
D:C-7.1 General HCI principles and guidelines

Deliverable Number: D37.1

Work Package: WP 37

Version: Final

Deliverable Lead Organisation: KAU

Dissemination Level: PU

Contractual Date of Delivery (release): 30th September, 2013

Date of Delivery: 30th September, 2013

Editors

Julio Angulo (KAU), Simone Fischer-Hübner (KAU), John Sören Pettersson (KAU)

Contributors

Julio Angulo (KAU), Simone Fischer-Hübner (KAU), John Sören Pettersson (KAU), Erik Wästlund (KAU), Leonardo Martucci (KAU) Eleni Kosta (TiU) Maartje Niezen (TiU)

Table of Contents

List of Figures	5
List of Tables	6
Abbreviations.....	7
Executive Summary.....	8
1. Introduction	9
1.1 Project Scope	9
1.1 Aims and Scope of this Deliverable	9
1.2 Relationship to other A4Cloud Work Packages	10
1.3 Deliverable Outline	10
2 Related Work	11
3 HCI Challenges and Related Research Questions	14
4 Research Methods.....	16
4.1 Human Centred Design	16
4.1.1 Stakeholder workshops	16
4.1.2 Focus groups	17
4.1.3 Semi-structured Interviews	18
4.1.4 Controlled experiments.....	18
4.1.5 Usability evaluations.....	19
4.1.6 Eliciting and mapping legal requirements.....	19
4.1.7 Eliciting requirements from trust issues mentioned in studies and surveys on cloud and Internet use	19
4.2 Ethical consideration	20
5 Eliciting HCI Requirements and Principles.....	21
5.1 Workshops, focus groups and interviews	21

5.1.1	Eliciting requirements from the initial stakeholders workshop (B-2)	21
5.1.2	Eliciting requirements from HCI stakeholders' workshop	22
5.1.3	Focus groups: advanced vs. lay users' mental models and attitudes of cloud services.....	28
5.2	Usability tests and controlled experiments	31
5.2.1	Background: Mental models of privacy and control of personal information	31
5.2.2	Exploring users' behaviours, needs and understandings through controlled experiments	32
5.2.3	Experiment 1: Understanding willingness to distribute personal data to cloud services.....	35
5.2.4	Experiment 2: Framing and terminology	37
5.2.5	Experiment 3: Desired features on cloud services	39
5.3	Evaluating visualizations of data disclosures and data traces	43
5.3.1	Background	43
5.3.2	Evaluation	46
5.3.3	Results.....	46
5.3.4	Summary of results.....	49
5.3.5	Limitations and next steps	51
5.4	Usability and Security for Access Control Rule Sets	52
5.4.1	Background	52
5.4.2	Experiment 1: Semi-structured interviews with system administrators for eliciting security and usability requirements.	52
5.4.3	Experiment 2: Between subject design to collect data regarding the use of our support tools for producing access control rule sets	53
5.4.4	Experiment 3: Expert opinion to rank the collected data according to their knowledge.....	53
5.4.5	Summary of results.....	53
5.5	Mapping legal principles	54
5.5.1	Legal Principles	54

5.5.2	HCI Requirements, Principles and Design Proposals	56
5.5.3	Summary of results.....	58
5.6	Mapping social trust factors	60
5.6.1	Literature review	60
5.6.2	Summary of trust factors:	63
5.7	Concluding words	66
6	Preliminary HCI Principles and Guidelines.....	67
6.1	Mapping HCI requirements to functional categories of A4Cloud tools	67
6.2	Towards HCI Guidelines for A4Cloud	70
6.2.1	Motivate users to make informed decisions	70
6.2.2	Help users comprehend policies and manage their preferences	71
6.2.3	Provide options for action	73
6.2.4	Frame in terms of consequences rather than technicalities	74
6.2.5	Consider differences in users (cultures, expertise, legal regimes, etc.)..	75
6.2.6	Make trustworthiness transparent	75
6.2.7	Provide privacy-friendly and useful defaults	76
6.2.8	Illustrate who is in control of the data	77
6.2.9	Plurality of input and output	78
7	Concluding Remarks	80
	References	81
	Appendices.....	87
	Appendix A.1: Experiment with fake cloud service	87
	Appendix A.2: Data Track usability tests	88
	Appendix B.1: Matching General HCI Requirements and Principles to the High-level Functional Analysis of the A4Cloud Scenarios	91

List of Figures

Figure 1. An illustration from a group of expert participants showing the entities involved in a transaction using the Skype service.....	29
Figure 2. SheepCloud registration page. Users were made believe they were registering and releasing personal data to a new storage cloud service.	33
Figure 3. Example of an offer to get double the cloud storage space if the user hands out control of his personal information to the cloud service provider.....	35
Figure 4. Screen shot from the baseline condition. In the two experimental conditions the subheading was changed to frame one of the two choices in a positive way.....	38
Figure 5. The functions for controlling data that SheepCloud offered at the time of registration.....	40
Figure 6. DataTrack user interface developed under the PrimeLife project.....	44
Figure 7. The trace view user interface of Data Track	45
Figure 8. Information about a user that a service provider has stored on their servers (service's side)	45
Figure 9. Post-questionnaire scale on the understanding of the Data Track trace view.....	47
Figure 10. Example of a service provider in the bottom panel of the Data Track's trace view, including storage icon to be clicked for getting online access to one's data stored at the service provider.	47
Figure 11. Data Track's timeline view of data disclosures.	49
Figure 12. Data Track mock-up for illustrating chains of information flows.....	52
Figure 13. Example of well understood PrimeLife policy icons	57
Figure 14. Icon proposals (alpha version) by Aza Raskin informing about how disclosure requests by law enforcement are handled	57
Figure 15. Example from Kelley et al. at making users decide on apps to install based on privacy facts.	71
Figure 16. Example of a multi-layered privacy policy complemented with icons by iubenda.	73
Figure 17. Example from the ghostery browser plugin.....	74
Figure 18. Example of the WOT plugin to indicate trustworthiness	76
Figure 19. Example of providing icons representing data in the cloud	78
Figure 20. Example of providing multiple ways for inputing data	79

List of Tables

Table 1. Summary of focus group sessions	17
Table 2. Summary of controlled experiments	18
Table 3. HCI requirements obtained from first stakeholder workshop done in WP B.2	21
Table 4. Participants of the HCI stakeholder workshop	23
Table 5. HCI requirements and design ideas obtained from HCI stakeholder workshop	24
Table 6. HCI requirements and design ideas obtained from focus groups	29
Table 7. Crosstabulation of the willingness to control data depending on the sensitivity of the data and the amount of storage offered.	36
Table 8. HCI requirements and design ideas obtained from Experiment 1	37
Table 9. Descriptive statistics showing the number of participants assigned to each conditions of Experiment 3.	38
Table 10. HCI requirements and design ideas obtained from Experiment 2	39
Table 11. The possible features for control of personal data and the participants' preferred features. 40	
Table 12. HCI requirements and design ideas obtained from Experiment 3	42
Table 13. HCI requirements and design solutions obtained from evaluating the transparency tool Data Track.....	49
Table 14. HCI principles and design solutions obtained from evaluating novel techniques for access control rules	53
Table 15. Mapping Legal Privacy Principles to HCI requirements and proposed solutions	58
Table 16. HCI requirements and design ideas obtained from literature review on trustworthy factors 63	
Table 17. Mapping HCI requirements and principles to functional categories of A4Cloud tools	67
Table 18. Functional categories for HCI requirements and principles mapped to the A4Cloud scenario functionalities for individual end users (cloud users).....	91
Table 19. Functional categories for HCI requirements and principles mapped to the A4Cloud scenario functionalities for business end users (cloud users)	94
Table 20. Functional categories for HCI requirements and principles mapped to the A4Cloud scenario functionalities for cloud auditors	97

Abbreviations

A4Cloud	Accountability For Cloud and Other Future Internet Services
CSP	Cloud Service Provider
DoW	Description of Work
EEA	European Economic Area
EU	European Union
GDPR	General Data Protection Regulation
HCI	Human-Computer Interaction
IaaS	Infrastructure as a Service
ISV	Independent Software Vendors
PaaS	Platform as a Service
PETs	Privacy Enhancing Technologies
SaaS	Software as a Service
SLA	Service level Agreement
TETs	Transparency Enhancing Technologies
UI	User Interface
WP	Work Package

Executive Summary

This deliverable elaborates HCI (Human Computer Interaction) concepts for making A4Cloud tools to be developed for different stakeholder groups comprehensible and trustworthy. A human-centred design approach is followed to elicit HCI requirements and to derive general HCI principles, guidelines, and proposals for user interface solutions. For deriving HCI requirements and principles, we conducted research and review work for addressing particularly the following HCI challenges:

- How can the users be guided to better comprehend the flow and traces of data on the Internet and in the cloud?
- How can individual end users be supported to do better informed decisions on how their data can be used by cloud providers or others?
- How can the legal privacy principle of transparency and accountability be enforced by the user interfaces of A4Cloud tools?
- How can the user interfaces help users to reassess their trust/distrust in services?

The research methods that we used comprise stakeholder workshops, focus groups, controlled experiments, usability tests and literature and law reviews.

Derived HCI requirements and principles were first grouped into the functional categories ex ante transparency (in form of policy notices which enable the anticipation of consequences before data are actually disclosed), exercising data subject rights, obtaining consent, policy preference management, ex post transparency (which inform about consequences if data already has been revealed), audit configuration, access control management and privacy risk assessment and then mapped to the functionalities of tools for different stakeholders in the A4Cloud use case descriptions.

Finally, some high level HCI guidelines are presented that are summarising a selection of key HCI principles with an emphasis on tools for individual end users. Even though these HCI guidelines are on such a high level also valid for many other privacy-enhancing technologies, it is nevertheless important to stress that they are especially relevant for the cloud context where developers have to apply them against the background of the complex picture of the cloud service chain. Moreover, user interfaces for transparency tools for the cloud should clearly inform users about additional aspects beyond the policy information that is legally required as a minimum, so that users can understand the implications very well. Such additional policy information may comprise information about contacts and obligations of data processors along the cloud chain, the geographic locations of data centres, applicable laws and consumer rights, how disclosure requests by law enforcement are handled.

Our high level guidelines recommend in particular that ex ante transparency tools should make the consequences of data disclosures more transparent. Privacy-friendly and useful default privacy settings should be provided, which can be adapted to the user's situation. Besides, ex post transparency tools have to make obvious who is in control or processing the data (the user, the service or cloud service provider) and what means exist for exercising data subject rights in what situations.

1. Introduction

1.1 Project Scope

The A4Cloud project deals with accountability for the cloud and other future Internet services. It conducts research with the objective of increasing trust in cloud computing by developing methods and tools for different stakeholders through which cloud providers across the entire cloud service value chains can be made accountable for the privacy and confidentiality of information held in the cloud. The A4Cloud stakeholders, for whom methods and tools will be developed, comprise so called cloud consumers in the form of individual end users or business end users (i.e., service providers outsourcing data processing to the cloud), further data subjects¹ whose data have been outsourced to the cloud, as well as regulators, such as data protection commissioners, and cloud auditors. The methods and tools that are developed are combining risk analysis, policy enforcement, monitoring and compliance auditing with tailored IT mechanisms for security, assurance and redress. In particular, the A4Cloud project is creating solutions to support cloud users in *deciding* and *tracking* how their data are used by cloud service providers (Pearson et al. 2012).

A4Cloud solutions will thus also include tools for enhancing transparency of data processing for the different stakeholders (so-called transparency-enhancing tools -- or in short: TETs). The concept of transparency, as it is considered by us in A4Cloud, comprises both 'ex ante transparency', which enables the anticipation of consequences before data are actually disclosed (e.g., with the help of privacy policy statements), as well as 'ex post transparency', which informs about consequences if data already has been revealed (what data are processed by whom and whether the data processing is in conformance with negotiated or stated policies) (Hildebrandt 2009).

1.1 Aims and Scope of this Deliverable

Task T:C-7.2 of A4Cloud work package C-7 on "HCI concepts for usable transparency and accountability" has the objective to elaborate general HCI (Human Computer Interaction) concepts for making A4Cloud tools comprehensible and trustworthy – which will be key factors for their successful deployment –, and to draw up user-interface design principles.

This deliverable aims at providing a first set of such general HCI principles and guidelines, which have a basis in human-centred design, and should be considered for User Interface (UI) design for the A4Cloud functions that gradually will be developed in the course of the project. The design principles have first been iteratively developed for generic interfaces and have then been extended and applied for the interfaces addressing the use cases published by WP:B-3 (Bernsmed et al. 2013).

For deriving such HCI principles and guidelines, Task T:C-7.2 conducted research and review work for addressing particularly the following HCI challenges that are of relevance for the tools to be developed for different A4Cloud stakeholders:

- How can the users be guided to better comprehend the flow and traces of data on the Internet and in the cloud?
- How can individual end users (i.e. data subjects) be supported to do better informed decisions on how their data can be used by cloud providers or others?
- How can the legal privacy principle of transparency and accountability be enforced by the user interfaces of A4Cloud tools?
- How can the user interfaces help users (in particular individual end users) to reassess their trust/distrust in services?

¹ A data subject is a natural person about whom personal data are processed.

For addressing these challenges, a human-centred design approach is taken in WP:C-7 (see Chapter 2). This deliverable documents the work conducted for addressing these HCI challenges and the results that we achieved in the form of derived HCI principles and guidelines.

This deliverable is however only the first deliverable of task T:C-7.1 and is focusing especially on general and generic HCI concepts for transparency and accountability, rather than on the concrete design proposal for A4Cloud tool user interfaces, as the functionalities of A4Cloud tools were not yet elaborated in detail during the first months of the project when the main work for this deliverable was conducted. At the end of the second project year, an HCI report on the perception of more concrete user interfaces to be developed for A4Cloud tools in WP:D-5 will be delivered.

1.2 Relationship to other A4Cloud Work Packages

This deliverable D:C-7.1, “General HCI principles and guidelines” has the objective to provide general HCI principles to populate the reference architecture developed by WP:D-2 and to provide guidance for the design of usable and trustworthy user interfaces for accountability and transparency tools in WP:D-5. Whereas the HCI work in task WP C-7 focuses on general HCI concepts, WP:D-5 will in its HCI-related task T:D-5.1 on “User interfaces for toolsets for different stakeholder groups” iteratively develop and test concrete user interface designs for the A4Cloud toolset.

This deliverable partly relies on work led by WP:B-3 and presented in deliverable D:B-3.1, “Use Case Descriptions”. In D:B-3.1, three uses cases were developed and analysed for the definition of the functionality that various kinds of user will interact with in a future cloud ecosystem where a satisfying level of accountability exists. The functionality compiled in D:B-3.1 have been analysed as to what design principles and guidelines are required to meet various known issues and problems for users, while the exact detailed designs will have to wait until the more definitive descriptions will be available about the tool functionalities.

1.3 Deliverable Outline

The remainder of this deliverable is structured as follows:

Chapter 2 on “Related Work” will present related previous work on HCI principles and guidelines for Privacy-Enhancing Technologies (PETs) and privacy-enhancing identity management including transparency-enhancing tools and functions. It is discussed how far these guidelines can also be applied to A4Cloud, and what the limitations of these guidelines are.

Chapter 3 on “HCI Challenges” motivates the choice of HCI challenges addressed in this deliverable mostly as an answer to these limitations. It also discusses the research questions that those challenges imply in more detail.

Chapter 4 on “Methodology” then discusses and motivates the different research methods that we have applied when addressing these HCI challenges and deriving HCI principles while following a human-centred design approach.

Chapter 5 on “Eliciting HCI requirements and principles” reports on the actual research work done for exploring the identified HCI challenges, for eliciting HCI requirements and discussing HCI solutions and principles.

Chapter 6 on “General HCI Guidelines for A4Cloud” is then deriving some overall HCI guidelines for A4Cloud from the HCI principles and proposed HCI solutions that we discussed in Chapter 5.

Finally, Chapter 7 “Concluding Remarks” will provide conclusions of this deliverable and provide an outlook into the future HCI work of work package C-7.

2 Related Work

This chapter presents an overview of related HCI principles, recommendations and guidelines for usable privacy and security, which are based on earlier research and that can be of relevance for A4Cloud technologies. The related work discussed in this chapter provides basic HCI rules that can also be applied or adapted to future A4Cloud technologies. We point out how far existing guidelines need further enhancements for the context of accountability and transparency in the cloud.

HCI guidelines for both security and privacy technologies have to address specific HCI challenges, as noted first by Whitten and Tygar (1999) for security, and later by many others for privacy:

- Security and privacy protection are typically secondary goals for ordinary users;
- They contain difficult concepts that may be unintuitive to lay users
- True reversal of actions is not possible.

Jakob Nielsen published one of the most referred to collection of general HCI principles, his so-called 10 Usability Heuristics for User Interface Design (Nielsen 1995), which are called "heuristics" because they are rather rules of thumb than specific usability guidelines. These HCI heuristics, which were originally derived from an analysis of 249 usability problems (Nielsen 1995), comprise: "Visibility of system status", "Match between system and the real world", "User control and freedom", "Consistency and standards", "Error prevention", "Recognition rather than recall", "Flexibility and efficiency of use", "Aesthetic and minimalist design", "Help users recognize, diagnose, and recover from errors", "help and documentation." Johnston et al. expanded and modified the Nielsen's list of principles to derive criteria for a successful HCI applied in the area of IT security ("HCI-S") (Johnston et al. 2003).

Further relevant HCI guidelines for aligning security and usability for secure applications were for instance proposed by Yee (Yee 2004) and by Garfinkel (Garfinkel 2005). Even though these guidelines are related to secure applications, some of them can be interpreted and adapted to privacy-enhancing transparency and accountability. For instance, Yee's guideline of "Explicit authorization" stating that "a user's authority should only be granted to another actor through an explicit user action understood to imply granting" can be translated to the guideline that informed consent to personal data disclosure should require an explicit user action understood to imply disclosure. Similarly, also his principles of "Visibility" and "Revocability" of authority could be applied to personal data disclosures. Dhamija and Dussault discussed flaws of identity management posing HCI and security challenges, and provide some HCI-related recommendations how to address them, which are partly based on Yee's guidelines (Dhamija & Dussault 2008).

Important domain-specific HCI requirements can be derived from privacy legislation. In the EU FP5 project PISA (Privacy Incorporated Software Agents), Patrick et al. have studied in detail how legal privacy principles derived from the EU Data Protection Directive 95/46/EC (European Commission 1995) can be translated into HCI requirements and what are possible design solutions to meet those requirements (Patrick & Kenny 2003; Patrick et al. 2003). Their research focussed on legal privacy principles of (a) transparency, (b) purpose specification and limitation and (c) data subject rights, as well as (d) informed consent as a basis for legitimate data processing. As concluded by the project, these legal principles "have HCI implications because they describe mental processes and behaviours that the data subject must experience in order for a service to adhere to the principles. For example, the principles require that users understand the transparency options, are aware of when they can be used, and are able to control how their personal data are handled. These legal requirements are related to mental processes and human behaviour, and HCI techniques are available to satisfy these requirements" (Patrick et al. 2003). Therefore, the HCI requirements that were derived comprised requirements on comprehension (to understand, or to know), consciousness (to be aware of or to be informed), control (to manipulate, or be empowered) and consent (to agree) in relation to the selected legal principles.

As a possible HCI solution for achieving informed consent and (ex ante) transparency, the PISA project proposed the concept of 'Just-In-Time-Click-Through Agreements' (JITCTAs), which instead of providing complex and lengthy service terms, should confirm the users' understanding or consent on an as-needed basis. JITCTAS therefore provide small agreements that are easier for the user to read and process, and that facilitate a better understanding of the decision being made in context.

The Art. 29 Data protection Working Party² has in its opinion on "More Harmonised Information Provisions" given the recommendation of providing information in a "multi-layered format under which each layer should offer individuals the information needed to understand their position and make decisions" (Art. 29 Data Protection Working Party 2004). They suggest three layers of information provided to individuals, which include the short privacy notice (basically corresponding to JITCTAs), the condensed notice and the full privacy notice. The short notice (layer 1) must offer individuals the core information required under Article 10 of the EU Data Protection Directive 95/46/EC, which includes at least the identity of the controller and the purpose of processing. In addition, a clear indication must be given as to how the individual can access additional information. "The condensed notice (layer 2) includes in addition all other relevant information required under Art. 10, such as the recipients or categories of recipients, whether replies to questions are obligatory or voluntary and information about the data subject's rights. The full notice (layer 3) includes in addition to layers 1 and 2 also "national legal requirements and specificities."

In the EU FP6 PRIME project on "Privacy and Identity Management for Europe", one built upon the legal privacy principles and HCI requirements from the PISA project along with HCI requirements for socio-cultural privacy principles to derive proposed UI design solutions for privacy-enhancing Identity Management systems (Pettersson 2008).

The PRIME project has also followed the Working Party's recommendations to use multi-layered privacy notices and the concept of a JITCTA in its design proposals for "Send Data?" dialogue boxes for obtaining the user's informed consent. However, a problem with click-through agreements including JITCTAs is that users have the tendency to automate behaviours so that the individual parts of an action are executed without conscious reflection (International Standard Organization (ISO) 1998). The PRIME HCI work package therefore also developed the alternative concept of Drag-And-Drop-Agreements (DADAs), by which users have to express consent by moving graphical representations of their data to a graphical representation of the receiver, and thus forces users to make better informed decisions while also allowing the system to detect erroneous conceptions of the user if data are dropped on the wrong recipient (e.g. credit card symbol is dropped on web shop symbol instead of on pay service symbol) (Pettersson et al. 2005).

Based on experiences gained from developing UIs for privacy-enhancing identity management systems over several years, the EU FP7 project PrimeLife provided an experience report "*Towards Usable Privacy Enhancing Technologies: Lessons Learned from the PrimeLife Project*" (Graf et al. 2011) which discusses HCI fallacies and provides HCI heuristics, best practice solutions and guidance for the development of usable PETs, which will be of relevance for A4Cloud. This report started with identifying major HCI fallacies that were experienced, which included the problem of many users to differentiate whether data are stored on the user side (under the user's control) and to

² Under Article 29 of the Data Protection Directive, a Working Party on the Protection of Individuals with regard to the Processing of Personal Data is established, made up of the Data Protection Commissioners from the Member States together with a representative of the European Commission. The Working Party is independent and acts in an advisory capacity. The Working Party seeks to harmonize the application of data protection rules throughout the EU, and publishes opinions and recommendations on various data protection topics.

comprehend to which network entities personal data flows during online transactions. Furthermore, the mediation of trustworthiness, intercultural differences and a well comprehensible terminology to be used in UIs are challenges to be taken into consideration. Many of the HCI issues that were experienced are mental model issues which are difficult to solve for novel PET concept, which are unfamiliar for the users. This is especially true for those PETs, for which no obvious real world analogies exist. Based on those experiences and lessons learned, the report provides HCI heuristics for PETs, which adapt, extend and exemplify the classical list of Nielsen's Usability Heuristics for the PET domain. Finally, the report also provides some evaluation guidelines for PET user interfaces, and what needs to be considered for the preparation and performance of usability tests.

In particular, *PET-USES* (Privacy-Enhancing Technology Users' Self-Estimation Scale) is introduced, which was developed in PrimeLife as a post-test questionnaire that enables users to evaluate PET-User Interfaces both in terms of the primary task and specific PET related secondary tasks (Wästlund et al. 2010).

In complementation to the HCI heuristics, the PrimeLife project also developed *HCI Patterns for PETs* which provide best practice solutions ("design patterns", after Alexander (1977)) for the PET user interface design (PrimeLife WP4.1 2010). Relevant also is the on-going Privacy Design Pattern project described by Doty & Gupta³.

While the existing HCI principles and guidelines presented in this chapter are still valid and applicable to the A4Cloud tools to be developed within the A4Cloud project, still some work is needed to elaborate and derive further HCI principles and guidelines addressing specifically HCI challenges for transparency and accountability technologies in the cloud context. Most HCI fallacies identified by the PrimeLife project in regard to the users' comprehension of his personal data flows and traces, trust in PETs and comprehension of novel PET concepts will also be important to address in the A4Cloud project when designing user interfaces for privacy-enhancing transparency and accountability tools for the cloud. Besides, legal privacy principles to be mapped into HCI principles and design solutions may be interpreted differently for the cloud and are currently re-discussed under the proposed reform of data protection legislation in Europe. Therefore, we have specifically researched related HCI challenges on comprehension of personal data flows, PET concepts such as policy notices, trust and the interpretation of legal privacy principles in the cloud context to derive further specific HCI principles and guidelines for A4Cloud.

³ <http://privacypatterns.org/>

3 HCI Challenges and Related Research Questions

This chapter briefly motivates and lists the HCI challenges and related research questions that we have addressed to derive specific HCI principles and guidelines for A4Cloud.

The A4Cloud project is creating solutions to support cloud users in *deciding* and *tracking* how their data are used by cloud service providers (Pearson et al. 2012). As discussed in Chapter 2, previous HCI research in the EU project PrimeLife had however revealed that many users have problems to differentiate whether data are stored on the user side (under the user's control) or on a remote services side and the problem to comprehend to which network entities personal data flows during online transactions (PrimeLife WP4.1 2010). Evoking the correct mental model in regard to where data are transferred to and where they are processed will especially be a challenge for the cloud with chains of cloud service providers that may be involved.

Hence, one major challenge for the HCI design of usable privacy-enhancing transparency tools in A4Cloud and related research questions that we addressed are:

1. How can the users be guided to better comprehend the flow and traces of data on the Internet and in the cloud?

- What are the mental models of different stakeholders and types of users in regard to the distribution of personal data in a complex cloud ecosystem?
- What HCI concepts are suitable for evoking the correct mental models of data flows and traces?

These questions will be significant for both ex ante TETs, e.g. in the form of privacy policy tools, as well as for ex-post TETs, which will allow users to track their data in the cloud.

However, for supporting individual users in making decisions on how their data are used by cloud providers, it has to be taken into consideration that previous research has shown that lay users often do not behave rationally with regard to decisions on personal data disclosure (Spiekermann et al. 2001; Gross & Acquisti 2005) meaning that we cannot assume either that they will do so when deciding on the disclose or outsourcing their data to the cloud. In order to design usable tools that offer transparency and accountability of the users' data in the cloud, we have to understand their attitudes, behaviours and mental models in relation to cloud services. Having these understandings can help to reveal what these users value, what they think is important, and what useful features that can be included in the user-friendly tools for transparency and accountability and how these features can be designed to be valued and well understood by individual users.

When it comes to the business end users, their security officers face the challenge generating and managing access control rule sets for controlling the use of data in the cloud.

These aspects have motivated us to research also the following:

2. How can individual end users be supported to make more informed decisions on how their data can be used by cloud providers or others?

- How much cognitive effort or time are people willing to spend in order to understand what happens to different types of personal information in the cloud?
- How can the user interfaces of ex ante TETs be designed to support and motivate users to take more rational and informed decisions?
- How can service providers obtain usable access control rule sets for data outsourced to the cloud that are reflecting the organisation's access control policy and are easy to understand and manage?

