participants in the study session. Participants tend to provide answers that they believe are satisfactory for the study, or pleasing to the moderator. Making cover stories can reduce this social desirability bias, as argued in [37]. On the other hand, some studies analysed the ethical considerations of deceiving test participants [85]. Still, Bortolotti and Mameli [17] argue that it is possible to use deceptive methods without causing serious harm to participants. In Papers III and IV, participants are not deceived in a way that they experience a completely irrelevant study comparing to the one to which they are introduced. However, the main objective is disguised regarding privacy and attention to data sharing, and it is presented as a study of testing the usability of a website which is actually a part of what participants accomplish.

**Role-playing.** In three of papers included in this thesis (II, III, IV), participants role-play a persona to complete their task(s). Participants receive some instructions about the task(s) and the persona they are required to role-play on a role-playing card. Using a persona has two main reasons: i) it allows a full control of what each participant encounters, providing a standard experience that can be compared between participants, and ii) due to ethical reasons, it helps to avoid handling sensitive participants’ information which has to be disclosed for the study such as birth dates or page likes on Facebook for Paper IV, and their locations over a period of time in Paper II. Although role-playing may affect the ecological validity of results it is not severely affecting the comparisons between the different test groups of participants (between-subject studies in Papers III and IV) as the premises remain the same. Also, in Paper II, the task defined for participants and the persona needed to complete the task serve as the starting point of discussing the interview questions. Role-playing the persona while doing the tasks, in paper II, provides all participants with a common ground and enables them to have a better insight into the meaning of downloading data from an SP and uploading the same data to another party. Nevertheless, the task is not used to measure the effectiveness of the UIs (i.e. Google archive managing interfaces in Paper II), or to time the participants for efficiency.

**Prototyping.** Prototyping enables designers to check their ideas with users and to gather feedback [74]. Prototyping is achieved using different tech-
niques, tools, and materials, ranging from paper, pens, and cardboards to wireframes and more advanced programming languages [19]. Prototypes are classified based on their levels of complexity and detail (e.g. paper-based lo-fidelity vs. some computer-based hi-fidelity prototypes) [93]. Some research compares user testing with low and high-fidelity prototypes and shows that low-fidelity prototypes are also good at uncovering usability issues [80, 93]. Walker’s et al. usability testing results also demonstrate to be independent of medium, despite differences in the interaction style [93]. Consequently, in this thesis, the medium and the level of fidelity are selected based on what suits the practical needs and design goals for prototyping. In this thesis, several hi-fidelity prototypes have been created and tested with real users under varying conditions. The prototype of the latest stand-alone Data Track tool described in Paper II is a hi-fidelity completely interactive prototype implemented to run in a browser using HTML, CSS, JavaScript and some JavaScript libraries (e.g. Leaflet for interactive maps) that evolved over various incremental and evolutionary iterations of pilot tests. Different interactive prototypes created with wire-framing tools are also used to test the design concepts presented in Papers III and IV.

4.2.2 Expert Cognitive Walkthrough

Cognitive Walkthrough (CW) is an expert review method in which interface experts imitate users, walking through a series of tasks [59]. In Paper IV, different tasks are defined considering various types of users to find as many user and interface problems as possible for current SSO interfaces of Facebook. Two of the paper’s authors (Farzaneh Karegar and Melanie Volkamer) worked as a group on the current user interfaces of Facebook SSO to identify usability and potential user problems. The detected problems in the CW and the literature review are used to elicit requirements which should be considered when designing for sign-up Uls in the context of IdPs.

4.2.3 Questionnaires and Surveys

Surveys are fundamentally used to ask well-defined and well-written set of questions, of a sample of people from a population, to explain, explore, and describe that population [32]. Surveys are rather easy methods to collect data from a large number of people cheaply, however, they cannot provide deep and detailed data. Although interesting concepts may appear in users’ answers, it is not possible to ask follow-up questions which can investigate more about the concepts in surveys. Consequently, careful considerations are required for inappropriate, ambiguous or biased responses. Moreover, questions, both open-ended and close-ended, should not be redundant, biased and incomprehensible and the number of questions should not cause respondents’ fatigue [42]. The terms surveys and questionnaires are often used interchangeably. Nevertheless, some people differentiate between them and define the questionnaires as a list of questions and surveys as a complete methodological approach including questionnaires as one of their elements
besides, for example, sampling and incentives [59]. This thesis does not differentiate between the survey and the questionnaire terms. However, this thesis differentiates between self-administrated, electronic questionnaires and surveys and moderator-supervised questionnaires and surveys combined with usability tests and user studies to collect more data from users and help to avoid misinterpreting questions as participants have the opportunity to interact with the moderator.