The EU Legal Data Protection Directive has defined legal principles for providing transparency and control to users. In the context of cloud computing, the existing legal requirements may partly need some re-interpretation. Currently, also new legal principles for providing better transparency and control for individual cloud users and increasing accountability for cloud providers have been discussed as part of the proposed EU data protection regulation (European Commission 2012). Therefore, a third HCI challenge that we addressed, which is also related to the other two HCI challenges mentioned above, is:

3. How can the legal privacy principles of transparency and accountability be enforced by the user interfaces of A4Cloud tools?

- What legal privacy principles for transparency and accountability for the cloud need to be taken into consideration by the HCI design of A4Cloud tools?
- How can legal privacy principles for transparency and accountability for the cloud be mapped to HCI principles and solutions?

Finally, as concluded by the PrimeLife project in its Lessons Learned report (Graf et al. 2011), trust plays a key role in the acceptance and uptake of PET solutions. Users may lack trust in novel PETs (such as the A4Cloud tools to be developed) with functionalities which may not fit their mental models of how the technology should work. For this reason, one more challenge to be tackled is:

4. How can the user interfaces help users (in particular individual end users) to reassess their trust/distrust in services?

- What are suitable HCI means for mediating trust in trustworthy services (as evaluated by A4Cloud tools)?
- How can user interfaces connect to known reliable sources for trust?

In the next chapter, we will discuss the research methodology that we have used for addressing these challenges following a human-centred design approach. Chapters 5 and 6 will then report on the actual research work done for exploring the identified HCI challenges and the results that we achieved in terms of elicited HCI principles and guidelines.

4 Research Methods

4.1 Human Centred Design

In A4Cloud's Work Package C7, we follow a human centred design approach for eliciting and testing HCI requirements and guiding the development of user interface design principles. Human-centred design is defined by ISO 9241-210, 2010 as “an approach to interactive systems development that aims to make systems usable and useful by focusing on the users, their needs and requirements, and by applying human factors/ergonomics, and usability knowledge and techniques” (International Standard Organization (ISO) 2010). User requirements are considered right from the start and included into the whole design and development cycle. In A4Cloud, we have elicited and refined such user requirements and related HCI principles through methods including stakeholder requirements, focus groups, controlled usability testing and other methods described in the subsections below.

For the choice of methods, we have taken into consideration that general concepts that are of importance for the comprehension of transparency and related risks, such as what information is stored and where it is processed, are usually difficult to understand for the lay users, while other end user groups such as regulators or security administrators usually have a clearer understanding. Therefore, different user-groups require different interfaces and interaction paradigms. This also means that the different user groups have to be involved using different approaches to human-centred design. For this reason, we have used controlled experiments and mock-up-based evaluations in addition to focus groups in order to explore the needs of lay users, while the needs of professional stakeholder groups were mainly investigated by means of stakeholder workshops and focus groups. The controlled experiments and mock-up-based evaluations had as an objective to analyse the user's mental models of A4cloud related technical concepts, since our earlier work has shown that many HCI issues are mental model issues which are difficult to solve for novel PET concept (Graf et al. 2011).

The following subsections briefly describe the methodologies applied and the reason they were chosen as suitable approaches for eliciting HCI requirements within the A4Cloud project.

4.1.1 Stakeholder workshops

Stakeholder workshops provide the opportunity for active face-to-face interactions between different influential actors who can express their opinions and needs for a system being developed. This method is strongly encouraged during the initial design processes, as a way of ensuring that the needs of those who might be impacted by the system are taken into account, as well as trying to achieve a common vision of the system (Maguire & Bevan 2002). An important step of this method is identifying those stakeholders that can have a say on the development of the system. Typically one stakeholder representative is selected from a user group and invited to participate in a workshop.

Once the stakeholders have been identified different approaches can be followed during the meeting in order to incite discussions, to promote the exchange of ideas and to identify the needs of the different user groups being represented by invited stakeholders. Such approaches can include general discussions, moderated interviews, focus groups, as well as Open Space (Owen 2008) and World Cafés (Brown & Isaacs 2005) methodologies, and others. Depending on the approach taken and the number of participants, the discussions might derive from one main question (as is often the case of Open Space), or from a series of questions. Also, participants might be divided into groups trying to identify challenges related to different themes, or they can be all exchanging ideas while a moderator leads the discussions. The results from the discussions can then be compiled, interpreted and expressed as a set of system requirements. Follow-up interviews or feedback from participants can also be setup in case the researchers need to complement or correct the information acquired during the workshop session.

In the A4Cloud project, Work Package B-2 has the task of planning and carrying out a series of stakeholder workshops focusing on different themes related to accountability, transparency and risk on cloud services. As a complement to the work done by WP:B-2 (Brede Moe et al. 2013), we have

carried out an additional stakeholder workshop concentrating on the HCI aspects of cloud services. The purpose of running such a workshop was to discover initial cloud related HCI requirements. These initial requirements would also serve as the bases and motivations for our subsequent experiments and tests that we conducted.

More information about the participants and the requirements gathered from that workshop can be seen in Section 5.1.

4.1.2 Focus groups

Focus groups are appropriate for bringing together a cross-section of users so that they can collaboratively share and unveil their opinions and needs regarding particular challenges foreseen in the design of a system. Moderators of a focus group can stimulate participants to discuss these opinions with the other group members by using different approaches, such as asking direct questions to participants, encouraging brainstorming, instructing them to work with various probes, etc.

To understand the different ways in which individuals with different levels of familiarity with technology perceive cloud services and comprehend the flow of their personal data on the Internet and in the cloud, we conducted three focus groups session (including a pilot session) with participants that were considered expert and non-expert users.

The group of expert users was formed of 16 Ph.D. students in computer science coming from different Swedish Universities (but with different nationalities) who were taking a graduate course on the topic of Privacy Enhancing Technologies. The non-expert users consisted of a group of 15 individuals from different age ranges, cultural and educational backgrounds, who were participants of project for personal development towards employment opportunities⁴. The following table summarizes the characteristics of the focus group sessions. More detailed descriptions of these focus groups and the requirements obtained from them can be found in Section 5.1.3.

The table below summarizes the structure and purpose of each of these focus groups:

Table 1. Summary of focus group sessions

Focus group	Participants	Purpose
Mental models of data sharing by Internet service providers	Approximately 15 students taking a course on Internet businesses at Karlstad University.	Pilot focus group session that served as planning for the latter focus groups.
Mental models of data usage, data flow and vulnerabilities in Internet services	16 participants considered expert users recruited at a PhD course on Privacy Enhancing Technologies.	To understand the needs and mental models of users with high knowledge of computers and experience with cloud services.

⁴ The project is called UMA (Utveckling Mot Arbete) taking place in the city of Kristinehamn, Sweden.

Mental models of data usage, flow and vulnerabilities in Internet services	15 participants considered non-expert users recruited through a program of personal development towards employment opportunities.	To understand the needs and mental models of users who have relatively little or no knowledge interacting with computers or who had little or no experience using cloud services.
---	---	---

4.1.3 Semi-structured Interviews

Semi-structured interviews are interviews where not all questions are designed or planned before the interview, allowing the interview to follow and explore new directions as they come up in the interview process (Bernard 1988).

Semi-structured interviews were considered a good method for capturing the challenges regarding the management of access control lists by system administrators, and how those challenges are commonly handled in their field of work. The application and results of using this method are reported in Section 5.4.

4.1.4 Controlled experiments

In experimental studies so called dependent variables of interest are identified. Then the factors in the study, or independent variables, can be controlled for checking the level of influence of these factors on the variables of interest. By performing experiments using control groups, different hypotheses about people's behaviours, actions, attitudes, opinions and performance can be tested. The ecological validity in an experiment measures the extent to which the setup of the experiment matches real world situations.

As part of WP:C-7 of A4Cloud, we have designed and carried out four controlled experiments in order to study the mental models, motivations and needs of lay users when subscribing to cloud storage services. In order to improve the ecological validity of the experiments, participants were deceived into believing that the cloud service was a real service. These are summarized in the following table:

Table 2. Summary of controlled experiments

Experiment	Participants	Hypotheses
Understanding willingness to distribute personal data to cloud services.	120	End users are more willing to release personal data to a cloud service in exchange for observable valuables (such as free cloud storage).
Framing and terminology	190	End users willingness to release personal data depends on how the cloud service expresses benefits at the moment of releasing data.
Desired cloud services' features	179	End users would have preferences over certain features for managing their data released to a cloud service.

Moreover, a between-subjects experiment design was deployed to gather evidence for the accuracy of the metrics proposed in Section 5.4.3 for creating usable access control rule sets, also explained in (Beckerle & Martucci 2013). This type of experiment was chosen because a control group was needed for comparing the results of the participants that were assisted by a tool that provided them with

measurements regarding the security and usability of their access control rule sets with the results of the participants that didn't have such a support.

4.1.5 Usability evaluations

Usability testing is a technique that can measure the actual performance of users when trying to achieve a tasks with a given user interface.

Usability testing of low-fidelity prototypes was considered a suitable method for our purposes since it has the advantage of letting lay users communicate their needs, opinions and expectations about new technologies. This is because lay users might not be very familiar with the terminologies and technologies related to cloud computing, and might not have a clear understanding of how Internet technologies and data handling works either.

During a usability test session test participants are typically presented with a graphical user interface and are given a set of instructions or tasks that they are asked to complete. A test moderator usually guides the participant through the tasks, while at the same time observing and annotating the interactions of the participants with the interface. The moderator also encourages participants to express aloud their opinions, actions and reactions to the prototype, in an approach commonly referred to as the "think aloud" protocol (Jaspers et al. 2004).

Earlier studies of a transparency enhancing tool called "Data Track" carried out during the PrimeLife project (Wästlund & Fischer-Hübner 2010) confirmed the difficulty for lay users to comprehend the flow and traces of their data on the Internet and in the cloud, the objective of the usability tests described in Sections 5.3 was to test whether graphical illustrations of data flows can improve the lay users' understanding of their personal data traces.

Besides usability testing done with lay users, expert evaluations are also considered valid usability studies which rely on the experience and knowledge of subjects that specialize on their field of expertise. Their opinion and suggestions based on their experience can be a valuable input on the design and evaluation of technology. As a way to evaluate the user control access mechanisms proposed in Section 5.4, expert opinions were obtained, whereby system administrators ranked a series of access control rules sets according to their security and usability properties.

4.1.6 Eliciting and mapping legal requirements

Legal principles that will have to be enforced by the user interfaces of A4Cloud tools were elicited from the stakeholder group workshops, by a review of relevant legal documents (including the EU Data Protection Directive 95/46/EC (European Commission 1995), the newly proposed EU data protection regulation (European Commission 2012), and relevant opinions published by Art. 29 WP (Art. 29 Data Protection Working Party 2004; Art. 29 Data Protection Working Party 2012)), by interviews with legal experts from the A4Cloud project, as well as from input from A4Cloud advisory board. The mapping of these legal principles to HCI principles and proposed design solutions were partly based on, and extending the work of, the PISA project (Patrick & Kenny 2003), the PrimeLife HCI patterns (PrimeLife WP4.1 2010), as well as other relevant HCI guidelines and heuristics.

4.1.7 Eliciting requirements from trust issues mentioned in studies and surveys on cloud and Internet use

For eliciting HCI requirements for mediating trustworthiness of services, including cloud services when they (in the future) have been evaluated by A4Cloud tools, a literature review was conducted. Many studies on Internet services and users, in particular those involving individual end users, have focused on the degree of confidence people have in e-commerce web sites and more recently in cloud services. Our literature review, as reported in the next chapter, concentrated on a few studies from which it has been possible to crystallise HCI requirements and, to some extent, map onto tentative HCI principles or UI examples. Many of the studies refer to other works on trust but it has not been within the scope here to report on every work. Rather, only one or a few references for an interesting

trust-related phenomenon have been deemed sufficient for this report to motivate the discussion of the phenomenon in question and its possible inclusion in the table of requirements.

4.2 Ethical consideration

Before the work with external participants in tests, focus groups and workshops commenced in WP C-7, a description of the work planned and the relation to the A4Cloud project in large was sent to the local board for ethical evaluations at Karlstad University, which evaluated the plan and allowed us to go ahead. The plan described the recruitment of participants of focus groups, workshops, tests and experiments where we only involved “adult (healthy) volunteers” who provided their informed consent. Besides, the plan described routines for handling and anonymising data at the earliest possible time, providing transparency and guaranteeing data subject rights to all participants. As no sensitive data were obtained and rules of the Swedish data protection act and the EU Data Protection Directive 95/46/EC were clearly followed, no ethical or legal privacy concerns were seen.

5 Eliciting HCI Requirements and Principles

Having listed the research methodologies in Chapter 4, this chapter describes more in detail how these methodologies were applied through different research activities as well as the results obtained. The different activities, presented in the subsections of this chapter, had the goal of tackling the main research questions presented in Chapter 3.

5.1 Workshops, focus groups and interviews

5.1.1 Eliciting requirements from the initial stakeholders workshop (B-2)

Within the A4Cloud project, Work Package B.2 is in charge of organizing a series of thematic stakeholder workshops at different stages of the project. Their first workshop, held in Brussels in the middle of January 2013, followed the Open Space (Owen 2008) and World Café (Brown & Isaacs 2005) methodologies, with the primary goal of identifying “initial accountability requirements from key stakeholders” (Brede Moe et al. 2013). From this first workshop resulting in the deliverable DB-B.2 some relevant HCI requirements can be extracted and summarized in the following table:

Table 3. HCI requirements obtained from first stakeholder workshop done in WP B.2

Rel. ID	Initial Accountability Requirement	Related UI Requirements
R22	The cloud provider is responsible to the cloud consumer for the provision of evidence of data segregation.	Data segregation. UI controls for displaying evidences of data segregation.
R23	The cloud provider is responsible to the cloud auditors, Regulators and Data Protection Authorities (DPAs) for the provision of evidence of compliance of data segregation with respect to legislative regimes.	
R5	The cloud provider is responsible to the cloud consumer for the implementation of different policies tailored to the nature of data, privacy laws and needs of the cloud consumer.	Understandable policies. A UI should make cloud consumers understand the policies under which their data are being collected, and allow them to express their needs in terms of policies.
R18	The cloud provider is responsible to the cloud consumer that data are used for the intended purposes.	Informed consent and purposes for data usage. UI should make the cloud consumer aware of the data management practices of the cloud provider and to obtained informed consent in an uncomplicated manner.
R26	The cloud provider is responsible to the cloud consumer for the provision of rights management on data.	
R50	The cloud provider is responsible to the cloud consumer for asking the explicit consent for any operation on data.	
R52	The cloud provider is responsible to the cloud consumer for revoking data consent if requested.	
R51	The cloud provider is responsible to the cloud consumer for asking the explicit consent every time any operation is performed on data.	
R35	The cloud provider is responsible to the cloud consumer for the provision of data classification mechanisms supporting different data security levels (e.g.	Security. The UI should allow cloud users to specify security of the data without hindering the usability of the cloud service. In addition, the UI should

	confidential or non-confidential).	provide the highest security level as the default option when appropriate.
R36	The cloud provider is responsible to the cloud consumer for the provision of custom-made data security levels.	
R40	The cloud provider is responsible to the cloud consumer for the provision of the highest data security level as default.	
R46	The cloud provider is responsible to the cloud consumer for allowing the use of data encryption.	Transparency features. UI should provide cloud consumers with understandable visualizations for different types of transparency features, such as the data gathered, aggregated or inferred by cloud providers
R37	The cloud broker is responsible to the cloud consumer for the provision of evidence of non-data aggregation (or effective data segregation).	
R54	The cloud provider is responsible to the cloud consumer for the provision of evidence of data collection practices.	
R57	The cloud provider is responsible to the cloud consumer for the provision of evidence of data gathered, inferred or aggregated.	

5.1.2 Eliciting requirements from HCI stakeholders' workshop

As a complement for eliciting specifically further HCI requirements in regard to usable transparency and accountability from experts representing all A4Cloud stakeholder groups, a second stakeholder workshop hosted at Karlstad University was organized by Work Package C-7, which took place on 27th of February, 2013.

5.1.2.1 Inviting participants

In order to select possible participants to invite to the workshop, members of the project created a list of professionals from Sweden who are representative of the envisioned stakeholder groups, for which tools in A4Cloud are to be developed. The idea was to organize a one-day workshop that was easy for local experts to attend and which was held in Swedish, the native language of the invited participants, to avoid any language barriers. The invitees included IT experts of service providers from the private and public sectors that are adopting or are planning to adopt cloud technologies as well consumer representatives who are well aware of the problems that individuals face regarding cloud computing and are thus representing the stakeholder group of individual cloud users. Besides, a lawyer from the Swedish Data Protection Agency (Datainspektionen) was also invited to represent not only the stakeholder group of regulators, but who was through her work also familiar with privacy concerns that data subjects have in regard to the handling of their personal data in the cloud.

Targeted participants received a personalized email of invitation in which a short description of the A4Cloud project was given along with a description of the intention of the workshop and a preliminary plan. Out of the ten invited professionals, seven confirmed their participation for the workshop. The participants represented all A4Cloud stakeholder groups and provided a good mix of regulatory authorities, business professionals, IT experts, consumer representatives, and data protection officers.

The participants, their professions and the A4Cloud stakeholder group that they are representing are listed below⁵:

Table 4. Participants of the HCI stakeholder workshop

Name	Organization	Position	Representative of A4Cloud Stakeholder Group
Ingela Alverfors	Swedish Data Protection Authority	Lawyer	Regulator, Data subjects/individual end users
Erik Mattson	European Consumer Centre Network	Consumer Legal Advisor	Individual end users
Niklas Nikitin	Karlstad University	IT Service Manager	Business end user (public sector)
Niklas Larsson	Landstinget (Regional Public Health Care Provider)	IT Planner	Business end user (public sector)
Farid Sajadi	Karlstad Kommun (Municipality of Karlstad)	IT Project Leader, Information Security	Business end user (public sector)
Mats Persson	Tieto AB	Senior Delivery Manager	Business end user (private sector)
Jan Branzell	Veriscan Security AB	Vice president	Business end user (private sector)

5.1.2.2 Approach

The workshop was divided into two main sessions, a morning and an afternoon session. The purpose of the morning session was to facilitate group discussions amongst all stakeholders in a relaxed manner. The objective was to encourage all participants to share their experiences and concerns regarding cloud computing. A moderator encouraged participants, without biasing the discussions, to elaborate on common questions, concerns and decisions regarding cloud computing services, such as client opinions, the considerations that are important when acquiring cloud services, the decision process of business and individual users surrounding adopting and using cloud computing services, as well as the issues encountered during the use of these services. Observers were assigned to record notes and occasionally ask questions to clarify points or to keep discussions alive. During the afternoon session participants were divided into two parallel groups, where the discussions in one group concentrated on business end users and on the other group focused on individual end users. Participants were free to choose which group they wanted to attend depending on their interests. A moderator was present in each group as well as an observer. In each of the parallel sessions, participants were encouraged to reflect over specific issues, concerns or benefits of cloud technology. In particular, the following participants were encouraged to discuss answers to the following questions:

- What problems do you observe?
- In which situation/environment/context do you observe such problem?
- Whom does this problem or issue affect?
- How can a computer tool help address this problem?

⁵ The informed consent of participants was obtained to publish their information

- What are legal and trust factors that should be considered?

Participants were given whiteboard markers and post-it notes to write down the ideas or important points that emerged while having these questions in mind.

After about one and a half hours of group discussions, all participants were brought together again to share their findings with the intention of complementing each other's discussions. The group discussions were collaboratively written on a blackboard and the notes from observers were compiled and analyzed after the workshop. The results obtained from this stakeholder workshop are summarized in the following section.

5.1.2.3 Results

Table 5 below summarises the problem in regard to usable transparency and accountability for the different stakeholder groups that were raised during the workshop and maps these problems to HCI requirements. Besides, for some of the elicited HCI requirements HCI principles and/or examples of design solutions are provided, which were partly suggested by the stakeholder workshop attendees and partly suggested by us.

Most notably, the workshop revealed problems for individual end users with respect to:

- Unclear responsibilities regarding: *Who is the data controller⁶? What liabilities do data processors, service brokers have? How do I get redress? What (national) laws apply?*
This is especially an issue if:
 - Swedish service brokers use services that reside in other countries
 - A Service provider appears to be located in Sweden (Website in Swedish, Swedish domain/address/telephone number, etc.), but is located in another country
- Insufficient support for service cancellation or data export
- Difficulties to understand trust seals and privacy policies

Furthermore, the workshop also revealed that business end users lack means to negotiate contracts and to view (mis-) matches of SLAs (service level agreements) along the cloud chain. All stakeholder groups require usable and selective audit and tracking tools.

Table 5. HCI requirements and design ideas obtained from HCI stakeholder workshop

Req #	Observation (or Problem)	HCI Requirement	Proposed HCI principles and/or sample design solutions
-------	--------------------------	-----------------	--

⁶ According to EU Directive 95/46/EC, a data controller is defined as the entity that alone or jointly with others determines the purposes and means of personal data processing.

R.1A	In contrast to traditional outsourcing, standard contracts are usually used for cloud Computing, which are often less negotiable for business end users in terms of security and privacy.	Make it possible for users to negotiate what is negotiable, and make the negotiation process clear and simple.	Provide opt-in alternatives, e.g. in regard to the country/legal regime of the data storage location.
R.1B	Often individual end users do not make really informed choice. It is easy to deceive people because they often do neither read nor understand the agreements.	Display privacy policies in a simple and understandable manner.	<p>Privacy policy statements could be explained in short videos clips (produced by consumer organizations), at the time when the user has to make choices.</p> <p>Display a graph view of personal data flow, showing how the service provider that users are contacting is connected to other services and the possible distribution of users' data for different purposes.</p> <p>Drag-and-drop data handling agreements can also help users to consciously understand what they are agreeing to.</p>
R.1C	<p>There are no seal/labels for security and trustworthiness for cloud services. If there were, how would the users know what labels to trust?</p> <p>Individuals are often not interested in understanding all details of trust seals, but would rather like to know in general whether their data are "secure".</p>	Information about trust seals should be displayed in an understandable manner. Further information about the meaning of the seal should be easily accessible.	<p>As suggested in (European Commission 2012) information about trust-related aspects of seals can be hierarchically structured in different layers (similarly as multi-layered privacy policies).</p> <p>Standardized and broadly used seals can be more easily recognized and understood.</p> <p>In-place information about what a seal means can be provided, e.g. via tooltips or information dialogs.</p>
R.1D	It is unclear for individual users how they can get redress or compensation if something goes wrong, and whom they should contact in this case, especially if sub cloud providers are used (for instance, a user signs up with the service "Box" providing a cloud service, and Box uses Amazon as a sub cloud provider).	It has to be clear and understandable for the user who the responsible parties are and how they can be contacted in case of disputes.	<p>Clearly display the contact address of responsible parties on the top layer of multi-layered policies.</p> <p>Redress tools to be developed in A4Cloud have to support end users in contacting the data controller or responsible party.</p>

R.1E	There is a lack of transparency along the chain of (cloud) service providers in regard to their location and applicable laws. The main services providers that are contacted may be located in Sweden, while back-end (Cloud) service providers are located in another country.	Users have to be informed about the country and legal regime of the data controller and data processors along the cloud chain.	Policy icons illustrating the storage location (e.g., inside or outside EEA) and/or legal rules or practices.
R.1F	Web services that target their business to Swedish customers (by having a Swedish website, a Swedish telephone support number, using SEK as a currency, etc.) fall under Swedish consumer and data protection laws, even if the business is located outside of Sweden and independent of what contracts say.	User should be informed about the applicable (national) consumer rights. Redress tools should (at least in these cases) allow users to contact the data controller in their natural language.	
R.1G	Services (such as hotels.com, resia.se) operate only as a mediator/broker, but take no responsibility if something goes wrong. Service brokers have to inform the users about who is the responsible data controller/service provider, with whom the agreement/service contract is actually made.	User interfaces of service brokers have to clearly inform the users about the identity of responsible data controller/service provider with whom the contract is made.	
R.1H	Individual users find it difficult to read and understand long and complicated contracts/terms & conditions that are posted online. Often data loss/unavailability of data is the greatest of consumer concerns, but limitations of availability (in terms of the amounts of time that data are accessible) mentioned in terms and conditions are not transparent to them.	Users have to be aware of and understand important service limitations	Use of UI elements for making users aware, e.g. suitable icons.

R.1I	It is often unclear for individual users what cloud providers really do with the data (e.g., if they are merging different registers) and whether they are following negotiated policies and contracts.	Users should understand data processing purposes and consequences. Users must be informed about serious risks of non-compliance and what this may imply before they disclose data, and about privacy breaches/non-compliance in regard to data that they disclosed.	Present consequences by "Speaking the user's language".
R.1J	Security and privacy risks are not very clear and comprehensible to many individual users. Even security incidents have no long lasting impacts on the user's risk awareness. On the other hand, they are not interested in policy details but just would like to know whether their data are "safe"	Users should be able to understand risk evaluation results, especially if they describe serious risks of non-compliance, and they should understand the implications before they disclose data. They must be informed about privacy breaches/non-compliance in regard to data that they disclosed, in a way that they are aware of and understand those risks.	An overall risk evaluation results can be displayed in a noticeable way, using a multi-layered structure (Art. 29 Data Protection Working Party 2004). The presentation is based on suitable metaphors.
R.1K	At the time of service registration, end users do not think about how to end the service in the future. While the registration for a service is usually made easy, it is often (made) difficult for end users/organizations to unregister/terminate a service contract, delete data or transfer data to other service providers. It is not always clear to end users whether they "own" their data, as they do not check the terms and conditions carefully.	Information about service termination, data deletion and portability should be easily accessible and comprehensible for end users.	Clearly present information about the option and rights of deletion and data portability in the context when it is relevant (e.g., when a service is terminated).
R.1L	It is difficult for individual and business end users as well as auditors to track data in the cloud and to find out who has or has had access to the data for what purposes.	There should be usable and selective audit and transparency tools which even make the handling of implicitly collected data (e.g. via the Facebook Like button) transparent.	Different visualizations of the users' previous data disclosures could be applied, using, for instance, a timeline view or a trace view.
R.1M	SLAs of different cloud services along the chain may not match.	Tools for auditors and business users should visualize the differences between different SLAs	Display a visual chain of SLAs and indicate with colors or icons when there is a mismatch of SLAs. Let users click on a particular mismatching connection to see the details and support his decisions.