In this work, moderator-supervised surveys and questionnaires are used in combination with the user studies performed in Papers III and IV in order to capture the participants’ opinions of various aspects of the experiment. Using existing scales which are examined and re-validated by other researchers is a good academic practice, advances the state of art, and disengage the researchers from developing their own measurement instruments [75]. Both existing questionnaires (e.g. SUS, see Section 4.2.1, and IUIPC questionnaire [62]) to collect data on subjective topics and researchers’ own questions to collect data on more objective topics are used in Papers III and IV.

Internet Users’ Information Privacy Concerns (IUIPC) questionnaire [62] is used in Paper IV to measure the participants’ privacy concern level and investigate if it has any effects on providing better-informed consents. IUIPC is developed to measure people’s general concern about information handling practices of organisations. Malhotra et al. [62] introduce a ten-item scale of IUIPC, which represents the dimensionality of privacy concerns, categorised as collection, control, and awareness [62]. IUIPC is a statistically valid instrument [20] and has been used extensively by researchers in statistical models. Also, Preibusch has highlighted IUIPC among other survey instruments for measuring privacy concerns because of its influence on other works and its maturity [75]. However, the IUIPC only measures an individual’s general concern across a restricted set of dimensions [66] and it suffers from straightlining and the absence of reverse-coded items [75].

4.2.4 Interviews

Contrary to surveys which are extensive but not profound, interviews can help to reach deep answers. Interviews are highly flexible when it comes to their structure. At one end, there are fully structured interviews with a firm script to present questions in a predefined order. At the other end, there are unstructured interviews which may be simply based on a list of topics or questions to guide the interview [59]. However, if researchers demand more clarifications and they want to avoid challenges involved in the interpretation of data collected from unstructured interviews and skills required for conducting them, they can use semi-structured interviews. In semi-structured interviews, some planned questions are asked and other questions may emerge based on participants’ answers and comments. A general guideline for conducting interviews and analysing collected data are presented in [35, 64].

In Paper II, in order to learn more about users’ perceptions of transparency and data export, ten participants are asked to answer different questions in semi-structured interviews. Guest et al. [38] operationalise saturation
and make evidence-based recommendations regarding non-probabilistic sample sizes for interviews. Guest et al. report that saturation occurs within the first twelve interviews, although basic elements for meta-themes are present as early as six interviews [38]. Guest et al. discuss that as long as the aim is to understand common perceptions and experiences among a group of relatively homogeneous individuals, twelve interviews should suffice.

In Paper II, the interview method is selected because it is required to investigate about users’ perceptions individually in detail. One of the shortcomings of the interviews is that collected data are separated from the tasks and the context under question and it suffers from the problems of recall. However, to reduce the effects of this drawback the semi-structured interviews are combined with other techniques. In Paper II, participants are given a task to complete and freedom of working with the Data Track tool while answering the questions. During the interview sessions of Paper II, a moderator observed participants, a note-keeper took note, and the screen was recorded. Notes are compared with corresponding screen recordings to reduce the observer bias and ensure the accuracy of data, which are analysed by grouping the codes showing similar concepts.

5 Contributions

This thesis contributes to the body of knowledge for bridging the gap between the legal transparency, intervenability, and informed consent principles and user-centric transparency and control. This general contribution is reflected in various partial contributions made in Papers I-IV, listed as follows:

Mapping legal privacy principles to HCI solutions. Legal privacy principles of GDPR for transparency, intervenability, and informed consent are discussed because they have HCI implications for users. It is shown (to a certain extent, see Section 8 for the future work) that how these legal privacy principles can be enforced by TETs with a special focus on a tool called Data Track (in Paper I) and by designing legally compliant authorisation dialogues (in Paper IV).

Unveiling HCI problems and solutions when aiming for legal compliance. Summary of results of the user studies conducted on Data Track versions raised some HCI challenges that remain to be addressed, as reported in Paper I. In particular, although users appreciate the transparency functions available in the Data Track tool, their perception and understanding of the flow of visualised personal data are challenges requiring further research.