R.1N	Users have the need to classify their data or groups of data (e.g., by marking sensitive personal data, confidential data). Data classification is needed in particular for risk analysis and by policy tools.	Users should be guided when defining and editing labels to classify their data in an easy and meaningful way. Moreover, the user should be able to browse through these data by the defined categories.	Provide a filter that allows users to select which categories (labels) are displayed. A tree view can be provided where users can check/uncheck the data to be shown. Alternatively, use tabs to divide the different categories.
-------------	--	---	---

5.1.3 Focus groups: advanced vs. lay users' mental models and attitudes of cloud services

To understand the different ways in which individuals with different levels of familiarity with technology perceive cloud services and comprehend the flow of their personal data on the Internet and in the cloud, we conducted three focus group sessions, one pilot session, one session with only expert users and another session with non-expert users.

The group of expert users was formed of 16 Ph.D. students in computer science coming from different Swedish Universities (but with different nationalities) who were taking a graduate course on the topic of Privacy Enhancing Technologies. It was assumed that these participants would have a similar level of understanding and experience as, for instance, system administrators or IT security professionals dealing with data handling and protection in Internet services. The non-expert users consisted of a group of 15 individuals from different age ranges, cultural and educational backgrounds, who were participants of a project for personal development towards employment opportunities⁷. Our collaboration with such project gave us the opportunity to carry out a focus group session.

During the focus group session participants were divided into different groups of approximately 3 to 4 people. They were asked to brainstorm about how their data were handled and transferred between common Internet services that they commonly use and that have required them to submit personal information (e.g. creating accounts, storing files, buying products, etc.). Each group wrote down these services in post-it notes of a given colour. Thereafter, a card-sorting exercise was performed in which all participants collaboratively classified the services that all groups had come up with into different categories and post it on the blackboard, and gave each category a name. This was done to find probable differences in people's beliefs in the kind of services that can potentially store, handle and share their personal information. Then, each group was asked to choose one of the online service providers and think about the information attributes that are required from the service they had chosen and write them down in a piece of paper. At the end, they were asked to discuss which other online services they believe could also get their personal information when carrying out a transaction with the chosen service and where attacks to their personal information can occur. This was done to get an idea on the users' mental models of how their personal information flows, other parties involved in a digital transaction and vulnerabilities of the transaction. At the end, participants were asked to complete a short post-questionnaire.

The focus groups session resulted in a series of illustrations from each group which resembled the way they visualized how personal information was being exchanged, the entities involved, when carrying out an online transaction, and the vulnerable spots of the transaction. The illustrations were then interpreted, annotated and analysed. Figure 1 shows an annotated illustration of one of the groups from the expert users' focus group session.

⁷ The project is called UMA (Utveckling Mot Arbete) taking place in the city of Kristinehamn, Sweden.

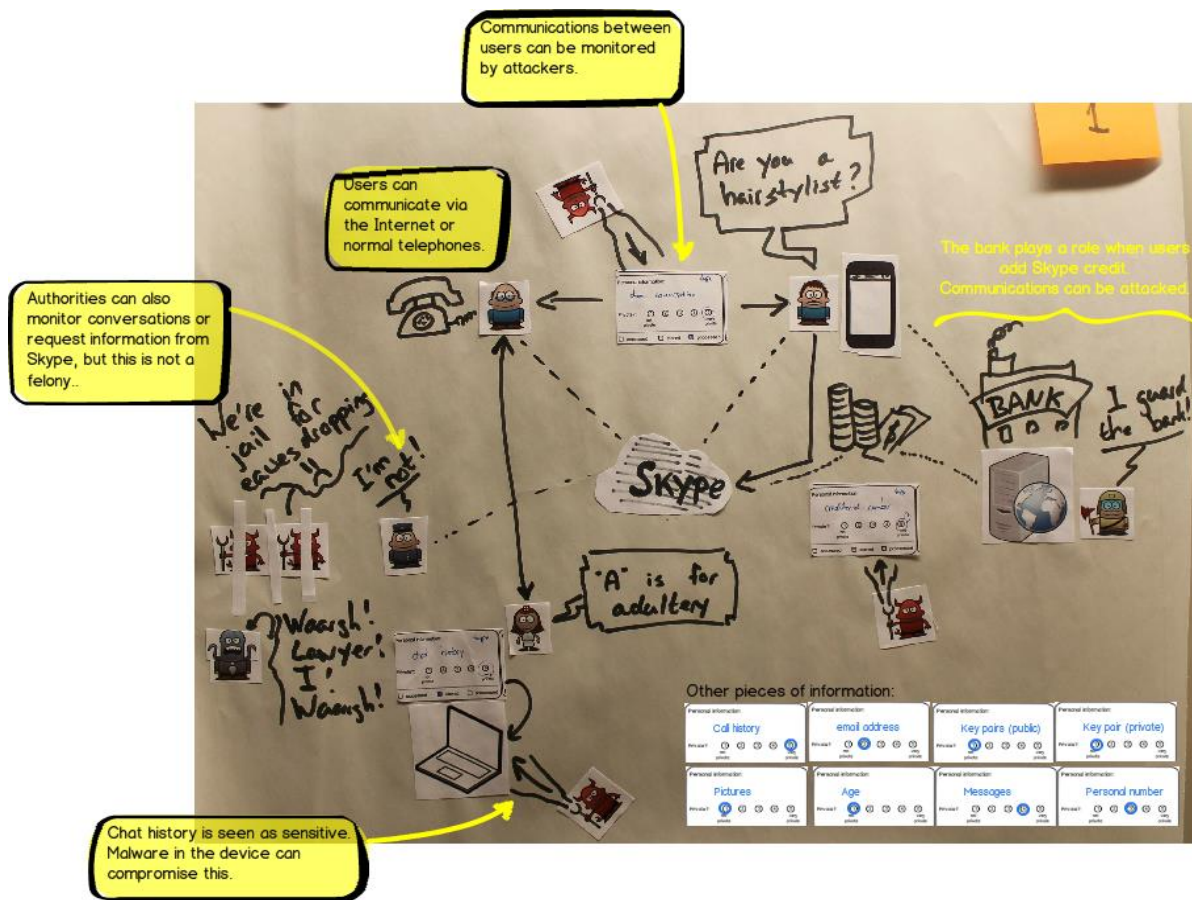


Figure 1. An illustration from a group of expert participants showing the entities involved in a transaction using the Skype service.

General comparison of the illustrations showed, as anticipated, that the participants considered as non-expert have a blurrier idea of how communication between the different entities work in reality, whereas expert participants have a much better understanding of the possible entities involved and the possible vulnerabilities that can occur in a digital transaction. Also, expert participants illustrations tended to go beyond relationship diagrams but they also included democratic statements, such as the power injustices, ideals of transparency, the control of information by powerful service providers, etc. The following table captures the results from the exercises done during the focus group sessions and maps them to UI requirements for the design of possible interfaces for protecting privacy and enhancing transparency.

Table 6 below summarises the results in terms of our observations from the focus group sessions, elicited HCI requirements and proposed HCI principles or design examples.

5.1.3.1

Table 6. HCI requirements and design ideas obtained from focus groups

Req #	Observation (or Problem)	HCI Requirement	Proposed HCI principles and/or sample design solutions
-------	--------------------------	-----------------	--

R.2A	Non-expert users believe that acting entities are more related to each other than they might be in reality. Tendency to believe that personal information is distributed among many of the entities represented. <i>"All internet companies can share information about me".</i>	The interface should clearly show the different entities that could get a hold of which kind of personal information.	Create a network visualization that clearly shows the entities (nodes) getting users' information and the pieces of information that each entity has (as the links).
R.2B	Both groups have an idea that data are being forwarded to third parties by service providers. However, non-expert users seem to have a less clear idea of who these third parties may be.	The interface should put emphasis on explaining the distribution of information to third parties in a clear way. The interface should explain that sometimes the third parties are not specified by the service provider. Present the purposes for which these third parties are allowed to use the data.	
R.2C	Expert users have a clearer idea of where attacks can happen and of possible counter measures. Non-expert users had an idea that information can be at risk, but it is very unclear for them what can be attacked, why is the information vulnerable and the approaches to mitigate the problems.	Lay users need help creating correct mental models of what is vulnerable/risky and what is safe. They should be able to understand when they are performing risky actions and feel comfortable or confident when their risks are minimal. Communicate risks by showing consequences of behaviours in a minimalistic way	Indicate different risk levels with colours and clear explanations. Use adequate language that would communicate the right message to the right user group. Provide layered explanations in an understandable way that can be read in more detail if users are interested, thus catering for the different experience of users.
R.2D	Non-expert users have the idea that their information is collected in a central repository (e.g. a cloud service), but they don't know anything about that repository (how secure it is, where it is, who controls it, etc.).	Inform lay users about some of the details of the location of their data and the properties that apply to it at that location (security, legislations, rights, etc.)	Provide short concise explanations of different aspects of the data in playful ways. Use icons to represent different things, maps to represent locations, and use a multi-layered approach for providing more information when desired.
R.2E	Both groups are aware that service providers can do analysis of their data to find out more information about them. However, non-expert users are less aware of the consequences of the possible misuse of their data.	Users could be informed about some of the possible inferences that a service provider (or a group of service providers) can make based on their previous and current data disclosures.	Show how different data items can be linked together to form new information or deduce information about them which they might not like to disclose. A series of small network visualization can be done showing common examples of combinations of data that can reveal more than people can imagine.

R.2F	Both groups are aware that it is not only the explicit release of personally identifiable information, but also what can be deduced from the data (like behaviours, attitudes, etc.). Difference between explicit and inferred data.	Show people the data that they have disclosed explicitly, and show some of the possible interpretations that a service can do based on that data.	Show a form where people enter data. Then a tool will present a list that shows the possible inferences about their behaviour and personal data based on simple search terms.
R.2G	Trust in the chain of services can play a role in a transaction or disclosure instance. Users can have misconceptions about the trustworthiness of a service based on the trust that they put on another service belonging to the chain.	Users should distinguish when they are interacting with a trustworthy service and be aware of the trustworthiness throughout the chain of cloud services.	Let users judge their trust level by presenting a visualization of third party services that the service provider has contact with.
R.2H	Expert users' concerns go beyond the use of personal data, but deal also with people's rights and democratic governments. Non-expert users are less aware of their rights concerning the protection of their data.	Interested users should be able to audit the chain of clouds. Who has accessed data, for what purpose, why did they access those data at a particular instance, with whom data were shared with, etc. It should be easy for people to exercise their rights regarding data protection and handling practices.	Make users aware of their rights with links to information, and help them exercise them by providing them with clear options for action and Show a list of logged data that users can query with various questions related to their personal information. Queries would filter relevant results. Display a visualization of the chain of clouds and their vulnerabilities.

From the observations of the focus groups it can be concluded that user interfaces for accountability and transparency in the cloud should adapt to the type of user that is interacting with them. For instance, using progressive disclosure, content-on-demand techniques, and multi-layered approaches, descriptive information can be shown only when users request it. Similarly, providing appropriate defaults can release non-expert users from having to modify settings for features and display options, while expert users can customize these options if they want. Also, it can be a good idea to provide different views that appeal to different types of users for displaying similar information. For example, non-expert users might like a more graphically colourful and interactive visualization of data releases, whereas expert users might prefer a log in form of a list of text that they can query. Learnability aspects can also be considered, in which the interface should promote the learning of the novice users so that, if interested, they can reach a higher level of understanding of what goes on with their data in the cloud.

5.2 Usability tests and controlled experiments

5.2.1 Background: Mental models of privacy and control of personal information

Previous studies have investigated the relationship between privacy concerns and the perceived control people have over their personal information on the Internet. For instance, Xu (2007) describes how the introduction of privacy-enhancing technologies (PETs), government legislations and industry self-regulations are factors that increase users' *perceived* control over their information, and thus

mitigate privacy concerns. Similarly, Hoadley et al. (2010) investigated the privacy concerns of Online Social Network users when an illusory loss of privacy control was introduced to a social network platform, suggesting that users' who believe their information is more accessible to others will present higher privacy concerns and show more willingness to adjust their privacy settings. Additionally, studies by Brandimarte et al. (2012) revealed effects on privacy concerns where increased *perceived* control over the release of personal information also increases the willingness of people to keep releasing sensitive information. That is, people often *perceive* that they are in control over their data releases without attaining *actual* control, nevertheless this illusory sense of control leads them to publish more sensitive information. (i.e., "more" perceived control can lead to "less" privacy in reality).

These studies suggest paradoxical and irrational behaviours by people when it comes to the value they place to their privacy when acting online. In particular, people who have an illusory sense of control over their data are less likely to protect their privacy in reality (Gross & Acquisti 2005). Besides, people seldom have an accurate perception on the actual amount of control they have over their information.

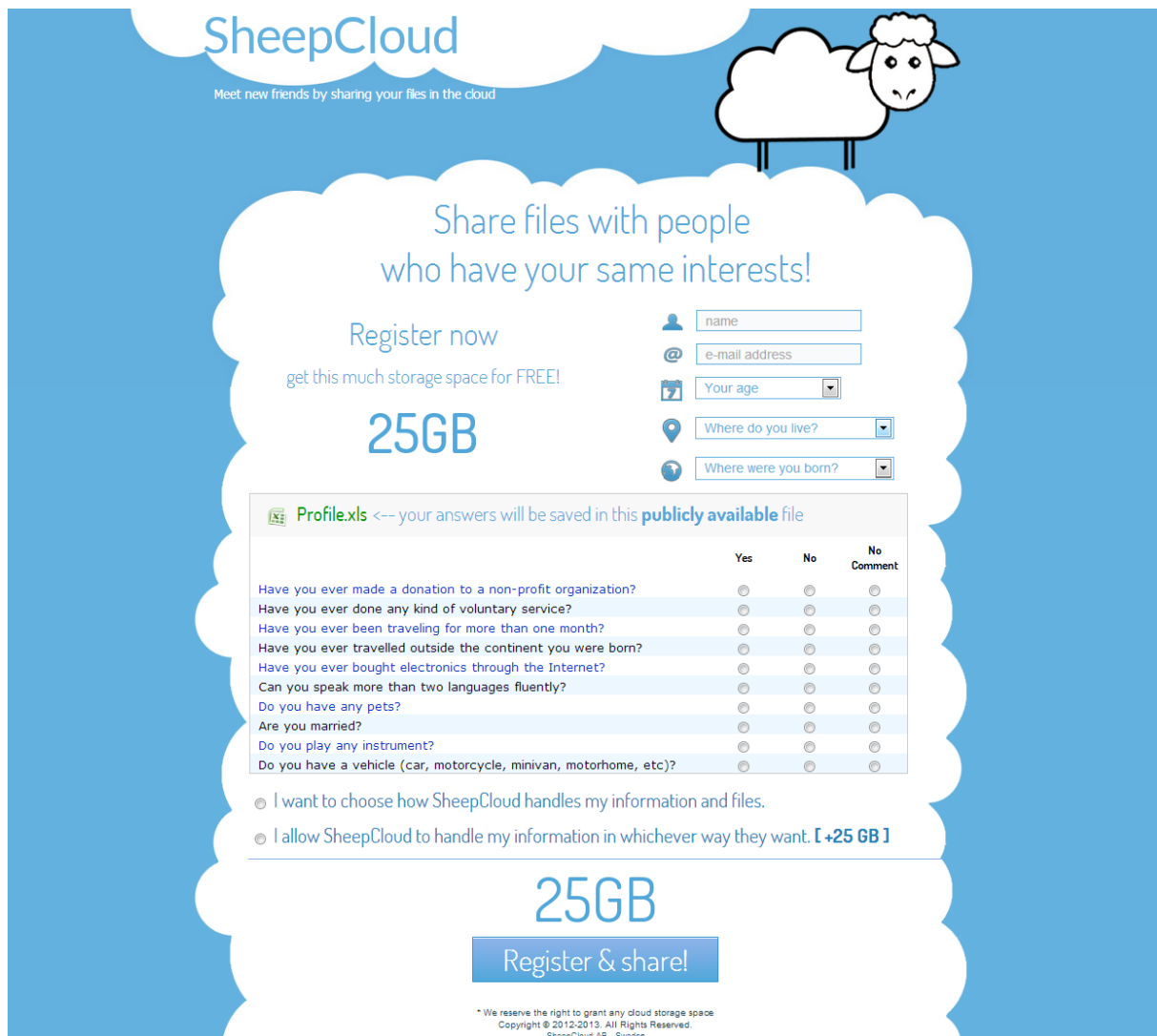
Moreover, Ion et al. (2012) have shown that individual end users of cloud services have strong privacy concerns, trust local storage devices more than cloud storage when dealing with sensitive data, but are not fully aware of the risks posed by cloud storage services. Marshall & Tang (2012) explore cloud services as file synching and sharing mechanisms, identifying five common use cases among individuals, including using the cloud as a repository to exchange files between own devices, using the cloud as a shared repository to collaboratively edit content in the cloud, backing up and editing content of own files offline, editing content of files reflects in others' devices, and synchronization of files.

The finding of the studies mentioned above can be complemented by investigating not only how "perceived control influences people's willingness to reveal personal information" (Brandimarte et al. 2012) but also the extent to which people are willing to reveal personal information in exchange for perceived control and other valuables, such as comfort and less cognitive burdens, or more cloud storage space and transparency features. There is also a need for understanding the kinds of features that cloud consumers need and appreciate at the moment of protecting the data stored and handled by cloud service providers.

5.2.2 Exploring users' behaviours, needs and understandings through controlled experiments

Previous studies have found paradoxical privacy behaviours of people when acting online, stating for instance that individual's desire for privacy is not necessarily reflected by their real actions (Spiekermann et al. 2001; Gross & Acquisti 2005). Motivated by the design of the investigations of these previous studies we set up a series of three experiments that had the intention of understanding the way people think about cloud storage services, their willingness to distribute their personal data (information and files) to cloud services, how they perceive and understand related risks and control options, the amount of trust they put in the service, the features and controls that they appreciate, and other factors related to the distribution and understanding of their information in the cloud. Analysing these factors can provide insights on how ex ante TETs can be designed to support users to make better informed decisions and to exercise control of the use of their data in the cloud.

The experiments were based on a scenario representing a fictitious cloud storage service, which we named "SheepCloud". An illustration of the registration page of this fake service is shown in Figure 2, in which participants of the experiments were made to believe that they were registering and submitting personal information to an unknown cloud service.



SheepCloud
Meet new friends by sharing your files in the cloud

Share files with people who have your same interests!

Register now
get this much storage space for FREE!

25GB

name
e-mail address
Your age
Where do you live?
Where were you born?

Profile.xls <-- your answers will be saved in this publicly available file

	Yes	No	No Comment
Have you ever made a donation to a non-profit organization?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Have you ever done any kind of voluntary service?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Have you ever been traveling for more than one month?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Have you ever travelled outside the continent you were born?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Have you ever bought electronics through the Internet?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Can you speak more than two languages fluently?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do you have any pets?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Are you married?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do you play any instrument?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do you have a vehicle (car, motorcycle, minivan, motorhome, etc)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

☐ I want to choose how SheepCloud handles my information and files.

☐ I allow SheepCloud to handle my information in whichever way they want. [+25 GB]

25GB

Register & share!

* We reserve the right to grant any cloud storage space
Copyright © 2012-2013. All Rights Reserved.
SheepCloud AB - Sweden

Figure 2. SheepCloud registration page. Users were made believe they were registering and releasing personal data to a new storage cloud service.

At each test session participants were asked to carry out through the following general steps, which were consistent across the different series of experiments:

1. Read introductory instructions:

As a first step, participants were directed to a webpage with instructions about the test. The instructions deceived participants into believing that they were going to register for a Beta test of a new cloud service with social capabilities.

2. Submit registration information and answer to questions:

Participants were asked to register by submitting some personal information (name, email, age range, place of birth, place of residency and profession). Participants were also told that in order to register they would need to answer a set of 10 questions with possible answers being “Yes”, “No” or “No comment”, and that their answers would be stored in a public file in their new storage space. Participants were assigned at random into two groups, in which the experimental group was shown ten questions that can be considered sensitive, and the

control group were shown ten questions that were considered non-sensitive. In order to minimize any possible bias due to the order of the questions, the order of these questions was also randomized. Moreover, the time participants spent in the registration page was also measured, as an indicator of the effort it took for them to register and take decisions.

3. Experimental part

This part of the design was varied between and within the three experiments. In experiment one the amount of gigabytes of storage was varied between the groups. In experiment two the choice between automatic and manual registration was framed in different ways between the groups. Finally, in experiment three the focus was to investigate which privacy features are desired by users.

4. Read debriefing information and confirm participation:

Once participants thought they had submitted their information and answers to the questions, they were then explained that SheepCloud was not a real cloud service and they could not get storage space. They were told which information was about to be collected and which personal information was not being collected by the experiment. In particular, no personal data⁸ was collected (such as name and email address), neither were the specific answers to the questions collected (only the total sum of “Yes”, “No” and “No comment” responses).

5. Answer a post-questionnaire:

After being debriefed, participants were asked to answer a few more questions⁹. The questions intended to measure aspects related to the credibility of the SheepCloud scenario, the sensitivity of the questions asked, the level of trust people will put in an unknown service, their online privacy concerns and behaviours, and other aspects surrounding the specifics of the test.

The idea of deceiving test participants into believing that they were about to disclose real personal data to an unknown cloud service had the purpose of improving the ecological validity of the experiments. To check if participants actually felt that they were submitting sensitive private information to this new cloud service when registering, the post-questionnaire asked participants to rate in a scale from 1 to 5 the level of sensitivity of the questions. A chi-square test of independence for the different test scenarios revealed that participants did indeed differentiate between questions that were sensitive and questions that were non-sensitive. The questions asked can be found in the Appendices. Although the use of deception in research is to be avoided based on the principle of informed consent it is permissible when the research question cannot be answered otherwise. According to the American Psychology Association¹⁰ ethical guidelines as well as the Swedish Etikprövningslagen¹¹ when using deception, researchers must ensure that participants do not experience physical pain or severe emotional distress and make sure that the participants are fully debriefed about the motive of deception after the data collection. In the reported experiments not using deception would have led to hypothetical responses from the participants. Furthermore all participants

⁸ As defined by EU Directive 95/46/EC

⁹ The specific post-questionnaires for each experiment can be accessed at these sites:

- Experiment 1: <http://goo.gl/Hco7S>
- Experiment 2: <http://edu.surveygizmo.com/s3/1139746/SheepCloud-v2-Survey>

¹⁰ <http://www.apa.org/ethics/code/principles.pdf>

¹¹ http://www.riksdagen.se/sv/Dokument-Lagar/Lagar/Svenskforfattningssamling/Lag-2003460-om-etikprovning_sfs-2003-460/

were debriefed regarding the deception and were again asked if they willingly conceded to continue with the study now that they knew the true set up of the experiment.

5.2.3 Experiment 1: Understanding willingness to distribute personal data to cloud services

In this first set of experiments we investigated the willingness of test participants to disclose personal data depending on the offers they get from a cloud storage service provider. In other words, would people be willing to give away control over their personal information if they perceived that they could get more value from the cloud storage provider?

To explore this question we used the SheepCloud scenario offering participants 25 GB of cloud storage space at the moment of registration, with the possibility to earn more storage if they were willing to hand over the control over their personal data. Participants were recruited on Karlstad University's campus as well as online through an international crowdsourcing service¹². On campus, participants were rewarded for their participation through small tokens of appreciations (such as candy or USB storage sticks), whereas remotely located participants got a small sum of money if their participation was satisfactory. The data were analysed to detect and exclude remote participants who did not take the test seriously. A total of 120 participants completed the test successfully.

During a test session participants, who were randomly divided into two categories of either sensitive or non-sensitive questions, were further assigned to three other subgroups at random, where they could get different *additional* amounts of storage space if they were willing to hand over control of their personal data and files to the cloud service provider. The three groups of additional storage space offers were:

- Group 1: No additional extra storage offered - control
- Group 2: Double the initial storage offered (+ 25 GB)
- Group 3: Large amount of storage offered (+ 100GB)

Figure 3 shows an example of group two, where a participant can double the initial offer of 25GB of cloud storage if she chooses the option "I allow SheepCloud to handle my information in whichever way they want [+25GB]", thus getting 50GB at the time of registration.

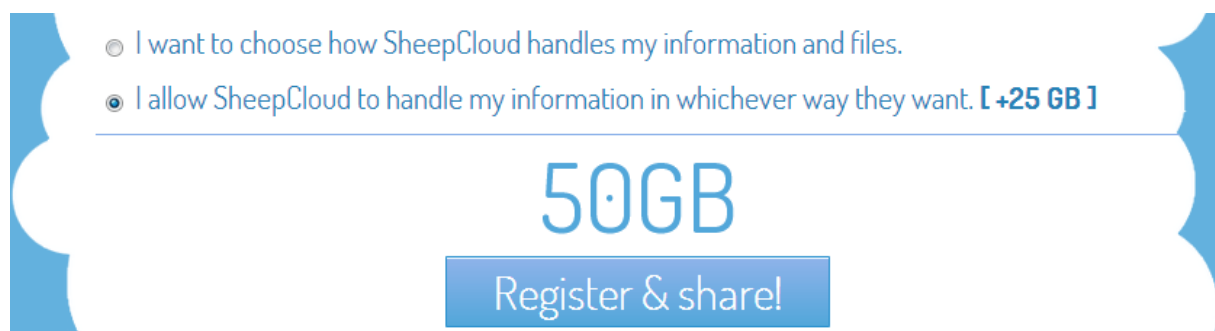


Figure 3. Example of an offer to get double the cloud storage space if the user hands out control of his personal information to the cloud service provider.

A logistic regression analysis showed no statistical significant influence on the willingness to control personal data by neither the level of the sensitivity of the questions, $p > 0.58$ (i.e. looking only at participants who were assigned to either sensitive or non-sensitive groups), or by the amount of cloud storage space offered, $p > 0.06$. No significance was found by the combination of these two

¹² Microworkers.org

independent variables (amount of cloud storage offered by the level of sensitivity), $p > 0.26$ (i.e. looking only at participants who were assigned to the three different GB storage offers groups).

Descriptive statistics of the results from this experiment are shown in Table 7, where no immediate obvious variations can be observed between the different groups.

Table 7. Crosstabulation of the willingness to control data depending on the sensitivity of the data and the amount of storage offered.

		Choice of control		Total
		User control	Cloud control	
Sensitive	+ 0 GB (control)	6	2	8
	+ 25 GB	15	16	31
	+ 100 GB	15	12	27
Non-sensitive	+ 0 GB (control)	13	3	16
	+ 25 GB	11	11	22
	+ 100 GB	4	12	16
Total		64	56	120

However, considering only the participants who were assigned to the non-sensitive data group as the baseline, it can be observed that there might be a difference on the willingness to control data depending on the amount of storage offered. In fact, calculating the chi-square statistic, $\chi^2(2, n = 54) = 10.19, p < 0.01$, it can be seen that the level of storage space offered is a predictor on the users' choice of control over their data in this case. Nevertheless, this is not the case when looking only at the participants assigned to the sensitive data group, $\chi^2(2, n = 66) = 1.84, p > 0.39$.

As mentioned above, participants were debriefed after registering for the fictitious cloud storage service, and they were asked to answer some additional questions. The answers to the post-questionnaire reveal that a 70% of participants who stated that they have never heard the term "cloud computing" did in fact use at least one cloud service, indicating that people do not comprehend the meaning of their data being stored in the "cloud". Also, the results indicate that 41% of participants would appreciate a lot if a service provider made it easier to understand their legal rights with regards to the use of their data.

Even though there is no concrete proof that people are willing to give away control over their personal sensitive information, the results from this experiment suggest that there are some factors that can influence a user's decision to adapt a certain cloud service, since they might be willing to give away some of the data that is not seen as sensitive in exchange for perceived valuables. These findings also indicate that a design effort has to be done for motivating users to become aware of the data they are releasing to unfamiliar cloud services and what the consequences can be, since users seem to be indifferent to whether they are releasing sensitive or non-sensitive data into the cloud. Moreover, there is a need to educate users on what the term "cloud service" entails and the implications of subscribing to such services.

Table 8 below summarises the results from experiment 1.

Table 8. HCI requirements and design ideas obtained from Experiment 1

Req #	Observation (or Problem)	HCI Requirement	Proposed HCI principles and/or sample design solutions
R.3A	Perceived sensitivity of data can influence people's behaviours in regard to exercising control. However, data that might be perceived as non-sensitive (or harmless) can become sensitive with changes in time and/or context.	<p>Users should be informed about possible scenarios in which data items could become sensitive.</p> <p>Users should also be aware about the different purposes for which their information might be used, as well as the possible recipients of their data, since this can affect their behaviour. The perceived sensitivity of data can be dependent on the context in which it is used.</p>	<p>On the user interface, provide inline examples of data aggregation or misuse of seemingly harmless data.</p> <p>Provide a visual indication of how their data might be transferred across the chain cloud or shared with third party services.</p>
R.3B	Users are willing to disclose personal data that is perceived as non-sensitive in exchange for a reward that seems valuable.	Users should be made aware of the risk and benefits of disclosing their data to a service.	Make users conscious about the value of the data they are releasing comparable to something they can relate to, like monetary value.
R.3C	Users are unaware or not well informed about the types of online services they subscribe to in regards with the handling of their data and personal privacy.	<p>Cloud providers should inform individual end users about the services' privacy policies and make the implications of data disclosures transparent to these users.</p> <p>Ex ante transparency awareness should be promoted, in order for users to know what type of service they are subscribing to.</p>	Make it explicit through the wording and the use of standard icons the consequences in terms of benefits and risks of having personal data in the cloud.

5.2.4 Experiment 2: Framing and terminology

Previous research on decision-making has shown that people are very much influenced of the description or framing of the problem (Tversky & Kahneman 1985). The aim of the experiment was to investigate influence of framing on the choice between preserving versus giving up one's privacy.

A total of 121 participants (62 female, 55 male (4 missing), 99 aged 19-30, 22 aged 31-60) partook in this experiment. All participants were recruited in public areas of the Karlstad University campus. The general design and procedure followed the outline described in Section 5.2.2.

At the end of the fictitious sign up process to SheepCloud the participants were asked to submit some personal information either manually (and thus retaining full control of their data) or automatically (without control of their data) by letting the system gather information that other popular Internet service already possess about them, such as the information on Facebook. Three conditions of framing were tested. One baseline, without any framing, one where the automatic option was framed as a way to save time and finally one where the manual options was framed as a way to control what data were actually being submitted. The participants were randomly assigned to one of the three conditions.



Figure 4. Screen shot from the baseline condition. In the two experimental conditions the subheading was changed to frame one of the two choices in a positive way.

The results of a Pearson chi-square test showed a significant effect $\chi^2(2, n = 121) = 13.20, p < .001$ of framing on the proportions of participants choosing either manual or automatic registration. A comparison of the three frames showed that participants were significantly (Bonferroni corrected p 's $< .05$) more willing to give up control when the automatic choice was framed as "time saving" compared to both the baseline and the control frame. There was, however, no significant difference $p < .05$ between the baseline and the control frame conditions (see Table 9 for descriptive statistics).

Table 9. Descriptive statistics showing the number of participants assigned to each conditions of Experiment 3.

Level of control	Frame		
	Full control	Save time	Baseline
	(n = 37)	(n = 41)	(n = 43)
High (manual)	81.1%	46.3%	76.7%
Low (automatic)	18.9%	53.7%	23.3%

The results clearly show the influence that the framing of the choice has on preserving versus making compromises on one's privacy. The effect of the time frame can be understood in two different

fashions: either in terms of effort or in terms of clarity of consequences. In terms of effort, the offer to save five minutes by compromising on one's privacy might not sound too attractive at first. However, given the time usually spent on a registration process in combination with the fact that privacy is a secondary task to, in this case, acquiring storage space, the offer to do so more effortlessly is rather alluring. In terms of known consequences, previous research has shown that concrete information on a given option has a very strong influence on choice (Tversky & Kahneman 1985). Other findings indicate that people's degree of perceived control is frequently overestimated (Langer 1975).

Table 10 below summarises the main results from experiment 2.

Table 10. HCI requirements and design ideas obtained from Experiment 2

Req #	Observation (or Problem)	HCI Requirement	Proposed HCI principles and/or sample design solutions
R.4A	Users' willingness to release personal data is influenced by the description of two or more choices.	Make users aware of all pros and cons of their choice in an unbiased fashion	Tooltips and/or help texts to clarify consequences of actions.

5.2.5 Experiment 3: Desired features on cloud services

The results from previous activities, namely the Stakeholder workshops, focus groups and the first round of experiments using the SheepCloud scenario, provided us with a narrow set of features of what users would value or what stakeholders representing user interests recommend for providing user control and transparency. In short as listed in Section 5.1.1, these features included the possibilities for data portability, specifying levels of visibility of data, specifying the data locations in the cloud and privacy and consumer laws applied in these locations, specifying the permissions for data usage and sharing, categorizing data, and defining the level of privacy and security for different data items. Most of these features were also suggested by stakeholders from consumer or data protection authorities (explained in Section 5.1.1) based on experienced privacy and consumer-related issues that arose to users when these transparency and control features were absent or when users were not sufficiently aware of them.

We saw it as relevant to understand which of these features individual users care about in order to suggest design guidelines that prioritize the availability, learnability and ease of use of such features. In particular, if individual users do not directly consider features listed above that were recommended by consumer organisations or data protection commissioners based on long-term experiences as valuable, more efforts may have to be put into the UI design for conveying the consequences of issues that may arise in absence of these features. In other words, instead of informing about more abstract technical terms, the UIs of A4Cloud transparency and control tools should rather mediate the practical advantages that these control options offer to the users.

For this reason, we performed another round of experiments using the SheepCloud scenario, where, as in the first experiments, participants were randomly selected into two groups, one group was asked a series of sensitive questions that were supposedly stored as a file in the cloud provider, and the other group was asked non-sensitive questions.

In this case all participants were recruited through the same crowdsourcing service as in the first experiment and were rewarded about \$1 if they completed the test satisfactorily. As in one of the previous experiments, the data entries were screened for possible demotivated participants. Submissions that were not judged as honest or seemed rushed were removed from the dataset. At the

end, the entries from 179 participants coming from different parts of the world were considered for this experiment.

This time, people who decided to be in control of their own data were then given the option to select four out of six control features that they would like to have. For every selected feature the amount of free cloud storage offered was reduced by 5GB. The idea of limiting the amount of features to be selected and reducing the storage space offered had the intention of forcing participants to select only those features that they really cared about. Figure 5 demonstrates the look-and-feel of the feature selection using the SheepCloud scenario.



Figure 5. The functions for controlling data that SheepCloud offered at the time of registration.

Out of the 179 people that completed the first part of the test, only 68 (38%) enabled the options for controlling their own information and files (62% of the participants indicated that they would let SheepCloud handle their information and files in whichever way this service provider wants).

Table 11 summarizes the preferred features as selected by the 68 participants who wanted control over their own data. The table presents the textual descriptions about the features that were shown to participants in form of a “cloud tooltip”. Results show that participants would value most the possibility to decide who will have access to their data (i.e. 52.9% would like to have a “Visibility” feature), followed by the power to decide how the cloud service provider will end up using and sharing their data with third parties (i.e. 38.2% value a “Usage” feature), as well as the opportunity to select the levels of security that apply to different data items.

Table 11. The possible features for control of personal data and the participants' preferred features.

Name	Detailed description	Frequency	Percent
------	----------------------	-----------	---------

Visibility	Control who will be able to see your data (Public, friends of friends, friends or only me)	36	52.9%
Usage	Determine the way SheepCloud uses and shares your data with other companies	26	38.2%
Security	Control the levels of encryption of individual data items or groups of data	22	32.4%
Location	Control where your data is stored and the laws that apply	16	23.5%
Portability	Be able to download all your data locally in a standard format	16	23.5%
Labelling	Tag your data with different labels, like “work”, “family holidays”, “high school”, etc.	9	13.2%

As with the first experiment, a Person chi-square test for independence revealed that there is no significant difference between the number of participants that wanted to have control over their own data and the ones giving away their control, when the sensitivity of the data was manipulated, $X^2(1, n = 179) = 0.181, p > 0.67$. In other words, 62% of participants choose to delegate the control over their data and information to an unfamiliar cloud service provider (i.e. SheepCloud), regardless of whether their data were sensitive or not, suggesting that many people are not willing to spend too much cognitive effort at controlling certain aspects of their data. One possible reason for this is that lay users might find it confusing, burdensome or time consuming to select controls that help them protect their data and preserve their privacy. Another explanation is that they might not be well aware about the practical consequences of releasing personal information to a cloud service, and thus they lack the motivation to spend cognitive efforts at setting these controls.

Curiously, results for this experiment also showed that the beliefs users had about SheepCloud being a real cloud service decreased as compared to the results of the first experiment (Section 5.2.3). This is probably because existing cloud services do not give users the opportunity to control their data, and thus people did not perceive this as a realistic service. These results are in accordance with findings described in (Lacohée et al. 2006), stating that unsubstantiated claims given by a service provider do not tend to build trust. Moreover, from mapping the questionnaire responses from participants who came from different parts of the world, it can be observed that there are cultural differences in the amount of trust people would place in an unknown cloud service provider. For instance, respondents from Sweden indicated in average that they would generally *not* trust SheepCloud with their personal data and files, whereas people from Southern Europe and Southeast Asia were more willing to trust SheepCloud with their data and files.

Table 12 summarises the results from experiment 3.

Table 12. HCI requirements and design ideas obtained from Experiment 3

Req #	Observation (or Problem)	HCI Requirement	Proposed HCI principles and/or sample design solutions
R.5A	Users are unmotivated to spend cognitive effort or time at setting up privacy controls.	<p>Users should be motivated to spend the necessary cognitive effort or time at adjusting their privacy preferences at a moment that is relevant to them and meaningful to their actions.</p> <p>Consequences are easier to grasp than technical features and terms. Inform users not only about how settings can be adjusted, but the consequences of adjusting such settings.</p>	<p>Provide appropriate privacy-friendly defaults for a set of situations in order to ease the users' burden of setting privacy preferences</p> <p>Let users adjust their preferences "on the fly" as needed. By providing brief but meaningful explanations as of why it is important to care about such setting in terms of the consequences to the users' privacy might motivate them to care about adjusting.</p> <p>In order to enhance users' comprehension and motivation, a cloud provider should present its privacy-enhancing features in a way that relates to users' everyday reality and try to reduce the technical explanations.</p>
R.5B	Knowing who is able to view/access and see users' data stored in the cloud as well as how their data are used are appealing features.	It should be easy for users to find and adjust functionality related to the visibility and usage of their data.	Provide privacy-friendly default settings for data access controls and usage.
R.5C	People may become sceptical towards unknown services that promise them to guard their privacy.	The cloud provider should motivate not only the benefits for users protecting their privacy, but also the benefits for the cloud provider itself when offering accountable and privacy-friendly features to its customers.	
R.5D	Trust on unknown cloud services might have a cultural component to it. Users from different cultures exhibit different levels of trust.	Cloud provider should consider their customers in terms of the culture, location of service, and legislative regimes and cater for their collective mental models and attitudes towards data in the cloud.	<p>When users are about to subscribe to a cloud service, appeal to their cultural background by emphasising features of security, accessibility and alike.</p> <p>Accountability and transparency features might balance the level of trust across different cultures.</p>

5.3 Evaluating visualizations of data disclosures and data traces

5.3.1 Background

One of the HCI challenges that we have been addressing in WP:C-7 is the question how we can guide the users to better comprehend the flow and traces of their data on the Internet and in the cloud. As was observed earlier in the PRIME and PrimeLife projects, many users have problems to differentiate whether data are stored on the user's side (under the user's control) or on a remote services' side, and to comprehend to which network entities personal data flows during online transactions (Wästlund & Fischer-Hübner 2010). Focus group sessions that we held (c.f. Section 5.1.2) confirmed these previous findings especially for lay users.

Therefore, as a next step we wanted to address the research question regarding what are suitable HCI concepts for evoking the correct mental model for users of their personal data flows and traces. We chose to examine this question by taking the prototype of a tool called the Data Track, which was developed in the PRIME and PrimeLife projects (Wästlund & Fischer-Hübner 2010), as a test case.

The Data Track is a user side ex post transparency tool, which includes both a history function and online access functions. The history function stores in a secure manner for each transaction, in which a user discloses personal data to a service, a record for the user on which personal data were disclosed to whom (i.e. the identity of the controller), for which purposes and under which agreed-upon privacy policy. The Data Track's user interface version developed under the PrimeLife project provided search functions, which allow users to easily get an overview about who has received what data about him, as well as online access functions, which allow end users to exercise their rights to access their data at the remote services' sides online and to correct or delete their data (as far as this is permitted by the services sides). By this, users can compare what data have been disclosed by them to a services side with what data are still stored by the services side, or what data have been implicitly been added (e.g., trust ratings of customers added by an eCommerce side) to the data records stored at the services side. This allows users to check whether data have been changed, processed, added or deleted (and whether this was in accordance with the agreed-upon privacy policy).

Figure 6 shows a screenshot of PrimeLife's Data Track user interface, which displays data that are stored locally in the Data Track as well as data stored at the remote services' side in a single table. Remotely stored data which were equal to data stored locally in the Data Track are displayed in green fonts.

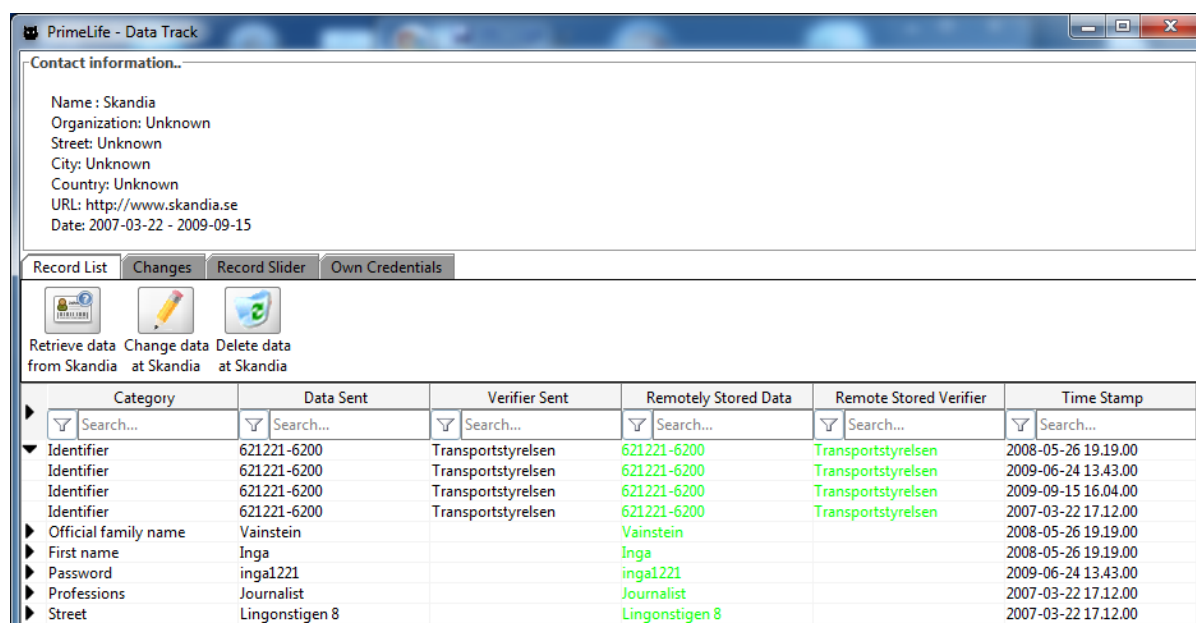


Figure 6. DataTrack user interface developed under the PrimeLife project

Complete descriptions of the Data Track proof-of-concept and user interfaces can be found in (Wästlund & Fischer-Hübner 2010). Usability tests of early design iterations of the PrimeLife's Data Track revealed that many test users had problems to understand from the Data Track table representation whether data records were stored in the Data Track on the users' side (under the users' control) or on the remote service provider's side.

Therefore, in A4Cloud we have tested alternative HCI concepts consisting of graphical UI illustrations of where data are stored and to which entities data have been distributed. Based on the usability heuristic suggesting a "match between the system and the real world" (Nielsen 1995), graphical illustrations of data storage and data flows have a potential to display data traces more naturally as in real world networks, as discussed in the PRIME deliverable D06.1.f, Section 5.8.1 (Pettersson 2008). Besides, previous research studies suggest that network-like visualizations provide a simple way to understand the meaning behind some types of data (Freeman 2000; Becker et al. 1995), and other recent studies claim that users appreciate graphical representations of their personal data flows in forms of links and nodes (Kani-Zabihi et al. 2012; Kolter et al. 2010).

Therefore, a new UI concept for visualizing the users' information in the Data Track tool has been proposed and prototyped by us¹³, as shown in Figure 7. This way of showing the tracking of the users' data has been called the "trace view", presenting an overview of which data (with data attributes) have been sent to service providers, as well as which service providers might have the users' data.

¹³ The development of new graphical user interfaces for the Data Track were co-funded by a Google Research Award Project on "Usable Privacy and Transparency".

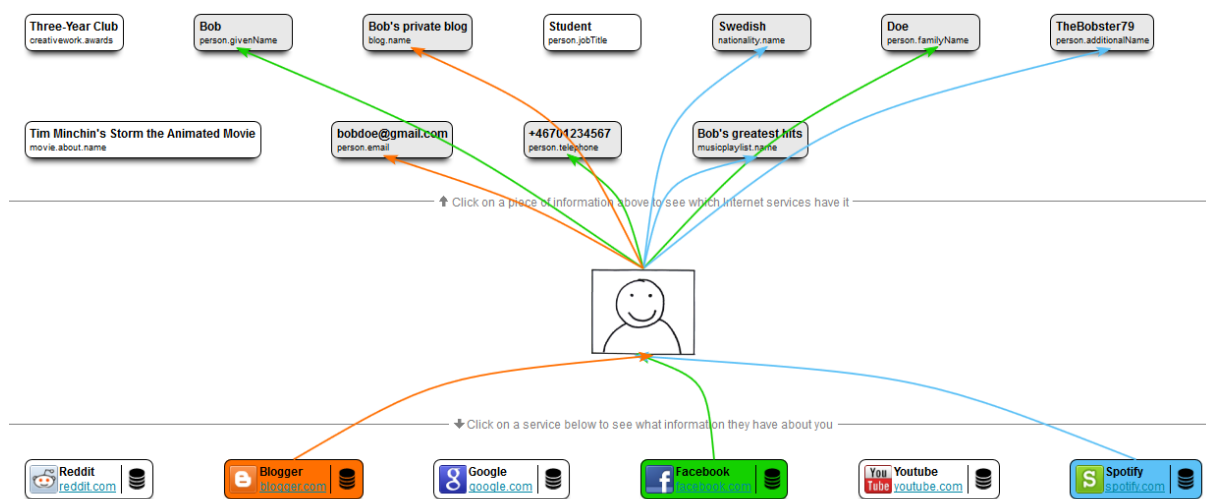


Figure 7. The trace view user interface of Data Track

The idea is that users should be able to see all the personal data items stored in the Data Track (displayed in the top of the UI) that they have submitted to services on the Internet (these Internet services are shown in the bottom panel of the interface). If users click on one or many of the Internet services they will be shown arrows pointing to the information that those services have about them, in other words they can see a *trace* of the data that services have about them. Similarly, if they select one or many data items (on the top), they will be shown arrows pointing to the Internet services that have those data items.

Users can also access the data about them stored on the services sides by clicking on the corresponding icons, and are able to correct it, or remove it if the respective service allows it. Figure 8 depicts an example sketch of the user's data stored at the service's side being showed.

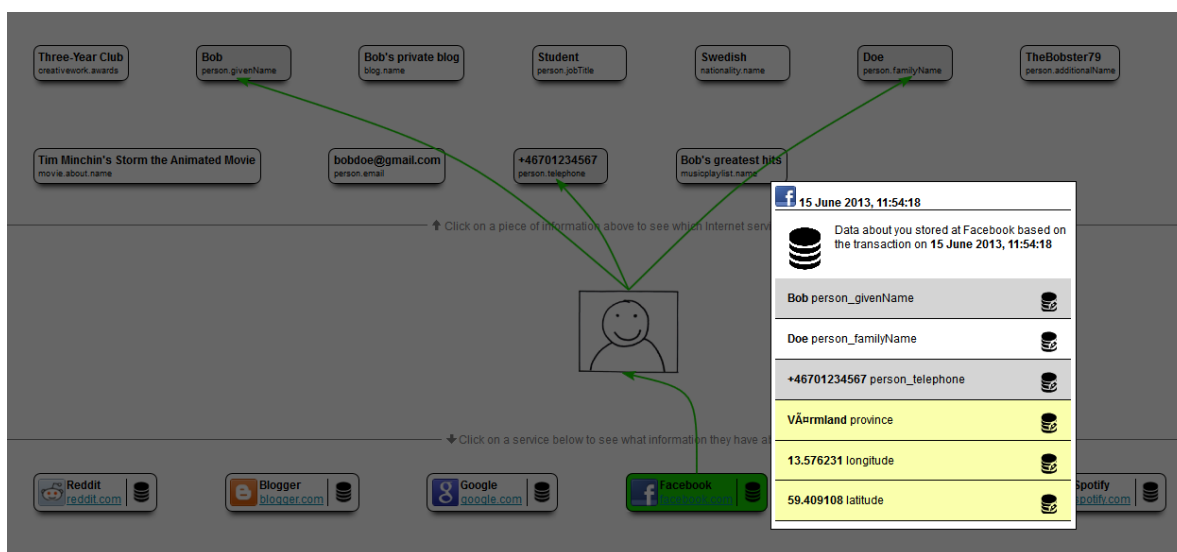


Figure 8. Information about a user that a service provider has stored on their servers (service's side)

5.3.2 Evaluation

In order to test the new Data Track user interface introduced above and the hypothesis of whether users more naturally understand graphical data flow illustrations, we implemented an interactive version of the Data Track's trace view based on the designed sketches and performed usability tests with 14 participants. Carrying out usability tests with a prototyped interface in this case was considered a more appropriate method for getting relevant responses from lay users, since it might be more difficult to reveal these types of users' concerns and needs through other methods, such as workshops described in Section 5.1.

The usability tests were setup using a scenario consisting of a fictitious online book retailer. A total of 14 participants between 19 and 40 years old were recruited in different parts of the city of Karlstad, Sweden. 12 of the 14 participants indicated that they were "experienced" or "very experienced" with computers, 7 of them were working professionals and 6 were undergraduate students (the rest preferred not to state their status). Participants of the tests were asked to read instructions about the test (found in Appendix 3.2), to sign a consent form, and then to pretend that they were purchasing a book from this online book store. In order to complete the transaction they were required to submit some personal data, such as their name, their home address, their email, their phone number, their credit card for payment (none of the information submitted was stored in reality and participants were given a fake credit card number for purchasing the book). After buying the book, participants were shown the Data Track trace view interface and a test moderator asked them to complete predefined tasks using the prototype (the tasks are listed in Appendix A.2).

In order to minimize the introduction of confounding variables in the series of tasks that participants are asked to complete, the order in which the tasks are presented was shuffled at random in every test session, in a technique known as *counterbalancing* (Rubin & Chisnell 2008). A test moderator annotated the observations made by participants and the success rate of the tasks. At the end of the test participants were asked to respond to a post-test questionnaire (Appendix A.2) where they could state their subjective opinions about the program.

5.3.3 Results

The analysis of the participants' responses during the test revealed several interesting results. First, 11 of the 14 participants clearly understood that the elements on the top panel of the interface represented their own information that was sent to online services, and all participants understood that the elements at the bottom represented online services to which they have sent information. Also, it was intuitive for all participants to find out the data items that they have sent to a particular service provider (by clicking on one of the services on the bottom panel). All participants but one found it easy to discover which services had a particular data attribute. These positive initial observations indicate that participants found the tracing feature of the interface easy to understand, intuitive and informative. The answers to the post-questionnaire corroborate that participants understood the basic elements of the trace view user interface, as can be seen from the bar chart in Figure 9.