Paper II investigates users’ perception and preferences of transparency of data exports and the concept of data portability. It confirms the problem of users’ perception of control over their data and understanding of locally and remotely stored data, even with exercising their right to data portability. As reported in Paper II, benefits of the right to data portability and the usage scenario are unclear for participants. Nevertheless, when informed they express their positive attitudes towards the stand-alone version of the Data Track to
Designing effective authorisation dialogues. Prototypes of user interfaces for authorisation dialogues on both mobile (in Paper III) and desktop (in Paper IV) are developed and tested. Paper III shows that users have a tendency to maintain control over the data which are requested to be shared. Studies conducted in Paper III reveal that users prefer to select even mandatory information themselves and not to have them selected by default. This tendency shows that speed is not always users’ first priority and can be explored as a way of slowing users down and have them reflect more. It is also aligned with the legal requirements of consent that is valid if given by an affirmative action. The results of the user study in Paper III show the potential of a confirmation screen to contribute to the improvement of users’ recall of what they shared. In Paper IV, transparency is improved and users are actively engaged in interacting with authorisation dialogues using drag and drop data selection method and interactive knowledge testing and feedback. The newly proposed interfaces of authorisation dialogues for social logins significantly help users to make better-informed decisions and decrease the level of uncertainty compared with the current practice of social logins.

6 Related Work

This section briefly describes the works related to improved transparency, both ex-ante and ex-post, and informed consent, complementing the description of the related work in the appended papers with updated information. In this section, it is also briefly shown how this thesis has advanced the state of the art.
6.1 Previous Work on Ex-ante Transparency and Informed Consent

As described in Section 2.1.1, while it should be freely given by affirmative actions, precise, and unambiguous the consent is just valid when it is also informed. Transparency of personal data processing and necessary information are usually provided via privacy notices which are supposed to help people to make informed privacy and consent decisions. In other words, privacy notices are public announcements of data practices involving personal information [81]. Privacy notices and policies have been used in practice and in literature interchangeably. Privacy policy, in this section, means the exact privacy principles, practices of data controllers, and information which should be provided to users. On the other hand, the term privacy notice is used to mention different channels and means via which those principles and information are communicated to data subjects. Consent forms and authorisation dialogues presented to users for the first time they want to use a service (e.g. Facebook social login dialogues or mobile app permission dialogues) are examples of privacy notices.

Problems of Consent. Users no longer make active, informed choices when confronted with a consent situation [24, 63]. Thus, in practice, most privacy notices and choice mechanisms are ineffective at informing data subjects [24, 63]. Instead, users simply give consent when consent is requested. Consequently, data controllers can also no longer trust that consent is doubtlessly a consent [83]. Schaub et al. [81] survey the existing literature on privacy notices and identify challenges and requirements for privacy notice design. Schaub et al. [81] also provide the design space for privacy notices. Schaub’s et al. taxonomy considers four main, general dimensions that can be leveraged in the design of notices: i) timing for notice appearance, ii) the channel via which notices are delivered, iii) modality of notices to communicate privacy, and iv) the control options they provide. As discussed by Schaub et al., the ineffectiveness of privacy notices originates from some impediments as follows: i) notice complexity due to different roles of privacy notices which aim at both fulfilling legal requirements and simultaneously informing users, ii) lack of choices to opt out certain practices such as sharing data for marketing purposes, iii) notice fatigue due to frequent exposure to apparently unrelated privacy notices, and iv) decoupled notices which usually happen in the context of novel systems, such as wearable devices.