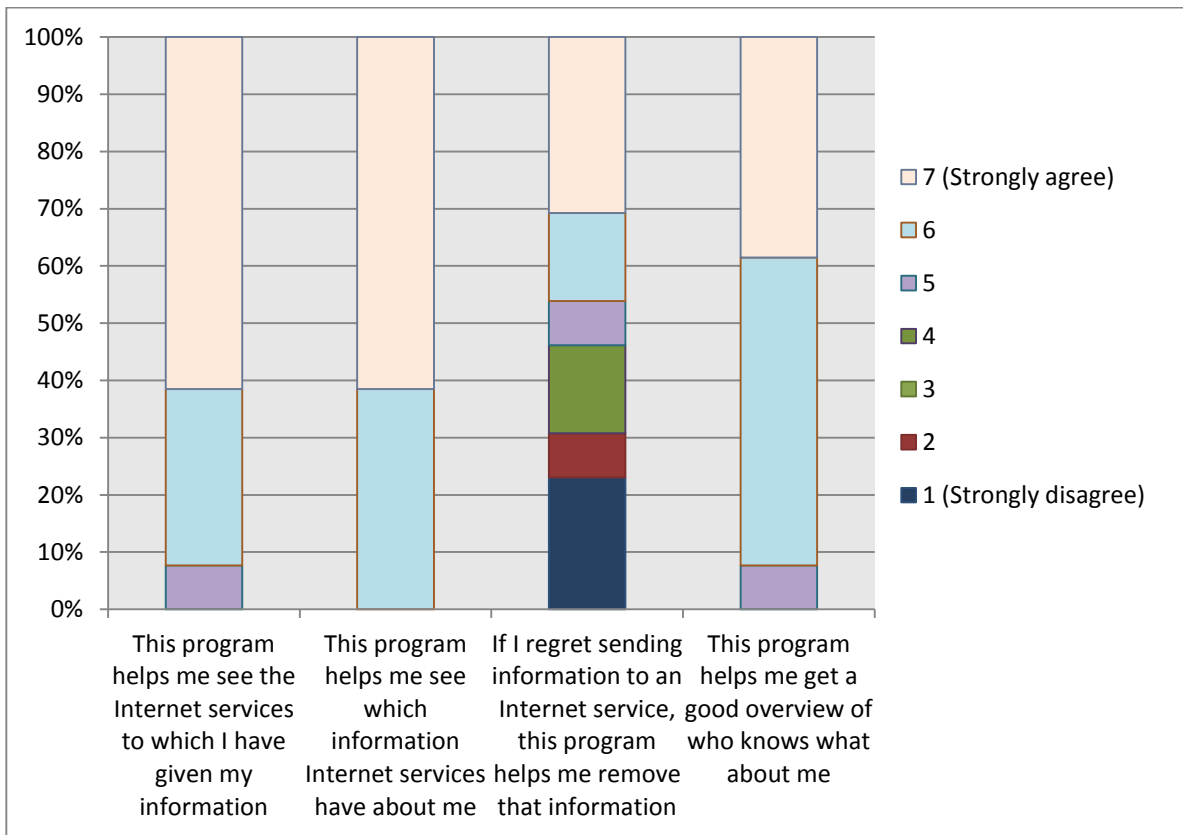


Figure 9. Post-questionnaire scale on the understanding of the Data Track trace view

On the other hand, participants had a harder time understanding that they could also access the data stored about them on the service's side, which was also a challenge in earlier versions of the Data Track interface. When asked the question "Where would you click to see the information that Adbokis has stored on their servers when you purchased the book?", participants did not immediately understand that there was a difference between the data saved by the Data Track program and the data stored on the service's side. Consequently, participants found it difficult to answer the question "What information about you does Adbokis have on their servers?", since they did not immediately grasp the idea that the Data Track allowed them to access this information. Once the test moderator explained that this was possible, only 3 participants succeeded at listing the information about them that Adbokis had stored on their servers. The reason for this poor result, besides the lack of users' mental models of transparency and control features on the services' side, was probably that the storage icon to be clicked to get online access to the service's side was not adequate, obvious and lacked visibility (see the database icon in Figure 10). Further redesigns need to address this issue, possibly considering alternative interaction paradigms and enhancing the learnability of the tool during first time use.



Figure 10. Example of a service provider in the bottom panel of the Data Track's trace view, including storage icon to be clicked for getting online access to one's data stored at the service provider.

Another important aspect to explore was the users' understanding of where their actual data were stored when using the Data Track program. The PrimeLife versions of the Data Track stored the logged data locally on the users' computer (Wästlund & Fischer-Hübner 2010). However, recent work has shown a privacy-friendly mechanism in which data could be stored remotely, for instance, at a cloud service, but still under the users' control (Pulls 2012). Answers from the usability evaluations showed that 8 out of the 14 participants understood that the data being displayed by the Data Track's trace view were stored either locally on their computer (6) or remotely stored (2), but under their control. The remaining 6 participants stated that these data were located only at the services that they have given them to, which is the wrong mental model. These ambivalent results indicate that work is still needed on helping users clearly differentiate what data are under their control and what is on the services' side.

To explore the mental models of participants regarding the possibility to delete data from the services side using the Data Track interface, participants were asked the question "What do you think happens when a piece of information is deleted from the Data Track trace view?" For this question, half of the participants (i.e. seven participants) stated that when deleting a piece of information from the top panel of the trace view such data get deleted only from the Data Track program, but not from the service's side. 2 participants stated that the information only gets deleted from the service's side, and 3 participants stated that deletion occurs in both places. This means that the interface successfully conveyed the idea to 10 participants about data being removed from the Data Track program, which implies that participants understood that there was a difference between their data located locally under their control, and remotely under the services' control. Similarly, participants were asked what would happen if one of the services was deleted from the Data Track, to which many responded that the service gets deleted from the Data Track program, but didn't state what happens to the information stored at that service. Only 3 participants mentioned that their data or account would be deleted from the service, or that a request would be sent to delete their data from the service.

Participants were also shown another way of visualizing data disclosures in a chronological order, which we called the "timeline view", shown on Figure 11. When asked for their preference between the *timeline* and the *trace view*, 61.5% were in favour of the trace view.

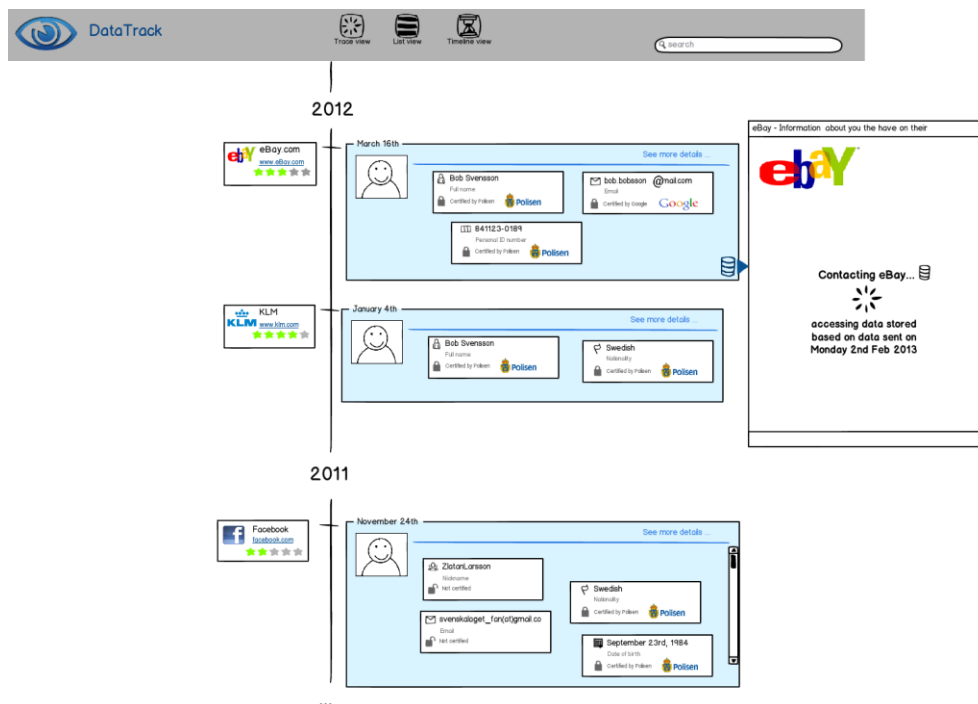


Figure 11. Data Track's timeline view of data disclosures.

5.3.4 Summary of results

Table 13. HCI requirements and design solutions obtained from evaluating the transparency tool Data Track

Req #	Observation (or Problem)	HCI Requirement	Proposed HCI principles and/or sample design solutions
R.6A	Visualizing data releases through a <i>trace</i> view was found useful, intuitive and informative. It seems to be preferred over a timeline view.	Users should have an intuitive and interactive way of visualizing previous disclosures of personal data.	Data releases could be visualized as a bipartite network, with one possibility having the user as a node in the centre and links branching on one side to the different services with whom he has had a relationship, and on the other side linking to the data items that have been released.

R.6B	Users have trouble understanding that they can access and correct their data on the services' side online with the Data Track tool.	<p>The interface should make it clearer that the possibility to review the data that has been disclosed is possible, and that it can be corrected or deleted</p> <p>Function to switch between users' side and services' side should be visible and obvious.</p>	The system could provide one view for visualizing data disclosures that are under the users' control and another view for visualizing data under the services' control. Making a clear animated transition between these views can help users understand that they are being connected to the services' side.
R.6C	The consequences of deleting a data item from the transparency tool could be made clearer and become more intuitive.	The interface should explain and warn about the consequences of deleting data items. It should be clear where data items are being removed from.	Alert with a dialog when information is about to be deleted. The dialog should contain an informative explanation and obvious icons representing what is being deleted and from where.
R.6D	The consequences of removing a service from a transparency tool are unclear	The interface should explain and warn about the consequences of removing a service. It should be clear where the piece of information is being removed from.	Alert with a dialog when information is about to be deleted. The dialog should contain an informative explanation and obvious icons representing what is being deleted and from where.
R.6E	In comparison to the PrimeLife Data Track UIs, improvements have been made about users' understanding of where their data are located, but further improvements can be done to make this difference more intuitive.	Users should be able to easily understand whether each data item is under their control or the services' control, and where it is stored.	Use colours to represent the different states of location of different pieces of information. Colour legends and mouse-over tooltips can provide more details explanations on demand.

R.6F	There is a difference between explicit and implicit collection of data.	Users should be made aware of implicit collection of data done by the service provider.	Make the look of the explicitly sent (i.e. information that user sent explicitly, for example during registration to a service) information different from the look of implicitly collected information or inferred information (i.e. information that the service provider collects without the user being fully aware of it, such as location, browser version, whether the customer is reliable, etc.).
-------------	---	---	--

5.3.5 Limitations and next steps

Testing the users' understanding that the services could collect and/or store more information about them than they explicitly release could not be tested on this occasion due to a problem with the prototype that was discovered at the moment of testing. However, this is an important question that will be investigated in future test iterations.

Although the usability evaluations have tested the understanding and user-friendliness of the user interface, the test was not designed to explore the realistic situations in which users might actually interact with the Data Track tool. Nevertheless, it will be important to know also the situations and motivations of users for looking into the data that they have previously released. Having a user-friendly interface does not mean that users would actually make use of such tool, and further evaluations with different methods are needed to test this aspect.

Additional future work on tools for transparency and visualizations of data releases include testing more realistic and more cloud-related scenarios, where users disclose many data items to many different service providers, who may in turn forward the data to (chains of) cloud service providers. This consideration forces some redesign for the Data Track user interface, where users should be able to navigate through various elements without the interface being cluttered. Figure 12 shows an example sketch of how a more realistic scenario for the Data Track could look like, depicting the flow of users' data through the chain of clouds.

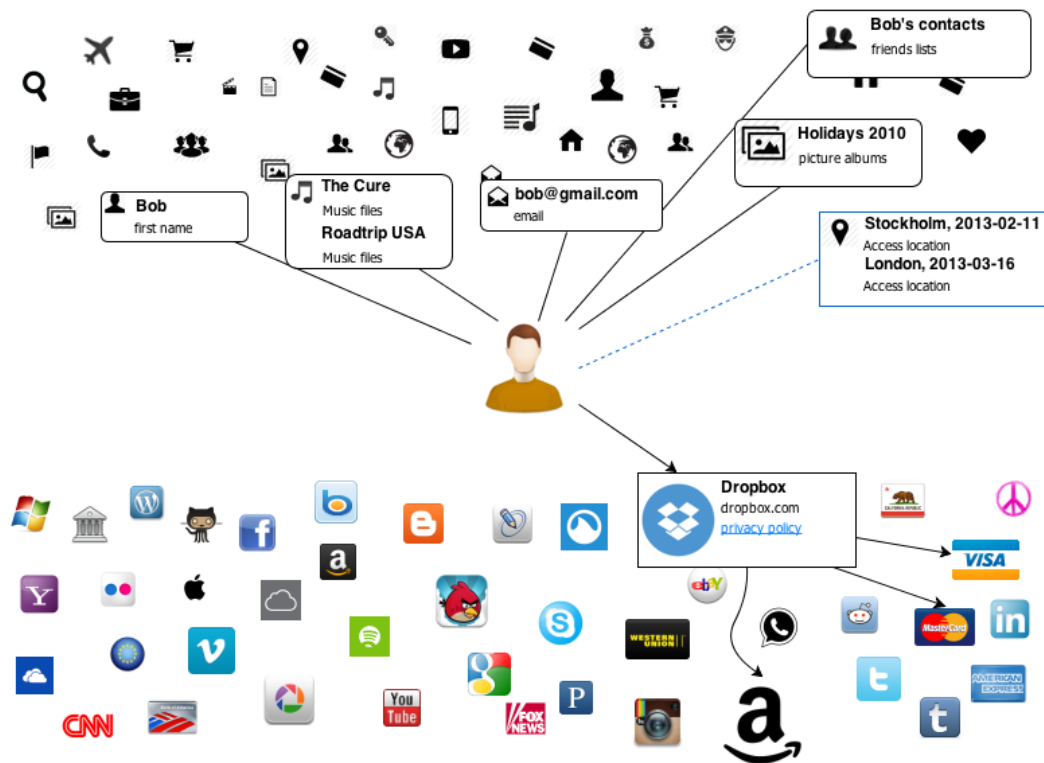


Figure 12. Data Track mock-up for illustrating chains of information flows

5.4 Usability and Security for Access Control Rule Sets

5.4.1 Background

Access control policies describe high level requirements for access control systems. Access control rule sets ideally translate these policies into a coherent and manageable collection of ALLOW/DENY rules. Designing rule sets that reflect desired policies is a difficult and time-consuming task. The result is that rule sets are difficult to understand and manage. To provide means for obtaining usable access control rule sets, i.e., rule sets that (a) reflect the access control policy and (b) are easy to understand and manage, we designed formal tools to handle access control policies more easily (Beckerle & Martucci 2013). The process to elicit HCI requirements for those tools and also their evaluation was carried out with three user studies, each with a different design and objective.

5.4.2 Experiment 1: Semi-structured interviews with system administrators for eliciting security and usability requirements.

A semi-structured interview design was carried out in a pilot study to find out what challenges do system administrators encounter when managing access control rule sets. Seven system administrators (IT support professionals) were individually interviewed. They worked in academia and industry and are from four different organizations (all located in Germany). All of them manage Linux- or Windows-based access control mechanisms. There were two objectives for these semi-structured interviews: to list the usability challenges regarding the management of access control rule sets and to look at how they handled those challenges. No financial incentive was offered to the participants in this user study, but we promised to inform them first-hand about our findings and conclusions.

The result of the semi-structured interviews is the identification of two general types of challenges related to the management of access control rule sets: those related to the accuracy of the rule sets

regarding their security requirements and those that result in poor manageability of the rule sets. The semi-structured interviews and their results are presented in (Beckerle & Martucci 2013) .

5.4.3 Experiment 2: Between subject design to collect data regarding the use of our support tools for producing access control rule sets

A between subject experiment was carried out in one of our user studies aiming to evaluate how helpful the security and usability metrics were for users when creating access control rule sets. 12 participants took part in this user study. Two-thirds of the participants were non-experts regarding access control configuration and management. The other 4 participants were IT support professionals (experts), who manage access control mechanisms on a regular basis. The age of the participants ranged between 20 and 25 ($\mu = 34.5$; $\sigma = 8.1$) and 4 of the participants were female. 7 of the participants were graduate students, 1 had a PhD degree, 3 held degrees from universities of applied sciences, and 1 had no university degree. All participants were volunteers and no financial incentive was offered to the participants in this user study.

The experiment was an exercise that consisted of a small file system populated by files and groups of users that should or should not have access to those files. The goal of the participants was to produce access control rule sets that reflected the desired access control policy. The participants were divided into two groups, a control group that was given the rule sets in the form of a “MS Excel spreadsheet” *without the support* of the sets, metrics and optimization criteria (this group was referred to as WOS), and another group who had a similar spreadsheet *with support* of our proposed sets, metrics and optimization criteria embedded in key cells of the spreadsheet. The rule sets obtained from the two groups of participants in experiment 2 were used as input to experiment 3, which is presented next.

5.4.4 Experiment 3: Expert opinion to rank the collected data according to their knowledge

We used the opinion from experts in the ranking of access control rule sets that were produced by participants of the second experiment (eighteen in total) according to the experts’ knowledge and experience. The evaluation criteria defined were: (a) how accurately the rule sets implement the access control policy and (b) how easily the rule sets can be understood and managed. Four system administrators took part in experiment 3, and they all ranked all eighteen rule sets. The system administrators were asked to rank the rule sets first according to criterion (a) and then according to criterion (b). They were also asked to provide a short description of their approach for evaluating the rule sets according to the defined criteria. No financial incentive was offered to the participants in this user study, but we promised to inform them first-hand about our findings and conclusions.

The evaluation of the results from experiments 2 and 3 showed that the participants in the group that had the support from our formal tools, sets and metrics performed significantly better than those on the group without support of our tools, sets and metrics ($t(3.629) = 7.621$, $p = 0.007$). The system administrators reported different approaches and methods used in their evaluation, but all produced similar rankings that showed a significant positive correlation between them and the rankings that were automatically produced (Beckerle & Martucci 2013).

5.4.5 Summary of results

The following table presents the summary of research questions and results from the experiments that were carried out in our studies on usability and security for access control rule sets.

Table 14. HCI principles and design solutions obtained from evaluating novel techniques for access control rules

Req #	Observation (or Problem)	HCI Requirement	Proposed HCI principles and/or sample design solutions
-------	--------------------------	-----------------	--

R.7A	It is very difficult for system administrators to verify the accuracy of access control rule sets regarding the access control policy. Thus rule sets need to be understandable and manageable to assist system administrators in their task.	Concise rule sets are better than large sets.	Tools, sets and metrics that can support users to evaluate and compare the security and usability properties of different rule sets.
		Redundant / contradicting rules are to be avoided.	
		Rule sets need to be designed to facilitate tasks for administrators.	

5.5 Mapping legal principles

Legal privacy principles influence Human Computer Interaction, and the PISA EU project showed earlier how privacy principles can be mapped into HCI principles and solutions (Patrick & Kenny 2003). In extension to the work by the PISA project, we will in Section 5.5.1 first discuss essential legal privacy principles for transparency and accountability for the cloud, for which we will then in Section 5.5.2 elicit HCI requirements and principles. Our legal analysis will mainly refer to the principles of the EU Data Protection Directive 95/46/EC, but we will also refer to other legal requirements deriving for instance from the opinions of the Article 29 Data Protection Working Party. In view of the ongoing review of the European legal framework on data protection, our analysis will also take into account legal principles that are being proposed in the draft EU General Data Protection Regulation (GDPR) (European Commission 2012).

5.5.1 Legal Principles

One important legal attribute of accountability is transparency. This section will put an emphasis on legal provisions for transparency and accountability for the cloud that have HCI implications and that thus need to be addressed by the user interface design. These legal provisions mainly comprise transparency rights and detective and corrective control rights that data subjects have in regard to data controllers¹⁴. The proposed EU regulation also highlights the importance of usable transparency and user control by requiring that data controllers have *“transparent and easily accessible policies with the regard to processing of personal data and for the exercise of data subjects’ rights”* (Art. 11 draft GDPR).

5.5.1.1 Information Rights (ex ante Transparency) and Consent

Ex ante transparency is a condition for data subjects of being in control and for rendering a consent¹⁵, which has to be informed, valid. Article 10 of the Data Protection Directive defines what information relating to the processing of their personal data needs to be given to data subjects, when information about them are collected and processed. This includes at least the identity of the data controller, and the data processing purposes. Moreover, further information needs to be given for example on the recipients or categories of recipients of the data, on whether replies to questions are obligatory or voluntary and on information about the individual's rights in so far as such further information is necessary to guarantee fair data processing. Such information has to be provided to the data subjects not only when the information is collected from the data subjects, but also when the data have not been obtained from them (Art. 11 Data Protection Directive).

¹⁴ According to EU Directive 95/46/EC, a data controller is defined as the entity that alone or jointly with others determines the purposes and means of personal data processing.

¹⁵ 'The data subject's consent' is defined by the Data Protection Directive as “any freely given specific and informed indication of his wishes by which the data subject signifies his agreement to personal data relating to him being processed”.

The processing of personal data has to be based on one of the grounds that are mentioned in Art. 7 of the Data Protection Directive. The consent of the data subject (Art 7(a) Data Protection Directive) can be taken as a legitimisation for personal data processing in the cloud. Information that needs to be given to data subjects for a valid (informed) consent should cover at least the elements of information required by Art. 10 Data Protection Directive.

The draft GDPR is in Art.14 extending the information that at least should be provided to data subjects by data retention periods, the right to lodge a complaint to the supervisory authority, and – what is especially of relevance in the Cloud context - information about the data protection level of a third country or international organisation to which the data controller intends to transfer data.

Recently, the Art.29 Working Party discussed in their Opinion 5/2012 on Cloud Computing (Art. 29 Data Protection Working Party 2012) a lack of transparency in regard to the Cloud Services' processing operations. Privacy threats may arise from the controller not knowing or not informing the data subjects about the:

- Chain processing that involves multiple processors & subcontractors
- Data being processed in different geographic locations within the EEA
- Data being transferred to third countries outside the EEA
- Disclosure requests by law enforcement

The last of the aforementioned threats is also important for the reason that even if data are processed at a services side located in the EEA, data transfers to the US may take place upon requests by US American law enforcement services.

Furthermore, increased transparency over the chain of data processors and subcontractors is important as in practice entities it cannot always clearly be assigned the roles of data controllers or processors. Art. 29 Working Party is in its opinion 1/2010 on the concepts of “controller” and “processor” arguing that these roles should therefore be determined by “factual elements and circumstances” (Art. 29 Data Protection Working Party 2010). Also the proposed EU data protection regulation recognises that data processors may under certain circumstances have increased control over the data processing and should be made directly accountable to the data subjects (cf. Art. 24, 26 IV).

Our stakeholder group workshop (see 5.1.1) revealed another transparency problem, namely that data subjects are often not well informed about the applicable consumer laws and rights, especially if cloud brokers or mediators are involved in cross-border eCommerce transactions.

Hence, in a cloud setting, it may therefore be argued that more policy information beyond the minimum that is required by Art. 10 of the Data Protection Directive should be displayed to the data subjects, including:

- Contacts & obligations of all data processors along the cloud chain (as far as data processors can be determined ex ante);
- Geographic locations of all data centres along the cloud chain and, in case that they are located outside the EEA, information about their data protection levels;
- How disclosure requests by law enforcement agencies are handled;
- Consumer rights and applicable laws.

5.5.1.2 Right of access (ex post Transparency) and other Data Subjects Rights

The EU Data Protection Directive provides data subjects with the right of access to their data, which comprises the right to information about the data being processed, data processing purposes, data recipients or categories of recipients, as well as information about the logic involved on any automatic

processing (Art. 12 (a)). This data subject right providing ex post transparency is a prerequisite for exercising the data subject rights to correct, delete or block data that are processed not in compliance with the Directive (Art. 12 (b)).

The proposed EU Data Protection Regulation is with its Art. 15 extending the information to be provided by the controller to include also information about data retention period, the right to lodge a complaint with the supervisory authority and “*the significance and envisaged consequences*” of the data processing at least in the cases of profiling. The data subjects shall also have the rights to obtain this information electronically if they have made their requests in electronic form. Besides, the proposed regulations extends the data subjects’ rights by the right to be forgotten (Art. 17) and the right to data portability (Art. 18) and introduces the obligation of data breach notification of the controller to the supervisory authority (Art. 31) and data subject (Art. 32).

Further more specific ex post transparency rights are for instance provided by the Swedish Data Patient Act (Svensk Författningssamling) to data subjects by requiring that health care providers have to inform patients upon request about who has accessed their medical information.

5.5.2 HCI Requirements, Principles and Design Proposals

As pointed out in (Patrick et al. 2003), legal privacy principles for transparency, consent and data subjects’ rights “have HCI implications as they describe mental processes and behaviour that the data subjects must experience in order for a service to adhere to these principles”. In particular, the principles require that data subjects *comprehend* the transparency and control options, are *aware* of when they can be used, *are able* to use them. As argued in (Patrick et al. 2003), HCI requirements mapping legal privacy principles can therefore be grouped into the 4 categories (1) comprehension, (2) consciousness, (3) control and (4) consent. In this section, we are following this grouping of HCI requirements and for discussing HCI methods for meeting these requirements. HCI methods proposed are partly presented in more detail in the PrimeLife HCI Pattern report (PrimeLife WP4.1 2010).

Comprehension: The category comprehension comprises categories that allow a user to understand the transparency and control options discussed above. For supporting the user’s understanding, HCI methods from cognitive psychology can be exploited that try to either evoke *appropriate mental models* or analyse whether already existing models can be accounted for. Hence, HCI principles that were elicited as described in Section 5.2 by analysing the users’ mental models of transparency and control functions can be helpful for making transparency and user control options well understandable.

Furthermore, user interfaces, which use real-world metaphors, e.g. in form of suitable icons, are easier to learn and understand (following Jakob Nielsen’s usability heuristics of a “match between system and the real world”). *Privacy policy icons* have been researched and developed for visualising policy elements in stated privacy policies with the objective of making the content of legal policy statements easier to access and comprehend. Policy icons should preferably be standardised in future and usable across different cultures.

Within the scope of the PrimeLife EU project, a set of policy icons addressing the legal transparency requirements of the EU Data Protection Directive has been developed. These icons can be used to illustrate core privacy policy statements, namely statements about what types of data are collected/processed, for what purposes, and what the processing steps are (Holtz et al. 2011). An intercultural comparison test of the policy icons conducted at Karlstad University with Swedish and Chinese students as test participants gave insights into which icons seem to be well understood by both cultures and which were understood differently by persons with different cultural backgrounds (Fischer-Hübner & Zwingelberg 2010). Icons easily understood by both Swedish and Chinese students were, for instance, the ones shown in Figure 13, displaying types of data (personal data, medical data, payment data), the purpose “shipping” and the processing steps (storage, retention).

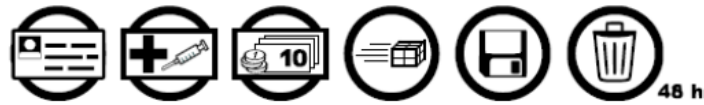


Figure 13. Example of well understood PrimeLife policy icons

Other Creative Common-like privacy icons have for instance been initiated by Aza Raskin in 2010 (Raskin 2010) and further developed by a Mozilla-led working group (who however stopped their work more than a year ago). Interestingly, it includes special icons informing end user about how easily services sides are cooperating with requests by law enforcement (see figure 14 for examples of the alpha release of icons). As already pointed out above and as it also became apparent after the revelation of the PRISM program, this is an important aspect that is often not transparent to cloud users.

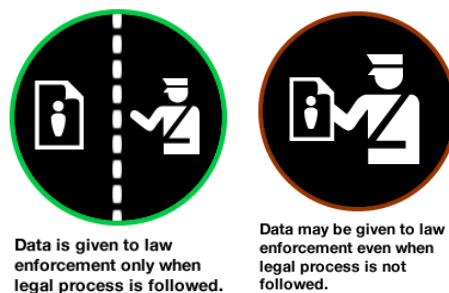


Figure 14. Icon proposals (alpha version) by Aza Raskin informing about how disclosure requests by law enforcement are handled (Raskin 2010)

For meeting the demand of higher transparency for data processing in the cloud, further policy icons can be helpful for informing about geographic locations of all data centres along the cloud chain, and in particular whether they are placed in the EEA, and, in case that they are located outside the EEA, information about their data protection levels.

Complex privacy notices are usually neither read and nor easily understood. This is also due to limited cognitive capacity that people usually have, such as limited attention spans memory, as well as a restricted ability to process a large amount of complex information at one time (Patrick & Kenny 2003). Comprehension of policy information can also be facilitated by a multi-layered structure of policy notices, as it was recommended by the Art. 29 Data Protection Working Party in their opinion on “More Harmonised Information Provisions”(European Commission 1995). This recommendation takes the approach to structure complex policies into different layers, where the top layer is only providing a short privacy notice with the policy information that is at least required by Art. 10 EU Data Protection Directive (i.e., at least the identity of the controller and data processing purposes) and further detailed policy information can be obtained from the condensed and full privacy notices on other layers (cf. Section 2). Each layer should offer to the data subjects the information needed to understand their position and make decisions.

However, as discussed above, if data are processed in the cloud, it may be argued that more policy information beyond of what is required by Art. 10 should be displayed to the data subjects for providing transparency, and depending on the circumstances. Such information listed at the end of Section 5.5.1.1, may also have to be displayed on top layer for allowing users to comprehend the implications.

Consciousness: The HCI requirement of consciousness requires that users are aware, or are paying attention, to the transparency and control options.

One common HCI technique for creating awareness is based on pop-up windows. However, privacy and security are usually only secondary tasks for users, who therefore tend to click away obtrusive messages. Another example for a more unobtrusive HCI method for reminding users is based on the arrangement of the interface. In particular, if a transparency or control function is available in a certain context, specific icons or messages can be placed nearby in the interface layout for ensuring that users are aware of these options (cf. Patrick & Kenny (2003)).

Control: The category of control refers to the ability of the users to actually carry out the transparency and control functions after they are aware of them and comprehend what to do.

Appropriate user actions can be supported by basing the user interface on real-world analogies and by HCI techniques that make these functions “obvious” (Patrick & Kenny 2003).

Consent: Informed Consent is both a legal and an HCI requirement, because it refers directly to mental processes and user behaviour that needs to be supported, as for providing informed consent, users must *agree* to privacy policies or terms and conditions and must fully *comprehend* what they agree to and what consequences it may have. Informed consent is thus related to comprehension and consciousness.

In practice, consent is often obtained by click-through agreements. However, a click-through agreement based on a long legal statements on one page, may formally fulfil the information requirements of the EU Data Protection Directive, but will not live up to its intention. Therefore, the proposed EU Data Protection Regulation also requires that privacy policies should be *transparent and easily accessible*. Hence, click-through agreements should at least be based on policies structured in multiple layers with click boxes appearing on the top layer displaying a short privacy notice.

Nevertheless, click-throughs still have the problem that users are habituated to easily click through without having read the text. Within the scope of the PRIME project, we have earlier introduced the HCI concept of “Drag and Drop Agreements (DaDAs)” (Pettersson et al. 2005), by which users express consent by moving icons graphically representing their data items to graphical representations of data receivers. By this, users are forced to make better informed decisions about what data they are releasing to whom. With the emerging use of touch screen devices, DaDAs may more conveniently and broadly be deployed as an HCI concept in future.

5.5.3 Summary of results

The following table extends and adapts the table in (Patrick & Kenny 2003) and summarises how legal principles for transparency and accountability for the cloud discussed in Section 5.5.1 can be mapped to HCI requirements and proposed HCI solutions as discussed in Section 5.5.2.

Table 15. Mapping Legal Privacy Principles to HCI requirements and proposed solutions

Req #	Legal principle	HCI requirements	Proposed HCI principles and/or sample design solutions
----------	-----------------	------------------	--

R.8A	Data subjects have the right to be informed at least about the controller's identity, purposes and possibly more details listed in Art. 10 EU Directive and should also be informed about any further information needed for making data processing in the cloud transparent.	The data subjects <i>know</i> at least who is the controller of their data, for what purposes the data are obtained plus other details (e.g., contacts and geographic locations of data centres along the cloud chain, applicable laws, how requests by law enforcement are handled, etc.), so that they can <i>understand</i> the implications.	Policy information is <ul style="list-style-type: none"> provided in a way that accounts for the users' mental models; structured in multiple layers following the Art. 25 WP recommendation; complemented with suitable policy icons.
R.8B	Personal data processing in the cloud can be legitimised by the data subject's unambiguously given consent pursuant Art. 7 (a) EU Directive	Users give <i>informed consent</i> and are <i>understanding</i> the implications	Consent is obtained by click-through agreements associated to short privacy notices (top layer notices of multiple-layered policies), or via DaDAs (Drag and Drop Agreements).
R.8C	Data subjects have the right to access their data pursuant Art. 12 EU Data Protection Directive. Data subjects may have further rights in regard to the processing of their data according to more specific laws, e.g. in Sweden a data subject has the right to information on who have accessed the data subject's data according to the Swedish Patient Act.	Data subjects are <i>conscious</i> of their ex post transparency rights, <i>understand</i> and <i>can exercise</i> their rights	Ex ante Transparency functions are displayed prominently and obvious to operate; Transparency functions are based on a suitable metaphor and/or account for the user's mental models; Transparency functions are made available at the right time/ in the right context, e.g. tracking logs display prominent online functions to exercise the right to access.;
R.8D	Data subjects have the right to correct, delete or block their data pursuant to Art. 12 (b) EU Data Protection Directive. Further rights, such as the right to be forgotten or the right to data portability, are currently proposed.	Data subjects are <i>conscious</i> of their control rights, <i>understand</i> and <i>can exercise</i> their rights.	Functions for exercising data subject rights are displayed prominently and obvious to operate; Transparency functions are based on a suitable metaphor and/or account for the user's mental models; Transparency functions are made available at the right time/context, e.g. at the time when users are accessing their data locally or online;

5.6 Mapping social trust factors

5.6.1 Literature review

The ISO/IEC 25010:2011 definition of *trust* runs: “Degree to which a user or other stakeholder has confidence that a product or system will behave as intended.” From this definition it may appear as if people’s notions of trustworthiness in an electronically delivered service were simply an issue of how well they trust an automaton. However, there is ample evidence that trust stems from previous encounters with the service provider (possibly in non-electronic form), from the general reputation of the brand of the service provider, as well as from statements made by friends, rather than any direct understanding of the privacy and security reliability of the service in question.

This complex picture has been drawn in some works including some previous EU-funded FP projects such as the PRIME project, see Andersson et al. (2005). Furthermore, several trust issues that individual Internet users may have with Internet services were revealed in that project: the Internet is believed to be **intrinsically insecure** (cf. also the cloud study by Ion & al.(2011)), **anonymity options are unknown**,¹⁶ and some users have a real difficulty in differentiating between the data stored in their own computer and the one controlled by the Internet services.

The Trustguide final report by Lacohee, Crane, and Phippen (2006) presents six *Trust Guidelines* which were derived from iterated focus groups in a project run by British Telecom and Hewlett Packard Labs. The issues raised do not only relate to trust in cloud services but also to some aspects of accountability. A very brief and condensed listing of the Trust Guidelines (TG) is given here (abstracted from the Executive Summary, but cf. Ch. 15, *ibid.*):

TG.1 “The fundamental foundation of the guidelines lies in education. [...] Currently education is sparse and disconnected, resulting in ill-founded beliefs hampering engagement.”

TG.2 “people will develop trust in a service through experimentation in a ‘safe’ environment prior to engaging in a potentially risky transaction.”

TG.3-4 “**Restitution Measures** [...] Citizens believe there is no such thing as a secure service and claiming so leads to mistrust. [...] clearly state the [restitution] measures”

TG.5 “**transparency** brings increased confidence”

TG.6 “trust is not built through **unsubstantiated claims** of security and protection”

Other observations from the Trustguide are also worthy of consideration: “In looking at building and monitoring trust relations we found that recommendations from trusted sources within an individual’s social network play a major role” (*ibid.* Lacohee 2006, p 19). This actually reflects a much more general characteristic than only the “cyber trust” that the Trustguide addressed. For instance, studying Ugandan farmers, Wamala (2010) found that technology inexperienced persons do not trust any “new” information source, whether the information comes through the Internet and web, or via SMS or even radio. This has been called “social proof” by Cialdini.¹⁷

¹⁶ Bold face are used here to highlight trust issues that have been particular focused when scanning the literature; most of these issues have resulted in the formulation of requirements in the table summarising this section.

¹⁷ How to understand the role institutions and institutional evolution play in privacy and security issues of cloud computing is further elaborated in a recent paper by Kshetri (2012) in which comparisons are made with more established industries. This article presents interesting perspectives on social and institutional factors but has no direct bearings on HCI guidelines in contrast to other papers reviewed in this section.

Thus, the discussion so far indicates that good **experience** with a service provider gives trust whether the experience is one's own or some friend's (or independent third parties; Turner et al., 2001). O'Neill (2002, p 76) points out that "*Well-placed trust grows out of **active enquiry** rather than blind acceptance*". However, do people seek the best information sources? Individual cloud users often replace inquiry by brand name reliance: "participants transferred trust in the company itself to trust in security of their cloud" report Marshall & Tang (2012) (cf. (Turner et al. 2001; Cloud Industry Forum 2011)). Trust seals and similar symbols have for similar reasons a very ambiguous value; if people blindly trust a logo on a web site, not much is gained by promoting such schemas. (Cf. e.g. discussion by Tsai et al., 2011) Marshall and Tang (2012) furthermore detected an unfounded belief that **cost of service would imply a higher trustworthiness**. One could of course argue that paying for a service should really give its user something more than what services free of charges do, but Marshall's and Tang's interview subjects could not prove their assumption that cost implies trustworthiness.

This naïve trust in cloud services should also be tested for business end users from the private sector: in the stakeholder workshops reported in Section 5.1 a representative from a data protection authority voiced concerns over the limited interest that companies pay to privacy issues when they have finally found the business value of adopting cloud services. We could call this a "**business first attitude**" and it could be worth to investigate if business people have the same naïve trust in cloud providers as individual end users have or how much this attitude is a result of pure neglect and disrespect of privacy laws.

To continue, even if many users are not able to make a complete risk assessment or may act in contradiction to their stated awareness of risks, their actions can often not be called irrational. Rather, as pointed out already in the Trustguide people find means to mitigate risks when they are unsure of the trustworthiness: "A commonly used technique to reduce the potential effects of credit card fraud was to have a credit card dedicated to Internet use or to use a different account" (Lacohée et al. 2006, p 20); "in every case of using potentially untrustworthy sites attendees cited the fact that their credit card company would foot the bill if something went wrong" (Lacohée et al. 2006, p 21). Another example is given by Bødker et al.: "Hans's trust in his friends, who ordered the original tickets, and the urgency of the situation caused him to ignore many aspects that he would otherwise be concerned with" (Bødker et al. 2012, p. 55). The Hans story had a happy ending. Conclusion: good prices are often worth the price of uncertainty. One might call this a **situation-dependent risk-taking** which includes a proportional risk assessment. Perhaps also the "selling" of information discussed in Section 5.2.2 of this chapter should be understood in a similar light – "the on-going work on the economics of privacy that stresses the costs and benefits to an Internet user of a specific situation requiring disclosure, rather than their disposition" (Joison et al. 2010, p18).

Naturally, this situational trust calculation is also of interest when it comes to how trust may influence people's adoption of ideas of transparency and accountability. These ideas are not dependent on a specific offer from a specific service provider at a specific time. One may thus think that cultural characteristics may play a greater role for such functions. Ion et al. (2011), as referred in Section 5.2.1, show how two different groups – Swiss and Indians – with almost opposite privacy concerns in some respects – Swiss distrust government surveillance; Indians like it because "national security comes first" – both have a tendency to distrust cloud storages and prefer physical backups at home, and at the same time, when it comes to cloud services, believe they have more rights and guarantees than what they actually have. But despite similarities, tendencies to different expectation was found and the authors suggest "changing the content and the presentation of privacy policies and Terms of Service agreements" to include "accounting for internationalization. The latter involves going beyond just translating the service interface and privacy policy. Companies should keep in mind that users from different countries may have **different privacy expectations and understanding** of privacy guarantees offered by the cloud storage system." (Ion et al. 2011). Thus unfounded trust must be met with a culturally sensitised mindset when evaluating terms and conditions.

Marshall and Tang conclude that “users’ uncertainty and misconceptions” affected their trust in cloud computing services (Marshall & Tang 2012) . “For almost all participants in this study, the primary consideration in deciding what digital identifier should be associated with information in the cloud was the **perceived longevity** of the digital identifier—not the identifier associated with or known to the most relevant collaborators. Participants wanted to ensure that they would always have access to their information” say Volda, Olson & Olson (2013) and give several examples from their study. This lack of trust in the longevity of some identifiers (such as work-related email addresses) results in that “information that would typically be associated with one digital identifier was managed under an account associated with another digital identifier with greater perceived longevity, blurring the distinctions between facets of identity” (Volda & Olson, Judith S Olson & Gary M 2013). Transparency and restitution controls that promise anonymity and build on pseudonymity can thus appear unreliable.

There is less literature on companies’ perspective on trustworthiness of cloud providers. Pearson (2013 , p 31) summarises surveys among CIOs showing that concerns and barriers include worries about **cloud security, performance, and availability** but also vendor lock-in. Worries about **vendor lock-in** is a distrust in the cloud concept but as it is not directly related to privacy or transparency of personal data processing we do not consider it as a requirement for our purposes (the “data portability” discussed in Art. 18 in the proposed EU Data Protection Regulation concerns data subjects rights; see Section 5.5.1.2 above).

The public sector has started to use cloud services (Wyld 2010). A recent article illustrates the concerns: “the empirical findings here demonstrate that **perceived availability, access, security, and reliability** would be key variables of cloud computing acceptance in public sectors since they were found to be influential in predicting the behavioural intention to use cloud technologies” reports Shin (2013 p 200) from a study on Korean public service workers’ views. A comparison with people from the private sector was included in this study, showing the trust problems to be similar between public and private sectors (cf. Pearson above) even if the emphasis on individual variables might differ somewhat. Shin ends his article saying that “The difference of [the Technology Acceptance Model] between private and public sectors may be an interesting topic for future cloud studies” (Shin 2013, p 202). The fears of breaking laws of data protection do not figure prominently in different studies on professional groups’ concerns (except for health care cases); it may of course be a factor behind the worries about security but may also differ between private and public sectors.

In addition to the voiced and investigated parameters just mentioned, one might wonder if not also a very explicit model of cloud service chains would be needed for the business end users (which we take to include public sector end users) to be able to understand assessments of these parameters. It could be worth to compare these findings about business end users with the conclusion of Marshall & Tang (2012) about file sync and sharing: “users’ uncertainty and misconceptions limited their ability to fully take advantage of the service’s features. Users needed more **accurate and robust models** to be able to discover and trust cloud computing services.” The authors also note that users’ “perceptions of privacy and security” are independent of users’ understanding of how cloud services work – these are thus independent dimensions and this might of course be true also for business end users. Nevertheless, if officers are given an understanding of reliable assessment measures of cloud services, this should make it possible to bring down barriers to cloud adoption. As previous sections of this chapter have shown, it can be a tricky question how to give users such understanding. For instance, the workshop with PET-PhD students, described in Section 5.1.3, displayed the **distrust in service chain** (cf. R.2C & G): in a transaction with one service provider, this service provider can send a confirmation email, but the email provider can be untrustworthy. Thus, the fear here is that a service provider may be evaluated as trustworthy, but behave riskily if his customers and clients are naïve enough to provide email addresses from an untrustworthy provider. A good presentation of an A4Cloud supported and assessed cloud service chain should be able to reassess the chain if an (business) end user adds specific circumstances to the (interactive) presentation.

There is in particular one point where different sources seem to contradict each other. The Trustguide talks about **unsustainable claims** of ICT mediated services: “trust is not built through unsubstantiated

claims of security and protection. Being clear about the benefits and issues related to a service will engender far greater trust” (Lacohée et al. 2006, p 2). This is the 6th Trust Guideline, “TG.6: Openness – honesty signifies and engenders trust” (Lacohée et al. 2006 p 85). “We have observed many instances where people have not engaged with a service simply because they did not like the terms and conditions because they did not inform them effectively in their risk assessment process.” The authors conclude that “it is crucial that the service provider is honest and they do not make unsubstantiated claims of security, and do not provide guarantees that are untenable. Today’s user is an informed, cynical individual; they cannot be bought with empty gestures” (ibid.). Joinson et al., on the other hand, found in their study that “A strong privacy statement, despite the presence of cues to lack of trustworthiness, increased participants’ reported trust” (Joinson et al. 2010, p 16; see also DMA¹⁸, and the discussion on perceived control in Section 5.2.1).

For the future one might investigate the hypothesis that more specific claims (among the unsubstantiated claims) instils more trust than general bragging if we take “general bragging” to also include complex privacy statements, as people regard them as being deliberately complex to obfuscate the terms and conditions according to the Trustguide (Lacohée et al. 2006, p 85); thus, they are not examples of specific claims because such claims need to be succinct enough for readers to grasp the specific claims being made.

Before presenting the summary of results, this section ends by restating some suggestions for further investigations which have been mentioned earlier:

- Professional users of cloud services: what is the difference between private and public sector officers as concerns trust in cloud services?
- Unsubstantiated claims do or do not build trust. Hypothesis: specific claims are more important than general bragging. (Still, A4Cloud user interfaces must signal the correct degree of trustworthiness.)
- Cost implies trust (individuals). Does this holds also for business end users? Or does the business first attitude make them immune for simple unsubstantiated trust signals?

5.6.2 Summary of trust factors:

Table 16. HCI requirements and design ideas obtained from literature review on trustworthy factors

Req #	Trust issue or factor	HCI requirements	Proposed HCI principles and / or example design solutions
R.9A	Well placed trust grows out of active enquiry <i>O’Neill (2002), Wamala (2007), Trustguide TG2;(Lacohée et al. 2006).</i>	Users should be able to pursue experimentation and enquiring. Users should be guided beyond enquiring only friends and relatives.	Safe environments for experimentations and enquiries (the environments must not oversimplify the complex cloud service ecology). Make it possible to enquire good sources

¹⁸ Direct Marketing Association (DMA) <http://www.dma.org.uk>

R.9B	<p>Transfer of trust: trust in the company itself is often transferred to trust in the security of their cloud services</p> <p><i>DMA (see footnote 18); not necessarily from experience: Marshall & Tang (2012) (Marshall & Tang 2012) .</i></p>	<p>Users should be clear about the difference between service performance and privacy performance.</p>	<p>Make evaluation results concerning trustworthiness prominent.</p>
R.9C	<p>There is an unfounded belief among individual end users that cost implies higher trustworthiness</p> <p><i>Marshall & Tang (2012) (Marshall & Tang 2012) .</i></p>	<p>User should be able to balance their impressions gained from pricing with other relevant information about trustworthiness.</p>	<p>Make evaluation results concerning trustworthiness as prominent as cloud providers' cost schemes.</p>
R.9D	<p>A situation-dependent risk-taking which includes a proportional risk assessment must be preferred over exaggerated risk-avoidance</p> <p><i>Bødker et al. (2012), Trustguide, (2006, p 20) (Lacohée et al. 2006), Angulo et al. (2012).</i></p>	<p>Users should not be frightened away from unattested sites if stakes are low (good prices are often worth the price of uncertainty).</p>	<p>Define a threshold for the desired risk levels and present information to users depending on this risk threshold. Promote the workflow of the application and the users' tasks if the risk lies below the defined threshold.</p>
R.9E	<p>Internet is intrinsically insecure: <i>Marshall & Tang (2012)(Marshall & Tang 2012), Ion et al. (2011), PRIME (2005),(Lacohée et al. 2006). Trustguide TG1 speaks of the necessity of taking measures also outside the user interface. This does not directly translate into HCI requirements, but the UI should relate to it.</i></p>	<p>"Users needed more accurate and robust models to be able to discover and trust cloud computing services." Marshall & Tang (2012)</p>	<p>In the user interface: users should be directed to sources they would normally rely on.</p>

R.9F	<p>“...perceived availability, access, security, and reliability would be key variables of cloud computing acceptance in public sectors since they were found to be influential in predicting the behavioural intention to use cloud technologies” Shin (2013, p 200)(Shin 2013).</p> <p>A business first attitude in cloud adoption where economic considerations far outweigh privacy concerns.</p>	<p>Business end users need to be correctly informed about cloud security, performance, and availability for individual cloud services they consider.</p> <p>This requirement holds for private sector (Pearson 2013) and public sector (Shin 2013) alike. For private sector this requirements also meets the problem of the business first attitude if accountability measurements are included in the information so that such aspects can easily be included in the decision process.</p>	<p>Display trustworthiness by evaluation results.</p> <p>Use graphical models of the A4Cloud model and match with graphical model of current chain.</p>
R.9G	<p>“Users from different countries may have different privacy expectations and understanding of privacy guarantees offered by the cloud storage system”</p> <p>Ion et al. (2011)</p>	<p>Internationalisation “involves going beyond just translating the service interface and privacy policy” (Ion et al. 2011).</p>	<p>When seeking customers outside EEA, seek expertise to cover different populations’ expectations.</p>
R.9H	<p>Restitution measures have positive trust effects</p> <p>Trustguide TG3-4 (Lacohée et al. 2006)</p>	<p>Clearly mark the possibility and ways of redress.</p>	<p>Users’ interfaces for transparency tools, such as the Data Track, could mark restitution measures.</p>
R.9I	<p>Transparency “brings increased confidence”:</p> <p>Trustguide TG5 (Lacohée et al. 2006); DMA (footnote 18)</p>	<p>Users should know when and where trustworthy transparency information is to be found.</p>	<p>Couple transparency with options for users’ actions.</p>
R.9J	<p>Anonymity option unknown:</p> <p>Unawareness of options for identity management has negative effects on trust in privacy-enhancing technology</p>	<p>Users must be able to understand the extent to which they can act under pseudonyms and that such identification schemas can provide access to transparency information.</p>	<p>Within the user interface demonstrate how anonymity options work.</p>
R.9K	<p>Perceived lack of longevity of identifiers make users blur partial identities:</p> <p>Preference for long-lasting identifiers (such as personal email addresses) Volda, Olson & Olson (2013)</p>	<p>Users must trust that they can manage in a life-long way the information associated with different identities (implications for transparency and restitution controls).</p>	<p>(This solution might result in conflicting interests using personal identifiers versus <i>appropriate</i> identifiers).</p>

R.9L	Unsubstantiated claims do not build trust Trustguide TG6 (Lacohée et al. 2006): <i>This issue concerns a long term perspective; one company's misconduct can affect a whole sector.</i>	Users must be able to put the right scope to their distrust.	Make privacy and security statements short and very clear.
R.9M	Unsubstantiated claims build trust Joinson et al. (2010): The problem here is that well-articulated privacy assurances make many individual end users trust a service competence and intentions.	As users do not check privacy statements etc., users must be made aware of trustworthy assessments of trustworthiness.	Make evaluation results concerning trustworthiness as prominent as cloud providers' privacy and security claims.

5.7 Concluding words

This chapter has demonstrated different research methods to derive HCI requirements and principles that will help to address the research challenges and questions that were outlined in chapter 3. In the next chapter, we will show which HCI requirements and principles will be relevant for the development of which A4Cloud tool user interfaces. In addition, a first set of high-level HCI guidelines will be derived.

6 Preliminary HCI Principles and Guidelines

6.1 Mapping HCI requirements to functional categories of A4Cloud tools

In order to propose a concise set of HCI principles and guidelines for A4Cloud tools, we group the HCI requirements and related HCI principles obtained from the different activities presented in Chapter 5 into general categories related to required functionality of possible accountable and transparent tools to be developed in A4Cloud, as also described in the use case descriptions by WP-B3. These comprise functionalities of A4Cloud tools for:

- Ex ante transparency (policy display incl. policy mismatches, mediating of trustworthiness or risks to individual end users);
- Exercising data subject rights;
- Obtaining consent;
- Privacy preference management (helping individual end users to manage their privacy preferences)
- Privacy policy management (for business end users)
- Ex post transparency (incl. display of policy violations and help with risk mitigation)
- Audit configuration (help with settings in regard to collection of evidences)
- Access control management
- Privacy risk assessment (for business end users)

Table 17 maps the obtained HCI requirements and related HCI principles into these functional categories of A4Cloud tools. This table has the objective to show the developers for each A4Cloud tool more clearly what HCI requirements need to be met and what HCI principles should be followed during the UI design.

Table 17. Mapping HCI requirements and principles to functional categories of A4Cloud tools

Functional Categories	General HCI requirements	For matching HCI principles refer to
1. Ex ante transparency	Make explicit data disclosures and implicit data collections transparent	R.2A, R.2E, R.2F, R.6F, R.9H
	Make data sharing and data processing along the cloud chain transparent, and provide the means to verify it	R.1E, R.2B, R.9A
	Provide indicators for the trustworthiness of nodes along the cloud chain	R1.G, R.2C
	Policies need to make the possible <i>consequences</i> of data disclosures in different recurrent situations transparent	R.3A, R.3B
	Make explicit that a service is a cloud-based service and what this implies in terms of privacy/security for the intended user	R.3C, R.9B

	Provide easily comprehensible policies informing data subjects at least about the identity of the controller, other responsible parties, for what purposes the data will be used plus other details needed, so that they can understand the implications.	R.1D, R.1E, R.1L, R.5B, R.8A
	Make trust-enhancing indicators intuitive, consistent and believable, as well as be appealing for the appropriate user group	R.1C, R.5C, R.5D, R.9E, R.9F
	Users should be able to know the approach and consequences when deciding to end the service	R.1K, R.9K
	Users should be aware of the extent to which they can act under pseudonyms	R.9J
	Inform users about the termination of their contract in a clear and straight-forward manner	R.1K
	Make reasonable claims about the privacy and security policies and technical capabilities of the service to promote trust	R.9L
2. Exercising data subject rights	Make users aware of their data subject rights, and support them to exercise their rights; in particular, make control options that are relevant in certain situations more obvious at those particular situations	R.1D, R.2H, R.8A, R.8C, R.8D
	Provide clear statements of what rights apply to individual users considering different factors, such as the users' culture or location and applicable legal regime	R.1F, R.8A, R.9G
3. Obtaining consent	Make users aware of pros and cons of their possible choices in an unbiased manner	R.4A, R.9M
	Obtain users' informed consent by helping and motivating them to understand policies and service agreements, so that they understand the implications	R.1B, R.1H, R.8B
4. Privacy preference management	UIs for preference settings need to make consequences in different recurrent situations & risks and benefits of disclosure transparent.	R.2C, R.9J
	Make users aware of pros and cons of choices in a comprehensible and unbiased manner	R.4A
	Offer appropriate default settings and choices that are privacy-friendly and reflect the users preferred options	R.1N, R.1J

	Let users do settings at the moment when it is relevant (on-the-fly management of privacy settings)	R.5A
	Explain consequences not in technical terms, but in practical terms ("speak the user's language")	R.1I
5. Privacy policy management	Make it possible for business end users to negotiate what is negotiable, and make negotiation clear and simple	R.1A
	Provide opt-in alternatives, e.g. in regard to the country/legal regime of the data storage location	R.1A
6. Ex post transparency	Make users conscious of their ex post transparency rights, so that they understand and can exercise their right of access	R.8C
	Make users aware of what information services providers have implicitly derived from disclosed data	R.2E
	Make users aware of the data processing and sharing practices of the service provider.	R.1I, R.2A, R.2B
	Help users making data traces transparent and promoting users' legal rights, e.g. by providing graphical interactive visualisations	R.6A
7. Audit configuration	Provide a standard way to perform audits across the chain of services. In particular, provide audit functions that visualise differences of SLAs along the cloud chain	R.1M
	Provide audit functions that make also implicitly collected data transparent	R.1L
8. Access control management	Allow users to classify their data items and easily provide access control rules for these data	R.1M
	Allow system administrators to verify the accuracy of access control rules in a straightforward and simple manner	R.7A
9. Privacy risk assessment	Provide different type of users (business users versus individual end users) with appropriate indicators obtained from risk assessment activities. Make risk awareness long lasting	R.1J, R.9E
	Provide clear visualizations of vulnerability of private data depending on different situations	R.2C

In Appendix B.1, we have mapped the functional categories in the table above to the functional requirements of the A4Cloud tools. This allows us to illustrate what HCI requirements and principles that A4Cloud tool functions in the A4Cloud scenarios will have to fulfil.