Schemer et al. [83] also describe the reasons why users rarely read notices, do not comprehend them, however, consent to the processing of their personal data. Similar to Schaub’s et al. work [81], Schemer et al. also report information overload, consent fatigue, and absence of meaningful choices as obstacles for effective privacy notices [83]. Further, Schemer et al. argue how stricter legal requirements with more focus on the autonomous choice of data subjects for giving consent weakens the effectiveness of the consent mechanism. Schemer et al. criticize the too much emphasis placed on the
active and informed role of people and autonomous authorisation in data protection [83]. On the other hand, Coles-Kemp and Kani-Zabibi [22] in their article about online privacy and consent show that service users would like to be active users of a dialogue system and not be passive actors in the management of privacy and consent in online transactions. Schermer’s et al. proposed alternatives for informed consent [83] have some strong problems: in fair transaction model of consent, legal requirements for giving and obtaining consent can be relaxed only if there is a common understanding among different users in different societies on i) what actions or inactions constitute consent, and ii) what is fair use of personal data. Both are very context-sensitive and very much dependent on the characteristics of different users which make the model difficult to be implemented in the real world by data controllers. In addition, before considering any changes in a consent model which may bring new problems, more research is required to investigate usable solutions and methods to actively and effectively involve people in legally compliant interfaces providing appropriate choices that are along with users’ expectations, and can fulfil their satisfaction of new, enhanced privacy rights pursuant to regulations.

**Proposed Solutions.** Researchers have proposed different solutions to solve the current issues with consent and privacy notices. Research focusing on the usability issues of privacy notices, improved interfaces of notice and useful suggestions and alternatives abound [9, 36, 52, 55]. Common proposals to improve the usability of privacy notices are the use of multi-layered notices [8] or just-in-time notices [69]. The proposed solutions, although important for solving information overload and notice complexity, cannot be a remedy for the lack of choice and consent fatigue. Furthermore, some research such as Acquisti’s research focuses on nudging users into making decisions about privacy when confronted with a consent request [1, 2]. Solutions provided by nudging users focus on helping user’s privacy and security choices with soft paternalistic interferences that allude users to more advantageous choices [2]. Although this approach may reduce problems of cognitive ability and consent fatigue by less obtrusive consents, it may not be effective in today’s scenarios where users have no choice other than consenting to all data processing purposes to use the service.

In particular, in the context of social logins, Egelman [28] emphasizes the fact that due to habituation, people do not pay attention to the exact content of the consent forms and have some preconceptions about what is shared. In Egelman’s study, no statistically significant differences are observed between conditions based on how the data are presented to participants (i.e. the control, list, or verbatim conditions); and participants who passively receive information which is presented in different formats fail to notice the changes made to the consent dialogues. A proposal from Wang et al. [94] suggests new interfaces based on the limitations of Facebook authorisation dialogues at that time. However, the extent to which the users may understand and pay attention to what is actually shared using the proposed new interfaces is not
evaluated. Javed and Shehab [50] investigate the effects of animated authorisation dialogues for Facebook. Another proposal by Javed and Shehab [51] utilise eye-tracking techniques in order to force users to read the permission dialogue but they do not report about the prices users may pay in terms of time and satisfaction. In Papers III and IV of this thesis, with the focus on permission dialogues and consent forms in the context of IdPs, the effectiveness of informed consent and privacy notices is improved by adjusting active and conscious role of users in interacting with dialogues. Papers III and IV, contrary to previous studies in the context of IdPs, investigate the effects of actively involving users in consent forms along with fulfilling legal requirements. Besides the reported effectiveness of proposed solutions on improving users’ awareness of data sharing, both Papers III and IV present the effects on the time to finish the relevant users’ tasks and their satisfaction.

6.2 Previous Work on Ex-post Transparency

Different Data Track versions, including the latest stand-alone version, described and studied in the first two papers for this thesis are just one group among many other ex-post TETs designed, implemented, and sometimes evaluated in various research. Recently, Murmann and Fischer-Hübner [67] conducted a survey on ex-post TETs in which reviewed ex-post TETs at least present implementation in a prototypical stage or in an evaluated mock-up. As discussed by Murmann and Fischer-Hübner, the underlying usage contexts of the reviewed TETs in [67] follow two patterns. The first category of TETs supports users in detecting and analysing a small number of different disclosed personal data items in a specific, clearly defined usage or application contexts, e.g. in the context of a location-based service or search queries. In contrast, the second category of TETs handles data as abstract information, tries to visualise the data in the best possible way and does not have a particular usage context. Different Data Track versions discussed in Papers I and II belong to the second category. Respectively, in this section, the ex-post tools listed in [67] oriented towards the visualisation of generic personal information, are discussed with two exceptions. i) Kani-zabibi’s et al. work [53] is considered generic rather than specific as what they propose is easily extendable to other stakeholders. ii) Riederer’s et al. work [77] is considered specific rather than generic as it receives the location data from online services, visualises it, and finally predicts users’ demographics based on their locations by comparing with census information (e.g. income and race derived from United States census). Consequently, seven different ex-post tools [15, 53, 56, 57, 71, 89, 97] along with Angulo’s et al. work [5] on the A4Cloud version of Data Track are selected among which just four conducted user studies [5, 15, 53, 57]. The user studies include either pre-studies before or during the implementation phase, or usability tests conducted during or after the implementation phase.