6.2 Towards HCI Guidelines for A4Cloud

In this section we present a set of high level HCI guidelines that are presented in a HCI pattern-like format and summarize a selection of the derived HCI principles that we think will be important for the A4Cloud tool development process. They are partly addressing the HCI challenges that we approached in task TC:7-2 in the first project year as laid out in Chapter 3.

As discussed in previous chapters, in comparison to traditional forms of outsourcing and internet services, transparency in the cloud requires that more complex information about the data handling along the cloud chain has to be provided to data subjects and other stakeholders. The guidelines presented in this section are in their objective similar to those for other kinds of usable privacy ecologies, even though TETs for the Cloud have at least to inform users about different additional aspects (as e.g. discussed in Sections 5.1.2 and 5.5.1). Nevertheless, for cloud usage it is important that developers apply them against the background of the complex picture of the cloud service chain.

The selected guidelines have so far an emphasis on tools for individual end users, even though they also include general guidelines that for all types of A4Cloud tools. They will be extended for the final report by work package C-7. Meanwhile, for a list of complete list of HCI principles for different stakeholder groups, the reader should refer to Section 6.1.

6.2.1 Motivate users to make informed decisions

Use different design techniques, UI elements and appropriate wording in order to encourage users to *care about* the information they are releasing to Internet services. Frame options and possible users' actions in terms of practical consequences that relate to the users and to their situation (see also 6.2.4). Use the influence of framing to motivate users to select those choices that are more likely to protect their information.

Users make decisions based on context. Present users with options or messages at a moment that is relevant to their actions at hand, such as the moment of uploading possible sensitive content to the cloud. Use the teachings from persuasive design in order to motivate users to take the appropriate actions by determining the targeted desired behaviour. For instance, if the targeted behaviour is to make users encrypt their data in the cloud even when there's a price or burden to it, then design for users to choose that option and frame the option in a way that portrays the benefits of encryption.

Functional categories:

Privacy preference management, obtaining consent.

Motivation:

From our experiments it can be seen that users lack the motivation to spend the cognitive efforts or time in carrying out appropriate measures to protect their privacy (R.3B, R.5A). As discussed above, this finding is in accordance to previous results which indicate that privacy and security online are not primary objectives of users.

Also, our results support the observation that users lack the motivation, and sometimes the appropriate knowledge, to make good judgments at the moment they are required to register to a service or disseminate their information. At the same time, results from the experiments described in Section 5.2.4 suggest the influence that the choice of wording can have on users' choices.

Examples:

A previous study has shown how users can be incited to care about tweaking their location privacy preferences by providing them with feedback about who, among their social network friends, were able to see their location (Tsai et al. 2009). Similarly, another study has explored the way to show privacy facts, for example, at the moment of installing a mobile app, which leads to users making better decisions about which apps to have on their device. See picture taken from (Kelley 2012).

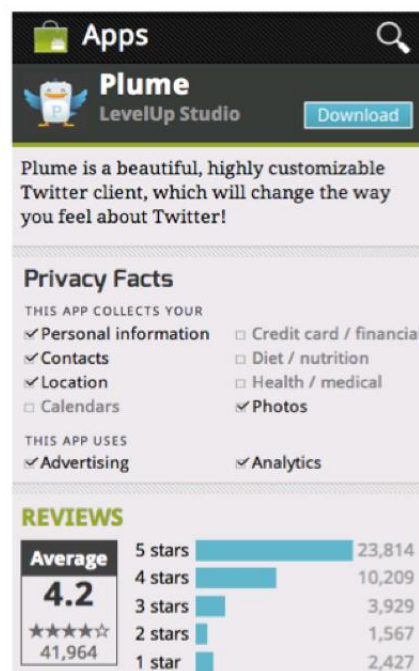


Figure 15. Example from Kelley et al. at making users decide on apps to install based on privacy facts.

Another study has shown the implications of displaying permissions on a mobile app at the moment that is more relevant, thus helping users understand what is being asked and how to act about it (Felt et al. 2012a; Felt et al. 2012b). For instance, when a mobile app needs to track the users' location it might be more appropriate to ask for permission at that moment instead of at the moment of installation of the app.

6.2.2 Help users comprehend policies and manage their preferences

A multi-layered approach for the display of policies, as recommended by the Art. 29 Working Party (Art. 29 Data Protection Working Party 2004), can help give a cleaner look to a policy by dividing the

information into layers of detail and importance. Policies can also be complemented by standard and meaningful icons representing data attributes, purposes of use, and data processing steps. As discussed in Section 5.5.2, proposals for privacy icons have been given by Mozilla¹⁹, Aza Raskin²⁰, and the PrimeLife project (Fischer-Hübner et al. 2010). As discussed in Section 5.5.2, further cloud-specific icons should be developed and used to inform about aspects that are often intransparent to users (such as geographic locations of data centres, applicable laws, how disclosure requests by law enforcement are handled). Provide privacy-friendly defaults of privacy preferences (see 6.2.7).

Use visual elements rather than plain texts to present summaries of privacy policies in understandable ways, for instance as tables to display purposes of data use (Kelley et al. 2009), *hoptrees* to navigate through hierarchies (Brooks et al. 2013), or branching trees to illustrate downstream data sharing through the chain of cloud services and other service providers. Whenever possible provide interactive elements within the policy visualizations that can help users adjust their preference on the fly.

Functional categories:

Ex ante transparency, privacy preference management.

Motivation:

Privacy policy information can especially be complex in the Cloud context. It is generally understood that people are not motivated to read privacy policies or manage their privacy preferences, but even those who are willing find these tasks difficult and time consuming (R.3C) (Gross & Acquisti 2005). An approach is needed to make the legibility and visualization of privacy policy easier and more comprehensible (R.5B). This is especially important when users' data is being shared among the chain of cloud services, which might have different policies for data handling and act under different legal regimes.

Examples:

The company iubenda²¹, creates privacy policies using an aesthetic design and multiple-layered approach, where in the first layer users are presented with a high level description of the data attributes collected and purposes of use (see figure below). Clicking in a data item takes users to the next layer displaying more detailed information about that item. Using meaningful icons (meaningful and informative also in a cloud context) and keeping text short can help people grasp the main aspects of the policy in an easier way.

¹⁹ https://wiki.mozilla.org/Privacy_Icons

²⁰ <http://www.azarask.in/blog/post/privacy-icons/>

²¹ <http://www.iubenda.com/en>

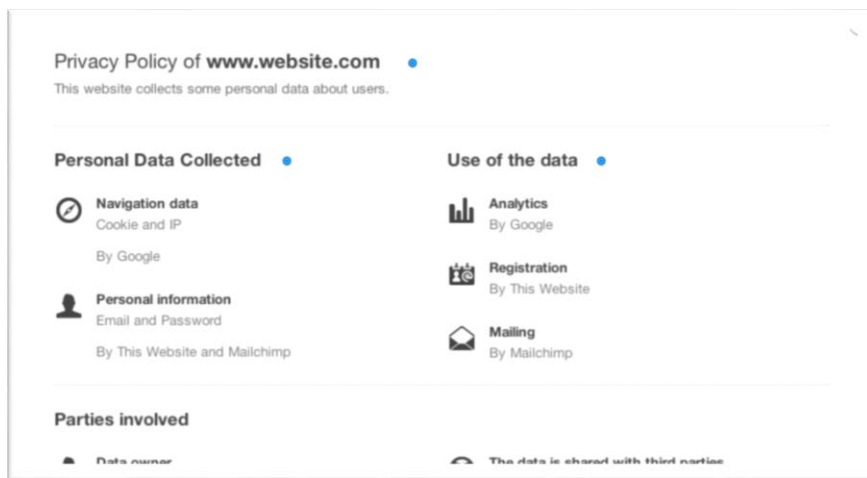


Figure 16. Example of a multi-layered privacy policy complemented with icons by iubenda.

Further example of UI designs presenting policies in multiple layers are design proposals based on a privacy nutrition label (Kelley et al. 2009) or PrimeLife's PPL (PrimeLife Policy Language) UIs that were elaborated for more complex PPL policy presentations (Angulo et al. 2012). See also the example of a *trace view* graphical interface in Section 5.3.5 for displaying data disclosures across the cloud chain, and/or sharing practices among other services.

6.2.3 Provide options for action

When alerting or providing information which might arouse doubts or unease on users do not only express messages in plain language and indicate the problem (Nielsen 1995), but also suggest understandable and effective approaches to act upon the presented message. Provide users with clear preventive measures, follow-up actions or exit strategies.

Sometimes, it will be too late to present users with feasible exit strategies or corrective measures. For instance, the action of disseminating data becomes very difficult to undo. However, in such cases, provide users with informative consequences of their previous data dissemination and possible solutions to protect their data rather than just explaining the problems. Moreover, in this context, inform users about how to exercise their legal data subject rights, and guide them in the process.

Functional categories:

Obtaining consent, exercising data subject rights, ex post transparency.

Motivation

Users should not feel they are left on their own when receiving notification of failures, or when decisions between two or more non-obvious choices have to be made.

The results obtained from the workshops presented in Section 5.1, corroborate findings stating that users are often not aware of their rights with respect to their personal data and lack a clear strategy to exercise these rights (R.1D). The user experience of an accountable tool could be improved when users are guided with the help of constructive solutions to their possible data-related dilemmas (R.9E, R9.H).

Examples:

The Ghostery plugin not only gives users a short preview of the trackers embedded in the visited website, but it also provides them with actions on how to act forward, such as investigating more information about the tracker, pause the blocking of trackers for a while, adjust their blocking options and tracking settings. With these alternatives users can feel more in control of their own decisions.

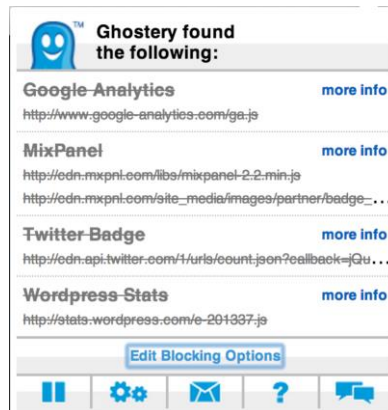


Figure 17. Example from the ghostery browser plugin

6.2.4 Frame in terms of consequences rather than technicalities

Provide meaningful choices that make the consequences or data releases more clear. Avoid using language that is difficult for average users to understand. Make an analysis of the possible implications of common user actions and explain various alternatives in terms of consequences that relate to the users' privacy. For instance, instead of simply asking users if they are sure about switching on or off a privacy-related setting, explain in clear and straightforward manner the ways their privacy could be affected by making the change. Also, look at *delayed gratification* (Singer 1955) strategies that can help explain the long-term benefits of protecting one's data in the cloud to users, as opposed to short term temporary rewards.

Functional categories:

Ex ante transparency, privacy preference management.

Motivation:

Most of the current privacy solutions and Internet service offerings are portrayed to users in terms that very few can actually understand or relate to. It is important that users understand the practical consequences of their actions when interacting with an Internet service and especially at the moment of releasing personal information (R.1K, R.3A, R.3B, R.9K).

Examples:

PViz is a tool that allows users to understand the visibility of their Facebook profile. With the use of such tools, users can be aware of the consequences of submitting a post on Facebook, since they can visualize the audiences reached by their posted content.

A User-Controllable Location-Sharing Tool, Locaccino²², developed at Carnegie Mellon University allows users to do more meaningful and expressive privacy settings in regard to their location based

²² www.locaccino.org

on time and location based rules data that make consequences more apparent to them (e.g., “My colleagues can only see my location when I’m on campus and only weekdays 9am-5pm”).

Another example is given by a report released by the SPION project (Gürses 2011) which lists a series of other “privacy feedback and awareness tools” for different purposes, some of which are good examples for explaining consequences of data releases to users that can also be very appropriate for the cloud.

6.2.5 Consider differences in users (cultures, expertise, legal regimes, etc.)

It is important to know the target users of a cloud service. In general, design for solutions that adapt to the types of intended users. When possible perform a cultural assessment to cover the expectations that different user groups may have in a cloud service. Provide sensitive defaults (as explained in Section 6.2.7) that are in harmony with the intended user groups. Use UI elements that are well understood and meaningful in different cultures and legislative regimes. Understanding the different laws that apply to a certain region is important, since data transfers across the chain of cloud services can often cross regional boundaries

Functional categories:

Ex ante transparency, ex post transparency

Motivation:

User groups are not the same and thus they should not all be provided with one solution. Our results and other studies have found that differences exist between cultures and users’ computer expertise with regards to the preferences, expectations and mental models of personal data in cloud services (R.2A-H, R.9G). Trivially translating the user interface and privacy statements would not have an effective impact in the different cultures (R.9G) where a privacy solution is being deployed.

Examples:

Certain browser features that are offered to customers from the US, like options for opting-out from cookies, are not as relevant in Europe where regulations protect customers by forcing opt-in cookies mechanisms on service providers.

As reported in Section 5.5.2, previous usability studies conducted in the PrimeLife project pointed out that some privacy icons that were well understood in some European countries, might not be understood by certain cultures where a particular image is not recognized (Fischer-Hübner & Zwingelberg 2010). For instance, the image of a post horn representing the purpose “shipping” which was well understood in Sweden or Germany, was not recognized by Chinese test participants. Moreover, the policy aspects for which icons can be helpful, vary across legal regimes.

6.2.6 Make trustworthiness transparent

Inform users about trustworthy practices in a transparent way at a service level rather than at an organizational level.

Use trust-enhancing elements in a consistent and efficient way. For instance, place standard trust seals in a visible way motivate users to investigate the meaning behind those seals in an easy manner. Consistency in brand of a product, on the look and placement of trust seals can make it easier for customers to recognize a trustworthy provider.

Functional categories:

Ex ante transparency, ex post transparency.

Motivation:

Many users' intrinsically believe that the Internet is insecure (R.9E) and that higher costs translate to more trustworthy services (R.9C). Trust in a service can be enhanced when the service makes its trustworthy practices transparent (R.9H).

At the same time, businesses tend to make decisions about which cloud providers to subscribe to based on the providers' ability to provide secure, stable environments at a low cost.

Transfer of trust in the company results in trust in all the company's services, even when these might be different (R.9B). For instance, if users trust Google's search functionality, they might transfer their trust onto other Google services, even when those other services might not be mature enough or not follow the same data handling practices.

Examples:

The Web Of Trust initiative (WOT)²³ is an example of showing a transparent trust evaluation of online services based on other users' reviews and experiences. A multi-layered approach is taken for displaying the evaluation results: The WOT browser plugin presents coloured icons and symbols to indicate the trustworthiness of a service to its users. By clicking on the icons, a website's WOT scorecard is shown that gives one detailed information about the site's reputation.

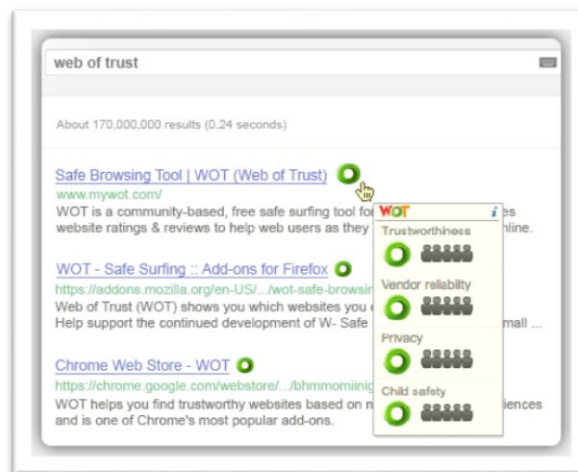


Figure 18. Example of the WOT plugin to indicate trustworthiness

6.2.7 Provide privacy-friendly and useful defaults

Depending on the context of a situation provide privacy-friendly defaults that will maintain the workflow of the application while at the same time protecting the users' data. Derive defaults that reflect the legal privacy principles of data minimisation and are based on techniques that adapt to the users' situation.

Functional categories:

Privacy preference management, access control management

Motivation:

²³ <http://www.mywot.com/>

Including good defaults when designing user interfaces is a recognized design principle (Tidwell 2005). The proposed EU Data Protection Regulation requires “Data Protection by Default” in its Art 23. This consideration becomes especially important when designing for privacy-enhancing technologies since privacy and security are often secondary tasks as compared to their primary objectives (this concerns all situations, not only cloud-dependent services).

Examples:

Earlier research studies have suggested the use of machine learning approaches to learn users’ privacy preferences under various contexts with the intention of relieving users of the complex tasks of having to specify their privacy preferences beforehand by creating sensible defaults that can match the actual users’ preferences (Tondel et al. 2011; Mugan et al. 2011; Ravichandran et al. 2009).

In PrimeLife, the approach was suggested of using default privacy preferences enforcing data minimisation, which can be adapted “on the fly” (if there is a mismatch with a website’s policy, the use can either accept this mismatch for the current transaction only or for all future transactions. In the latter case, the user’s preferences will be adapted accordingly) (Angulo et al. 2012).

6.2.8 Illustrate who is in control of the data

Make users feel empowered and confident by providing them with information about what data they can control and the reasons some of their personal data lie outside their personal control. Illustrations, consistent icons, tooltips and animated transitions could be used to demonstrate the differences between user-controlled data and service-controlled data. Make it clear what laws apply on the data controller’s side and what control rights technical control options the users have (compare R9.9F).

Functional categories:

Ex post transparency.

Motivation:

Our results (R.6B, R.6E) confirm earlier findings about the challenge for users to distinguish what data are located remotely under the services’ side and what is locally under their control. Many design alternatives have proven ineffective at communicating this aspect to users, who sometimes are unaware or sceptical about the possibility of being able to access and alter data on the services’ side.

The introduction of cloud technologies has created a blurry line for who is in control of data. Users are growing accustomed to the possibilities and conveniences of accessing their files and information from wherever they are and from multiple devices. This introduces users’ confusion at the moment of trying to understand where the actual data are located and who is in control. For users to be able to exercise their rights, they must understand the possibilities and limitations with regards to the manipulation of the data located at the services’ side outside their control, or locally under their control.

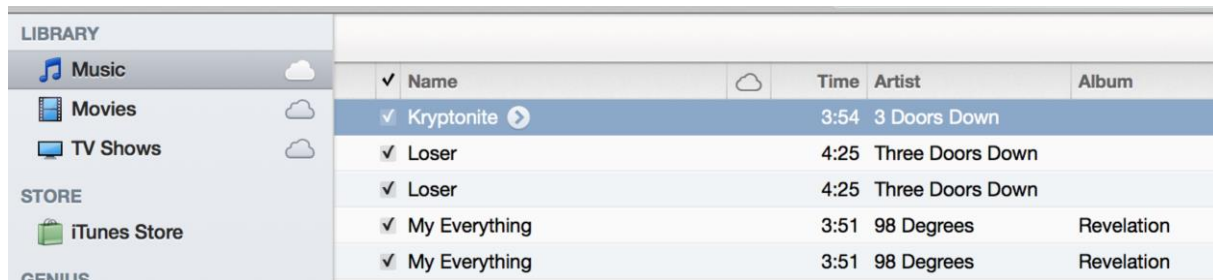
The connection of different services makes the situation even worse. For instance, services can fetch information and images from Facebook, while synchronizing them and displaying them in their own service. Instead of segregation of identities, there’s a coupling of the users’ identities throughout various services, which sometimes creates confusion.

Besides, a survey done by Wakefield Research²⁴ commissioned by Citrix, showed that 51% of respondents believed that the weather can affect cloud computing, and 95% of people who do not believe to be using cloud services, actually are subscribed to some kind of cloud provider. This is in accordance with our findings explained in Section 5.4, which indicated that it is not clear to users what constitute a cloud service and the implications of having their data in the cloud.

²⁴ <http://www.citrix.com/news/announcements/oct-2012/cloud-confusion-survey.html>

Examples.

The music players iTunes and Spotify have ways to tell users which music files are located on their devices and which are located remotely. In iTunes a cloud icon is used to represent songs that are accessed remotely through the iCloud. Similarly, Spotify's left menu panel has a dedicated music list where users can locate the files stored in their computers.



LIBRARY					
✓	Name		Time	Artist	Album
✓	Kryptonite	⬮	3:54	3 Doors Down	
✓	Loser		4:25	Three Doors Down	
✓	Loser		4:25	Three Doors Down	
✓	My Everything		3:51	98 Degrees	Revelation
✓	My Everything		3:51	98 Degrees	Revelation

Figure 19. Example of providing icons representing data in the cloud

6.2.9 Plurality of input and output

Consider different types of input and/or output modes, which can not only help to convey privacy implications and avoid preventing errors, but also improve the learnability and accessibility aspects of the system.

Functional categories:

All functional categories.

Motivation:

People are different and have different preferences for interacting with a program. By providing different input and output modes users might be more likely to adapt to the system and find it more enjoyable and suitable to their needs.

Examples:

Cloud services like Evernote allow users to “clip” web content through the right-click context menu, and also directly through the icon on a browser's plugin, thus providing multiple modalities for input.

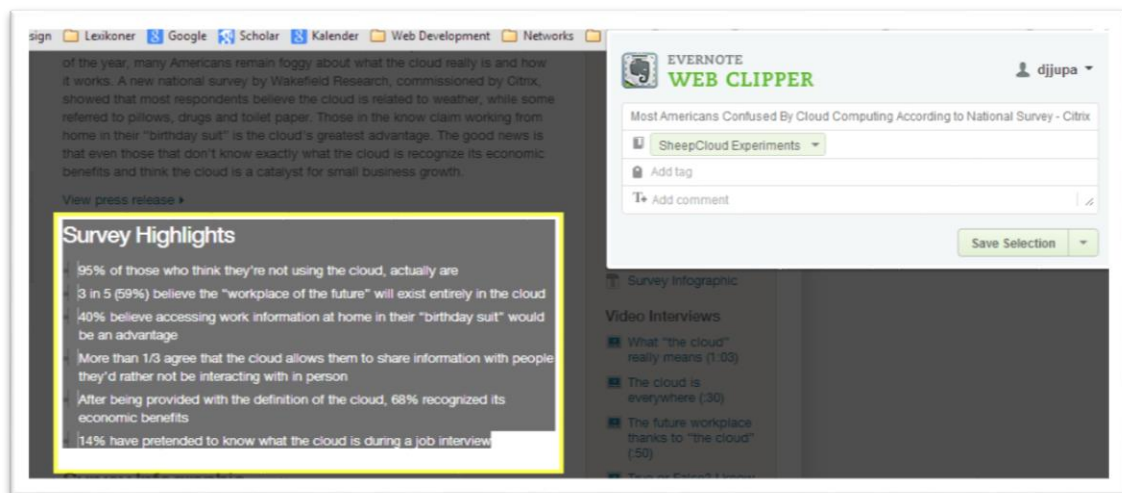


Figure 20. Example of providing multiple ways for inputting data

The prototyped Data Track program, introduced in Section 5.3, gives users different options to visualize their previous data releases using different views, such as a trace view or a timeline view, providing multiple output modalities. Each of these might be useful depending on what users are trying to find out or on the way they prefer to visualize data.

7 Concluding Remarks

In this deliverable, we have elaborated HCI concepts for making A4Cloud tools comprehensible and trustworthy. The deliverable reports on how we have followed a human-centred design approach to derive HCI requirements and related HCI principles for A4Cloud tools developed in the project for different stakeholder groups. For this, research has been conducted to analyse how users can be guided to better understand their data traces, how they can be supported to make better informed decisions in regard to the use of their data by cloud and other service providers, how legal privacy principles and social trust requirements can be enforced by A4Cloud tool user interfaces. A set of high level guidelines is finally presented that summarise a selection of the derived HCI principles and proposed design solutions.

This deliverable has also revealed more specific open HCI research challenges with regards to the implementation of our proposed HCI principles and guidelines, which we plan to address in the second project year. In particular, we will tackle the following research questions: How can ex ante transparency tools better inform users about the consequences of data disclosures? How can we derive and easily adapt good privacy default settings that are both privacy friendly and matching the user's preferences? How can ex ante transparency tools best illustrate and make obvious who is in control of the data and/or who is processing data under which conditions, and what means of legal or technical control exist in which situations? How can mismatches of policies or SLAs along the cloud chain be best presented to individual and business end users?

Further questions that we have not addressed yet and that we will take up in the second project year concern the level of detail with that users of ex ante transparency tools are interested to track the processing along the chain of cloud providers, as well how to communicate risk perceptions calculated by risk assessment tools to business end users in a transparent and comprehensible manner.

Research results on these questions will be reported in D:C-7.3 "Report on end user perceptions of privacy-enhanced transparency and accountability" that will be published at the end of the second project year.

References

- Alexander, C., Ishikawa, S. & Silverstein, M. (1977). Pattern languages. *Center for Environmental Structure*.
- Andersson, C., Camenisch, J., Crane, S., Fischer-Hübner, S., Leenes, R., Pearson, S., Pettersson, J.S. & Sommer, D. (2005). Trust in PRIME. In *Proceedings of the Fifth IEEE International Symposium on Signal Processing and Information Technology*. IEEE.
- Angulo, J., Fischer-Hübner, S., Wästlund, E. & Pulls, T. (2012). Towards usable privacy policy display and management. *Information Management & Computer Security*, 20 (1), 4-17.
- Art. 29 Data Protection Working Party (2004). *Opinion 10/2004 on More Harmonised Information Provisions*. (November 25th, 2004). European Commission.
- Art. 29 Data Protection Working Party (2010). *Opinion 1/2010 on the concepts of "controller" and "processor"*. (February 16th, 2010). European Commission.
- Art. 29 Data Protection Working Party (2012). *Opinion 5/2012 on Cloud Computing*. (July 1st, 2012). European Commission.
- Becker, R.A., Eick, S.G. & Wilks, A.R. (1995). Visualizing network data. *IEEE Transactions on Visualization and Computer Graphics*, 1 (1), 16-28.
- Beckerle, M. & Martucci, L.A. (2013). Formal definitions for usable access control rule sets from goals to metrics. In *Proceedings of the Ninth Symposium on Usable Privacy and Security (SOUPS '13)*. New Castle, United Kingdom, July 24-26, ACM.
- Bernard, H.R. (1988). *Research methods in cultural anthropology*. Sage Newbury Park, CA.
- Bernsmed, K., Felici, M., Santana De Oliveira, A., Sendor, J., Brede Moe, N., Rübsamen, T., Tountopoulos, V. & Hasnain, B. (2013). *D:B-3.1 Use Case Descriptions*. A4Cloud Project.
- Brandimarte, L., Acquisti, A. & Loewenstein, G. (2012). Misplaced confidences: Privacy and the control paradox. *Social Psychological and Personality Science*. 4 (3), 340-347. SAGE publications.
- Brede Moe, N., Gilje Jaatun, M., Haugset, B., Niezen, M. & Felizi, M. (2013). *D:B-2.4 Stakeholder Workshop 1 Results (Initial Requirements)*. A4Cloud Project.
- Brooks, M., West, J.D., Aragon, C.R. & Bergstrom, C.T. (2013). Hoptrees: Branching History Navigation for Hierarchies. In *Hoptrees: Branching History Navigation for Hierarchies. Human-Computer Interaction–INTERACT 2013*. Springer. 316-333.

- Brown, J. & Isaacs, D. (2005). *The world café: Shaping our futures through conversations that matter*. Berrett-Koehler Publishers.
- Cloud Industry Forum (2011). Transition to the cloud: The case for a Code of Practice. *CIF Report*, [Online].
- Dhamija, R. & Dussault, L. (2008). The Seven Flaws of Identity Management: Usability and Security Challenges. *IEEE Security and Privacy*, 6 (2), 24-29.
- European Commission (1995). Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. *Office Journal L*. 281. 23.11.1995.
- European Commission (2012). Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). *COM (2012) 11 Final*. Brussels, 25.1.2012.
- Felt, A.P., Egelman, S., Finifter, M., Akhawe, D. & Wagner, D. (2012a). How to ask for permission. In *Proceedings of the USENIX Workshop on Hot Topics in Security*.
- Felt, A.P., Ha, E., Egelman, S., Haney, A., Chin, E. & Wagner, D. (2012b). Android permissions: User attention, comprehension, and behavior. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*. ACM.
- Fischer-Hübner, S. & Zwingelberg, H. (2010). *UI Prototypes: Policy administration and presentation - Version 2*. PrimeLife Deliverable D.4.3.2. PrimeLife Project.
- Freeman, L.C. (2000). Visualizing social networks. *Journal of social structure*, 1 (1), 4.
- Garfinkel, S. (2005). *Design principles and patterns for computer systems that are simultaneously secure and usable*. Massachusetts Institute of Technology.
- Graf, C., Hochleitner, C., Wolkerstorfer, P., Angulo, J., Fischer-Hübner, S., Wästlund, E., Hansen, M. & Holtz, L. (2011). *Towards Usable Privacy Enhancing Technologies: Lessons Learned from the PrimeLife Project*. PrimeLife Deliverable D4.1.6. PrimeLife.
- Gross, R. & Acquisti, A. (2005). Information Revelation and Privacy in Online Social Networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*. Pittsburg, PA, USA. ACM.
- Gürses, S. (2011). *Deliverable D2.1 - State of the Art*. SPION project.
- Hildebrandt, M. (2009). *Behavioural biometric profiling and transparency enhancing tools*. FIDIS Deliverable, D7.12. March 2005. FIDIS EU project.

- Hoadley, C.M., Xu, H., Lee, J.J. & Rosson, M.B. (2010). Privacy as information access and illusory control: The case of the Facebook News Feed privacy outcry. *Electronic Commerce Research and Applications*, 9 (1), 50-60.
- Holtz, L., Nocun, K. & Hansen, M. (2011). Displaying privacy information with icons. In *PrimeLife/IFIP Summer School Proceedings 2010. Helsingborg, August 2-6, 2010*. Springer.
- International Standard Organization (ISO) (1998). *Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11: guidance on usability-Part 11 (ISO 9241-11:1998)*.
- International Standard Organization (ISO) (2010). *9241-210: 2009. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems (formerly known as 13407)*.
- Ion, I. (2012). *User-centered security mechanisms for protecting information sharing in the cloud*. Diss., Eidgenössische Technische Hochschule ETH Zürich, Nr. 20702, 2012.
- Ion, I., Sachdeva, N., Kumaraguru, P. & Capkun, S. (2011). Home is safer than the cloud!: privacy concerns for consumer cloud storage. In *Proceedings of the Seventh Symposium on Usable Privacy and Security*. Pittsburg, PA, USA. ACM. 13:1.
- Jaspers, M.W.M., Steen, T., Bos, C.v.d. & Geenen, M. (2004). The think aloud method: a guide to user interface design. 73 (11-12), 781-795.
- Johnston, J., Eloff, J.H. & Labuschagne, L. (2003). Security and human computer interfaces. *Computers & Security*, 22 (8), 675-684.
- Kani-Zabihi, E., Helmhout, M. & Coles-Kemp, L. (2012). Increasing Service Users' Privacy Awareness by Introducing On-line Interactive Privacy Features. *IAAC Symposium 2011*, [Online].
- Kelley, P. (2012). Privacy as Part of the App Selection Process. In Sonia Chiasson & Jaeyeon Jung (eds.). In *Proceedings of the Workshop on Usable Privacy & Security for Mobile Devices (U-PriSM)*. Washington DC, USA: Symposium On Usable Privacy and Security (SOUPS) 2012.
- Kelley, P.G., Bresee, J., Cranor, L.F. & Reeder, R.W. (2009). A "Nutrition Label" for Privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09)*. Mountain View, CA. USA. ACM.
- Kolter, J., Netter, M. & Pernul, G. (2010). Visualizing past personal data disclosures. In *ARES'10 International Conference on Availability, Reliability, and Security, 2010*. IEEE. 131.

- Kshetri, N. (2012). Privacy and security issues in cloud computing: The role of institutions and institutional evolution. *Telecommunications Policy*.
- Lacohée, H., Crane, S. & Phippen, A. (2006). Trustguide: Final Report. *Trustguide*. October 2006.
- Langer, E.J. (1975). The illusion of control. *Journal of personality and social psychology*, 32 (2), 311.
- Maguire, M. & Bevan, N. (2002). User requirements analysis. In *Proceedings of IFIP 17th World Computer Congress*.
- Marshall, C. & Tang, J.C. (2012). That Syncing Feeling: Early user experiences with the cloud. In *Proceedings of the Designing Interactive Systems Conference*. ACM.
- Mugan, J., Sharma, T. & Sadeh, N. (2011). Understandable Learning of Privacy Preferences Through Default Personas and Suggestions.
- Nielsen, J. (1995). Usability inspection methods. In *Conference companion on Human factors in computing systems*. ACM.
- Owen, H. (2008). *Open space technology: A user's guide*. Berrett-Koehler Pub.
- Patrick, A.S. & Kenny, S. (2003). From privacy legislation to interface design: Implementing information privacy in human-computer interactions. In *Privacy Enhancing Technologies*. Springer. 107.
- Patrick, A.S., Kenny, S., Holmes, C. & van Breukelen, M. (2003). Human Computer Interaction. In G.W. van Blarckom, J.J. Borking & J.G.E. Olk (ed.) *Handbook of Privacy and Privacy-Enhancing Technologies: The Case of Intelligent Software Agents*. College Bescherming Persoonsgegevens, Den Haag, The Netherlands. 249-290.
- Pearson, S. (2013). Privacy, Security and Trust in Cloud Computing. In *Privacy, Security and Trust in Cloud Computing. Privacy and Security for Cloud Computing*. Springer. 3-42.
- Pearson, S., Tountopoulos, V., Catteddu, D., Sudholt, M., Molva, R., Reich, C., Fischer-Hübner, S., Millard, C., Lotz, V. & Jaatun, M.G. (2012). Accountability for cloud and other future Internet services. In *IEEE 4th International Conference on Cloud Computing Technology and Science (CloudCom), 2012*. IEEE.
- Pettersson, J.S. (2008). *HCI Guidelines. PRIME Deliverable D06.1.f. Final Version*. PRIME project.
- Pettersson, J.S., Fischer-Hübner, S., Danielsson, N., Nilsson, J., Bergmann, M., Kriegelstein, S.C.a.T. & Krasemann, H. (2005). Making PRIME usable. In *Proceedings of the 2005 Symposium on Usable Privacy and Security (SOUPS '05)*. Pittsburg, PA, USA. ACM.

- PrimeLife WP4.1 (2010). HCI Pattern Collection -- Version 2. In S. Fischer-Hübner, C. Köffel, J. Pettersson, E. Wästlund & H. Zwingelberg (eds.), *PrimeLife Deliverable D4.1.3*. PrimeLife (<http://www.primelife.eu/results/documents>).
- Pulls, T. (2012). Privacy-Friendly Cloud Storage for the Data Track. In *Privacy-Friendly Cloud Storage for the Data Track. Secure IT Systems*. Springer. 231-246.
- Raskin, A. (2010). Privacy Icons: Alpha Release.
- Ravichandran, R., Benisch, M., Kelley, P.G. & Sadeh, N.M. (2009). Capturing social networking privacy preferences. In *Privacy Enhancing Technologies*. Springer.
- Rubin, J. & Chisnell, D. (2008). *Handbook of usability testing : how to plan, design, and conduct effective tests*. Indianapolis, Ind.: Wiley Publ.
- Shin, D. (2013). User centric cloud service model in public sectors: Policy implications of cloud services. *Government Information Quarterly*.
- Singer, J.L. (1955). Delayed gratification and ego development: implications for clinical and experimental research. *Journal of consulting psychology*, 19 (4).
- Spiekermann, S., Grossklags, J. & Berendt, B. (2001). E-privacy in 2nd generation E-commerce: privacy preferences versus actual behavior. In *Proceedings of the 3rd ACM conference on Electronic Commerce*. Tampa, Florida, USA. ACM.
- Svensk Författningssamling Riksdagen. *Patientdatalag (2008: 355)*.
- Tidwell, J. (2005). *Designing Interfaces : Patterns for Effective Interaction Design*. O'Reilly Media.
- Tondel, I.A., Nyre, A. & Bernsmed, K. (2011). Learning privacy preferences. In *Proceedings of the Sixth International Conference on Availability, Reliability and Security (ARES), 2011*. IEEE.
- Tsai, J.Y., Kelley, P., Drielsma, P., Cranor, L.F., Hong, J. & Sadeh, N. (2009). Who's viewed you?: the impact of feedback in a mobile location-sharing application. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2003.
- Turner, C.W., Zavod, M. & Yurcik, W. (2001). Factors that affect the perception of security and privacy of e-commerce web sites. In *Fourth International Conference on Electronic Commerce Research, Dallas TX*.
- Tversky, A. & Kahneman, D. (1985). The framing of decisions and the psychology of choice. In *The framing of decisions and the psychology of choice. Behavioral decision making*. Springer. 25-41.

- Voida, A. & Olson, Judith S Olson& Gary M (2013). Turbulence in the Clouds: Challenges of Cloud-Based Information Work.
- Wamala, C. (2010). Does IT count?: complexities between access to and use of information technologies among Uganda's farmers. *Sort*, 20 (50).
- Wästlund, E. & Fischer-Hübner, S. (2010). End User Transparency Tools: UI Prototypes. PrimeLife Deliverable D.4.2.2. PrimeLife project.
- Wästlund, E., Wolkerstorfer, P. & Köffel, C. (2010). PET-USES: Privacy-enhancing technology - users' self-estimation scale. In Bezzi, M., Duquenoy, P., Fischer-Hübner, S., Hansen, M. and Zhang, G. (eds.) *Privacy and Identity Management for Life (IFIP Summer School 2009 Proceedings)*. Springer. 266-274.
- Whitten, A. & Tygar, J.D. (1999). Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. In *the Proceedings of the 8th USENIX Security Symposium*.
- Wyld, D.C. (2010). The Cloudy future of government IT: Cloud computing and the public sector around the world. *International Journal of Web & Semantic Technology*, 1 (1), 1-20.
- Xu, H. (2007). The effects of self-construal and perceived control on privacy concerns. In *Proceedings of the 28th Annual International Conference on Information Systems (ICIS 2007)*.
- Yee, K. (2004). Aligning security and usability. *IEEE Security and Privacy*, 2 (5), 48-55.

Appendices

Appendix A.1: Experiment with fake cloud service

1. Introduction text

"A new cloud storage service called SheepCloud lets you meet new friends with your same interests by letting you share files with your creative content.

For example, you can upload music files of songs you have created, or text files with poems you have written or photoshop files of pictures you have edited, etc.

We need some people to register to SheepCloud in order to test the program that matches people's profiles. This will allow us to improve the SheepCloud service before is ready to the public.

As a reward for registering early to SheepCloud you will get additional cloud storage space (Gigabytes) for free!"

2. Non-sensitive vs sensitive questions

The intention of differentiating between non-sensitive and sensitive questions was to have a control and an experimental group, in which participants of the experimental group were made to believe that they were releasing sensitive data into a new and unfamiliar cloud storage service provider.

Non-sensitive questions	Sensitive Questions
"Are you married?"	"Are you married?"
"Have you ever done any kind of voluntary service?"	"Do you believe hitting children when they misbehave is appropriate?"
"Have you ever made a donation to a non-profit organization?"	"Have you ever had sex in a public venue (e.g. public restroom, airplane, etc.)?"
"Have you ever been traveling for more than one month?"	"Have you ever cheated on your partner?"
"Have you ever travelled outside the continent you were born?"	"Have you ever downloaded or bought pirated movies or music?"
"Do you have a vehicle (car, motorcycle, minivan, motorhome, etc)?"	"Have you ever been caught stealing something?"
"Do you have any pets?"	"Have you ever used drugs of any kind?"
"Can you speak more than two languages fluently?"	"Are you insecure or ashamed about certain parts of your body?"

"Have you ever bought electronics through the Internet?"	"Have you ever contracted any sexually transmitted disease (STD)?"
"Do you play any instrument?"	"Do you practice any sort of religious activities more than twice per month?"

Appendix A.2: Data Track usability tests

1. Introduction to the test

"You are about to test a computer program called the "Data Track" that lets you see a history of the information you have given to different companies on the Internet. By using "Data Track" it is possible to see which information the companies have stored about you and also it is possible to know if that information is the same as the information you have sent (or if the companies have changed it). If the company allows it, it is also possible to correct or delete the data they have stored about you.

In this test you are going to pretend that you have previously given information to some Internet companies, like Facebook, Spotify, Google, and few others. This time, you are about to buy a book on the Internet with [adBokis.com]."

2. Test tasks

#	Question / task
1*	What do you think the elements on the top represent?
2*	What do you think the elements at the bottom represent?
3	Using the Data Track's trace view, how can you see the information that you have sent to adbokis?
4	How can you see to which Internet services have you given your email address?
5	Where would you click to see the information that adbokis has stored on their servers when you purchased the book?
5.1	In your opinion, can others access your data that adbokis has stored on their servers?
5.2	What information about you does adbokis have on their servers?
5.3	Does adbokis stored the location you were in when you bought the book?
5.4	Is the information that adbokis have about you more or less that what you gave to them? Why?
6	The Data Track gives you an overview of the information you have given to different Internet services. Where is this information stored?

7	Rate how secure do you think is your information being shown by the Data Track?
8	In your opinion, who other than you has access to the information being shown by the Data Track?
9	[Possibly] Another way of showing the information you have given to Internet companies is in chronological order. What do you think of this "timeline" view?
10*	How would you remove a piece of information from the Data Track?
10.1*	If you click and hold a piece of information, a trashcan appears that lets you delete that piece of information. What do you think happens when the information is deleted from the Data Track?
10.2*	If you click and hold a service, a trashcan appears. What do you think happens when you delete a service from the Data Track?
11	What would you do to delete or correct the information that you sent to adbokis?

* means that the questions were always presented in that order (the rest of the questions were shuffled to account for counterbalancing)

3. Post-questionnaire²⁵

- Rate how much you agree or disagree with each of the following statements concerning the Data Track program
 - This program gives me an idea of the risk to have my identity stolen
 - This program helps me see which information Internet services have about me.
 - If I regret sending information to an Internet service, I can remove that information with the help of this program.
 - My personal information that is shown in the program is completely secure
 - This program helps me see how much I have used a particular user name or email address
 - This program helps me get a good view of who knows what about me
 - Nobody else can access the personal information that is shown in the program, only I have access
 - This program helps me see the Internet services to which I have given my information.
- Which of the following would best describe your emotions when looking at the information displayed by the Data Track program
 - Relaxed or Calm

²⁵ Look for possible updates of the questions here: <http://edu.surveymizmo.com/s3/1268306/Data-Track-Trace-View-Postquestionnaire>

- b. Scared or distressed
 - c. Sad or grouchy
 - d. Happy or pleased
 - e. Quiet or not caring
 - f. Tired or drowsy
 - g. Excited or enthusiastic
 - h. Astonished or surprised
3. How often do you believe you would use the Data Track program to check or modify the information you have given to different Internet services?
- a. Very often (almost always)
 - b. Often (around two to four times per week)
 - c. Sometimes (a few times per month)
 - d. Rarely (a few times per year)
 - e. Very rarely (almost never or never)
4. If you would have the Data Track program available, how often do you think you would have the program turned on so that it tracks the information you give to Internet services?
- a. Always tracking (100% of the time)
 - b. Often tracking (75% of the time)
 - c. Sometimes tracking (50% of the time)
 - d. Rarely tracking (25% of the time)
 - e. Never tracking (~0% of the time)
5. Which of the two views of the Data Track would you prefer to use?
- a. Trace view
 - b. Timeline

Appendix B.1: Matching General HCI Requirements and Principles to the High-level Functional Analysis of the A4Cloud Scenarios

The functionalities derived from the high level functional view of the A4Cloud scenarios in the A4Cloud Deliverable D:B-3.1 (Use case descriptions) are tabulated in the next subsections, with one table for each A4cloud stakeholder.

The functional categories of HCI requirements and principles derived in chapter 5 and listed in Section 6.1 are mapped to the high-level functionalities of the A4Cloud scenarios to illustrate the HCI requirements and principles that A4Cloud tool functions in the A4Cloud scenarios will have to fulfil.

F1 – Functionality for individual end users (cloud users)

Table 18. Functional categories for HCI requirements and principles mapped to the A4Cloud scenario functionalities for individual end users (cloud users)

ID	Functionality (cf. A4Cloud scenarios)	Description	Categories for HCI requirements and principles (cf. Table 17)
<i>Policy Management</i>			---
F1-1	Edit policy	Create, modify and delete a user policy about the use of personal data	Privacy preference management
F1-2	Edit access rights	Set, view and modify access rights to personal data	Access control management; Policy preference management
F1-3	Configure time period of use	Set the time period for keeping personal data in the cloud	Privacy preference management;
F1-4	Delegate right to reconfigure policy	Allow another cloud actor change the configuration of a specific user policy	Privacy preference management; Access control management
F1-5	Accept policy	Accept the policy of a cloud provider/cloud service user	Ex ante transparency; Obtaining Consent
F1-6	Accept purpose of use	Accept the purpose of use of personal data from specific cloud provider / cloud user	Ex ante transparency; Obtaining consent; Privacy preference management
F1-7	Select policy	Browse sample policies and select policy for the use of personal or confidential data	Ex ante transparency; privacy preference management

ID	Functionality (cf. A4Cloud scenarios)	Description	Categories for HCI requirements and principles (cf. Table 17)
F1-8	Receive policy notification	Receive notifications on the status of the policy enforcement of the cloud provider / cloud user, including policy violations	Ex post transparency
F1-9	Report violation	Report any policy violation experienced in the use of cloud services	Ex post transparency
F1-10	Report infringement	Report a misuse experienced in cloud provider/cloud service user implementing accountability practices	See F1-9 Exercising data subject rights
<i>Data Governance</i>			---
F1-11	View policy settings	Request to explore the fields comprising the user policy on governing the use of personal data	Privacy preference management
F1-12	Select data [that can be externalised]	Decide which personal data can be transferred outside the primary service provider's own IT systems	Privacy preference management; access control management
F1-13	Edit data	Correct or delete the personal data used (even if they are "in the cloud")	Exercising data subject rights
F1-14	Track data	Track the use of personal data (including data "in the cloud")	Ex post transparency
F1-15	Analyse use	Analyse the trace on the use of personal data with respect to how data are stored by the cloud provider, what data have been collected, for what purposes and when and who accessed this data	Ex post transparency; audit configuration
F1-16	Request data tracking	Select which personal data used "in the cloud" should be tracked	Ex post transparency

ID	Functionality (cf. A4Cloud scenarios)	Description	Categories for HCI requirements and principles (cf. Table 17)
F1-17	Receive notification on data management	Receive notifications on actions with respect to data management, based on user policy (e.g. deletion of expired data)	Ex post transparency
<i>Compliance Check</i>			
F1-18	Request compliance check	Request a compliance check of a cloud provider or cloud user	Exercising data subject rights; audit configuration
F1-19	Receive compliance check results	Get the results of the compliance check of a cloud provider / cloud user	Ex post transparency
F1-20	Request role obligations	Explore the actor's responsibilities, based on the policy for handling corporate data	Ex post transparency; audit configuration
F1-21	Request conformance	Request compliance with policies on the use of confidential data	Ex post transparency
F1-22	Summary of actions	Request the actions with respect to policy enforcement and the relevant incidents for a given period of time	Ex post transparency
F1-23	Navigate through actions	Filter the list of actions with respect to policy enforcement, based on performer and incident	Ex post transparency
F1-24	Risk notification	Receive notifications on potential risks derived from the policy settings of the cloud provider / cloud user	Ex ante transparency

F2 – Functionality for business end users (cloud users)

Table 19. Functional categories for HCI requirements and principles mapped to the A4Cloud scenario functionalities for business end users (cloud users)

ID	Functionality (cf. A4Cloud scenarios)	Description	Categories for HCI requirements and principles (cf. Table 17)
<i>Policy Management</i>			
F2-1	View regulation framework	Explore the provisions and restrictions of the data protection law	Ex ante transparency
F2-2	Request for regulation framework	Search for the appropriate regulation framework governing the execution of a specific application scenario	Ex ante transparency
F2-3	Receive policy notification	Receive notifications on the status of the policy enforcement for personal and corporate data, including policy violations	Ex post transparency
F2-4	Analyse violation	Track the policy violation data to identify which parties are affected and which personal and/or corporate data are violated and how	Ex post transparency; audit configuration
F2-5	Edit policy	Create, modify or delete a policy about the use of corporate data and devices	Privacy policy management
F2-6	View redress actions	Explore the list of recommended actions in case of receiving a policy notification, such as a policy violation	Ex post transparency
F2-7	Implement redress actions	Select and implement the action(s) to remediate and redress the incident caused the notification alert	
F2-8	Inform users	View and submit automatically generated notifications for infringements on the use of corporate data subjects	Ex post transparency;

ID	Functionality (cf. A4Cloud scenarios)	Description	Categories for HCI requirements and principles (cf. Table 17)
F2-9	List users	View the list of individual end users associated with a policy on the use of corporate data	Ex post transparency
<i>Data Governance</i>			
F2-10	View policy settings	Request to explore the fields comprising the user policy on governing the use of personal data	Privacy policy management
F2-11	Track personal data	Track the reference to the personal data (but not the contents of the personal data) of those involved in the execution of corporate processes	Ex post transparency
F2-12	Analyse use	Analyse the trace on the use of personal and corporate data with respect to how data are stored by the cloud provider, what data have been collected and when and who accessed this data	Ex post transparency
F2-13	Request data tracking	Select which personal data and corporate data used in the cloud should be tracked	Ex post transparency
F2-14	Select data	Decide which corporate data can be placed in the cloud	Access control management
F2-15	Edit data	Correct or delete the personal data used in the cloud	Access control management
<i>Compliance Check</i>			
F2-16	Match data	Match personal data and corporate data collected with the terms of the contract established with the cloud provider	Policy management; ex post transparency
F2-17	Negotiate contract	Negotiate the contract terms to establish agreement with the cloud provider	Ex ante transparency

ID	Functionality (cf. A4Cloud scenarios)	Description	Categories for HCI requirements and principles (cf. Table 17)
F2-18	Collect data for evidence	Collect data from the cloud as evidence to configure the proper policy enforcement	Ex post transparency
F2-19	Share evidence	Share results on the evidence collection data with individual cloud end users	Ex post transparency
F2-20	Request compliance check	Check corporate data governance policies with respect to regulation	Ex ante transparency
F2-21	Select processes	Define corporate data governance policy process	
F2-22	Report on compliance	Prepare reports on corporate compliance to legislation bodies	
<i>Risk Analysis</i>			
F2-23	Perform risk analysis	Define which data will be used for risk assessment and request risk analysis	Privacy risk assessment
F2-24	Define risk model	Select which risk analysis model (including configuration thresholds) should be adopted to run risk analysis	Privacy risk assessment
F2-25	Define trust model	Select which trust model (including configuration thresholds) should be adopted to run risk analysis	Privacy risk assessment
F2-26	Explore cloud providers	Explore the list with the associated cloud providers	
F2-27	View risk results	View risk analysis results	Privacy risk Assessment; ex ante transparency

F4 – Functionality for cloud auditors

Table 20. Functional categories for HCI requirements and principles mapped to the A4Cloud scenario functionalities for cloud auditors

ID	Functionality (cf. A4Cloud scenarios)	Description	Categories for HCI requirements and principles (cf. Table 17)
F4-1	Collect data for evidence	Collect data from the cloud, including corporate incident handling procedures, as evidence that accountability practices are being followed	Ex ante transparency, audit Configuration
F4-2	Accountability support	Report on the results on accountability checks, provide recommendations towards accountability compliance and legal guidance for redress	Ex ante transparency
F4-3	Certify accountability	Certify compliance with data protection legislation	Ex ante transparency
F4-4	View policy notifications	Explore the list of the policy-related notifications, which have been generated, in order to assess their severity	Ex ante transparency
F4-5	Verify risk analysis	Review process on risk assessment	Privacy risk assessment
F4-6	Verify mitigation actions	Check privacy impact assessment and mitigation plan and review on remediation and redress actions	Privacy risk assessment
F4-7	Accountability alert	Generate alerts and notifications in case that a cloud actor is not accountable	Privacy risk assessment
F4-8	List accountability actions	View the list of responsibilities for the involved cloud actors, associated with liabilities	Ex post transparency
F4-9	Suggest compensation	Decide on sanctions in case of infringement	---
F4-10	Revoke certification	Revoke certificates from cloud actors	---