Previous Work on Ex-post Tools. In 2009, Kolter et al. [56] introduced a user-centric, provider-independent privacy architecture employing a collabo-
rative privacy community to share and exchange privacy-related information, ratings, and experiences among Internet users. This architecture enables Internet users to control the disclosure and management of personal data by a browser plug-in. The representation is mainly in text and includes both implicitly and explicitly disclosed data. Later, in 2010, Kolter et al. [57] introduced a usable solution for the visualisation of past personal data disclosures within a user-centric privacy architecture. This new proposal extends one of the three local privacy components of the provider-independent privacy architecture that was first introduced in [56]. Based on a browser extension that detects and stores personal data submitted to service providers, the presented visualisation tool displays the resulting transaction log in multiple views and perspectives: i) A basic view based on recipients of transferred data, ii) the chronological view allowing for a temporal analysis of past personal data disclosures, and iii) a flexible graph view facilitating the visualisation of various relationships between entities. The visualisation includes implicit, explicit and derived data. A user test conducted on the Kolter’s et al. proposal demonstrates the usability and the user acceptance of the solution [57].

Trabelsi et al. [89] propose a cloud service (called SPACE) that provides all the tools to a data subject to impose her privacy preferences during the virtualisation phase. The design of Trabelsi’s et al. solution was inspired by lack of transparency and data control in different commercial cloud infrastructures. SPACE is based on the sticky policy technology [13] and offers access and usage control functionalities to the data anywhere in the cloud. The different modes of presentations (e.g. access history map and geolocation of data) and control functionalities make the user aware of the storage condition of her explicit private information in the cloud.

Kani-Zabibi et al. [56] developed an interactive social translucence map that visualises the flow of explicitly disclosed personal information. The user can highlight the flow in the presented graph per data category. The user study conducted in [56] evaluated the usability of the prototype and its effect on users’ privacy awareness, understanding, and attitude. Findings show that privacy features are welcomed and necessary to enable users to manage their privacy concerns but some groups need further support by social and institutional privacy management processes.

Zavous et al. [97] introduce a tool called Cloudopsy, a service with the goal to provide a visualisation of the exchange of user implicit and explicit data in the cloud premises. Cloudopsy offers an interface based on the graphs produced with the Circos visualisation tool to the customers of the cloud services to independently monitor and get a better understanding of the processing of their sensitive data disclosed implicitly and explicitly to cloud services. As discussed by [97], one of the benefits of the proposed visualisation is the ability to better understand ongoing events, detect anomalies, and reduce decision latency. But the proposed solution is not tested in a user study.

Pistoia et al. [71] present Labyrinth, a run-time privacy enforcement system for the mobile environment that automatically detects leakage of private data originating from standard and application-specific sources. The
Labyrinth allows for visually configuring the application-specific sources of confidentiality, through which implicit and explicit private data enters the application. Visualisation is textual and the results on evaluating Labyrinth experimentally have only been reported.

Bier’s et al. [15] work which is the closest study to the A4Cloud version of the Data Track [5] presents a privacy dashboard called PrivacyInsight considering some legal and usability requirements. PrivacyInsight allows i) automated collection and processing of the explicit, implicit, and derived personal information, ii) visualisation of personal data, information flow, and selective views on data of interest, iii) customizable depth of information, and iv) immediate exercise of the data subject’s rights of control and intervenability like the right to rectify and erase. PrivacyInsight is evaluated in comparison with the trace view of the A4Cloud version of Data Track (GenomSynlig) [5] and a JSON document representing the common response from companies. The results of the user study conducted in [15] show that PrivacyInsight is the most usable of the privacy dashboards included in the Bier’s et al. study with more complete features.

Finally, a brief overview of different Data Track versions is presented. The first version of the Data Track is developed within the PRIME project and includes a history for each transaction storing, among other information, the disclosed personal data, the recipients of the data, and purposes of data disclosure [70]. PRIME Data Track is augmented in the PrimeLife project with online access functions allowing users to exercise the right of control and the right to intervene [30]. Angulo et al. [5] improve the UIs and interaction concepts of next versions of Data Track implemented in the A4Cloud project, e.g. by replacing the tabular presentations of the PrimeLife Data Track with two graphical UI illustrations: i) trace view, and ii) timeline view. Prototypes of the A4Cloud Data Track have been evaluated with usability tests and two focus group sessions in a workshop [5]. Evaluations reveal that while the majority of test subjects consider Data Track to be a potentially useful tool with appreciated transparency options, many participants have difficulty differentiating the client-site from the service-side accessible via the UI when exercising their rights.

The implementations presented by Bier et al. [15] and Angulo et al. [5] work similarly regarding how disclosed personal data are visualised. Among all existing ex-post tools with the general usage context which are briefly described in this section, Data Track presented by Angulo et al. [5] and PrivacyInsight presented by Bier et al. [15], display derived data when technically feasible. Additionally, these two works discuss GDPR legal privacy rights and enable data subjects to exercise their legal rights to rectify and erase data.

**Latest Stand-Alone Version of the Data Track.** Aiming to solve the problem of confusion about locally and remotely stored data reported in [5] and investigating the users’ perception of visualising data exports in a TET with a special focus on the new data subject’s right of data portability, the latest version of the Data Track has been developed at the end of the A4Cloud
project. The latest stand-alone version of the Data track is an open source program presented in Paper II. The latest stand-alone version provides users with the visualisation of Google location history exported from the Google archive managing service, as a proof of concept. After successfully exporting the location data from Google and importing it to the Data Track, participants see a newly developed map view that allows visualising location, activity and movement patterns as described in the location history file provided by Google. Activities are data derived by Google, based on the locations reported by users' devices. Contrary to the previous user studies of the Data Track, Paper II of this thesis presents the first evaluation of users' perception and understanding of a TET based on exports of personal data from an SP. Moreover, Paper II investigates the perceived value of the latest version of Data Track when users exercise their rights to receive an electronic copy of the personal data or data portability, pursuant to GDPR.

7 Summary of Appended Papers

Paper I - Transparency, Privacy and Trust - Technology for Tracking and Controlling my Data Disclosures: Does This Work?

Transparency is a basic privacy principle and a social trust factor. However, in the age of cloud computing and big data, providing transparency becomes increasingly complex. In this paper, we discuss legal privacy principles pursuant to the GDPR providing transparency and intervenability for users and how these principles can be enforced by TETs and, in particular, by the Data Track tool. We also elaborate on open research challenges from a Human-Computer Interaction (HCI) perspective. We show that functions developed in different Data Track versions are complementing each other because they address several legal privacy requirements of the GDPR enabling ex-post transparency and intervenability. The users' perception of control and security in Data Track remains a challenge to be tackled besides promoting users' trust. In the future, we would like to tackle the potential consequences of big data profiled by service providers improving transparency. In particular, we are interested in analysing how the right to data portability and the right of access together with the right of being provided with the meaningful information about the logic involved in profiling, can enable citizens to aggregate their data and to infer and understand what service providers, including government agencies, might deduce from them via profiling as well as the possible consequences.

Paper II - Visualizing Exports of Personal Data by Exercising the Right of Data Portability in the Data Track - Are People Ready for This?

A transparency enhancing tool called Data Track has been developed at Karlstad University. This paper reports on a user study investigating the perception of a new function visualising exports of personal big data for the data subjects, which we added to the latest stand-alone version of the tool.
Towards Improving Transparency, Intervenability and Consent in HCI

27

Analysing the users’ perception of Data Track and its transparency features, and the concepts of data export and data portability, we conducted a qualitative user study in which users experienced the latest stand-alone version of the tool. We observe that although users have little interest in the visualisation of derived data activities revealed in the Google location file, they are interested in other kinds of derived data like movement and travel patterns, usage patterns for different service providers, statistical data based on their behaviours, and information about to whom their data are disclosed, how it is exchanged, and how they might receive related advertisements. Also, we confirm that it is confusing for users to differentiate between locally and remotely stored and controlled data. Finally, in spite of being concerned about the security of the data exported to their machines, for exercising data portability rights pursuant to the GDPR, most participants prefer to first export and edit the data before uploading it to another service provider. Users appreciate Data Track for helping them in this context. Clearly users would like to be in control when exercising the right of data portability and they do not prioritize the convenience of having their personal data transmitted directly from one controller to another when it is feasible technically. In the future, we would like to extend the tool to visualise data exports of other service providers and to expand its functionality to support users in all the steps involved in exercising the right of data portability.


CREDENTIAL is an EU-funded Horizon 2020 project that involves developing, testing and presenting cloud-based services to manage digital identity information and personal data with a higher level of security than existing technology. The CREDENTIAL Wallet is the central component of the tools developed in the project and it supports users with its functionalities in a mobile application acting as an identity provider and a data access manager. The user interfaces of this app can be used to evaluate not only people’s ability to use such an app, but also for important general questions concerning people’s understanding and preferences of consent-giving to data-disclosure and their appreciation of more privacy-friendly single sign-on solutions than what is presently offered to the general public. In this paper, we conducted usability tests for the mock-up of the CREDENTIAL mobile app as an identity provider. The goal of conducted studies is twofold: to assess 1) users’ consciousness of data disclosures and flow of data in authorisation dialogues (consent forms), and 2) users’ understanding of authenticating to service providers and authorising service providers to access personal data in the context of identity providers accessible via mobile apps. The study covers a set of three user tests made of the core functions of authorisation and authentication. Results show that using one’s fingerprint for giving consent is easy, but most participants do not have a correct view of where the fingerprint data are used and which entities may have access to it. Familiarity with identity apps appears to aggravate misunderstanding. The comparison of BankID users with
others indicates that trust is trust in an actor (bank), not in the system, which can possibly lead to a skewed mental model of what should be trusted. This calls for educating the general public, besides designing usable user interfaces. In addition, participants could not easily recall details of personal data releases and settings for disclosure options. An evaluation with a confirmation screen which improved recall rate suggests that a confirmation screen can be a default option in authorisation (consent) dialogues. Most of our participants voiced a desire to have control over their data and expressed a wish to manually select mandatory information. This can be a way of slowing users down and make them reflect more. But effective ways to unobtrusively slow them down to reflect more are subject to the future work.

Paper IV – Helping John to Make Informed Decisions on Using Social Login

Users need to make two privacy-related decisions when subscribing to a new web service: (1) whether to use an existing Single Sign-On (SSO) account of an Identity Provider (IdP) and (2) the information the IdP is allowed to share, and for how long, with the Service Provider (SP). From a privacy point of view, the use of existing social network-based SSO solutions (i.e. social login) is not recommended. This recommendation, however, is accompanied by drawbacks regarding security, usability, and functionality. Thus, in principle, it should be up to the user to consider all advantages and disadvantages of using SSO and to consent to requested permissions, provided that she is well informed. Another issue is that existing social login sign-up interfaces are often not compliant with legal privacy requirements for informed consent and ‘Privacy by Default’. Accordingly, our research focuses on enabling users to make informed decisions and consent in this context. To this end, we identify users’ problems and usability issues from the literature and an expert cognitive walkthrough and we elicit end-user and legal privacy requirements for user interfaces (UIs) providing informed consent. We use this input to develop a tutorial for informing users about the pros and cons of sign-up methods. We also use this input to design SSO sign-up UIs for enabling informed consent, following the approaches of human-centred and privacy by design by addressing end user-specific and legal privacy requirements from the beginning and throughout the UI development cycle. We test both the tutorial and the UIs in a between-subject laboratory study with 85 participants. The results indicate that the level to which users are informed when deciding and providing consent in the context of social login is increased. Particularly, the tutorial notably helps users to improve their knowledge about the advantages of options they have for sign-up, however, more investigations are required to ideally communicate the advantages and disadvantages of services that may threaten the users’ privacy. For our newly developed UIs, informed consent is enforced with the help of an active involvement of users via drag and drop and question and answer methods. A between-subject user study evaluation of the effects of our UIs show that the new UIs are significantly more effective in helping users to provide informed consent than the current authorisation
dialogues of the social network. In conclusion, affirmative actions such as 
drag and drop requiring users to carefully check opt-in choices to be made as 
well as interactive knowledge testing and feedback are examples of effective 
HCI concepts for informed consent UIs.

8 Conclusion and Future Work

From May 2018, GDPR will be effective and services should offer enhanced 
and new legal privacy rights including transparency and control rights to peo-
ples. However, it is of paramount importance that privacy principles be ap-
plied in a way that accounts for HCI implications related to mental processes 
and human behaviours. This thesis examines how ex-ante and ex-post trans-
parency, informed consent, and intervenability can be improved in a usable 
way by designing and developing transparency enhancing tools and employ-
ing new design solutions, while considering users as stakeholders.

To improve ex-post transparency and intervenability, an ex-post trans-
parency enhancing tool is designed, which visualises exports of personal data. 
Users’ perception of data transparency and portability in the Data Track, the 
designed tool, are investigated. Users cannot differentiate between the data 
that are stored server-side and the data on their machines. Although users ap-
preciate the role of Data Track as an intermediary tool providing them more 
control over their data while exercising the right to data portability, they are 
more interested in the visualisation of behavioural information services col-
lect about them. Moving forward towards improving ex-post transparency 
and intervenability, future research will continue to map the users’ right to 
intervene in the processing of their data to HCI solutions. In particular, a spe-
cial focus of the future research will be on proposing and testing usable tools 
that provide algorithmic transparency, inform users about potential biases in 
the context of automated decision-making, and support them to exercise their 
intervenability rights, including the right to object, pursuant to the GDPR.

To improve ex-ante transparency and consent, new design solutions are 
prototyped and tested to make effective consent forms in the context of iden-
tity providers considering both user and legal requirements. Results of the 
studies conducted on the designed consent forms of a mobile app for a cloud-
based identity management show that confirmation screens can slightly help 
users to recall better what they shared. Further, results of user studies con-
ducted on the new legally compliant UIs for the sign-up process of social 
logins demonstrate that new affirmative actions for selecting personal infor-
mation such as drag and drop help users significantly to recall what they share 
and reduce the level of uncertainty about disclosed personal information. 
Equally important is the question and answer method employed in the new 
UIs which helps users to make informed consent. Although many requests 
for information sharing in consent forms are blocking (i.e. users cannot opt 
out mandatory information if they want to use services) this thesis shows 
that trying to actively involve users can capture their attention to what they 
disclose and prevent them from repeated clicking without paying attention.
However, utilising *drag and drop* and the *interactive knowledge testing and feedback* as effective HCI concepts for informed consent in legally compliant UIs affects efficiency and satisfaction compared to non-compliant UIs lacking active involvement of end-users. Efficiency is decreased because affirmative actions demand more activities from users. Reasons for decreased satisfaction require more investigation. Moving forward towards improving ex-ante transparency and informed consent, future research will continue to compare different modes of affirmative actions to select the data in the consent forms for their effectiveness in obtaining informed consent, robustness against habituation, appropriateness for inclusive design, and their usability in terms of efficiency and users’ satisfaction. Particularly, it is planned to include data processing purposes in the *drag and drop* method. Users will be provided with the consent forms which give them the ability to control and opt in purposes. It will be examined if these consent forms are more effective in helping users to make informed consent and are less vulnerable to habituation compared to the consent forms utilising, for example, checkboxes.
References


Towards Improving Transparency, Intervenability, and Consent in HCI

The new General Data Protection Regulation (GDPR) strengthens people’s rights for transparency, intervenability, and consent. The legal privacy principles have Human-Computer Interaction (HCI) implications. Besides aiming for legal compliance, it is of paramount importance to investigate how to provide individuals with usable and user-centric transparency, intervenability, and consent.

The objective of this thesis is to propose usable tools and solutions, to enhance people’s control and enforce legal privacy principles, especially transparency, intervenability, and informed consent. To achieve the goal of the thesis, different ways to improve ex-ante transparency and informed consent are investigated by designing and testing new solutions to make effective consent forms. Moreover, ex-post transparency and intervenability are improved by designing a transparency enhancing tool and investigating users’ perceptions of data portability and transparency in the tool. The results of this thesis contribute to the body of knowledge by mapping legal privacy principles to HCI solutions, unveiling HCI problems and answers when aiming for legal compliance, and proposing effective designs to obtain informed consent.