Assessments of a Cloud-Based Data Wallet for Personal Identity Management

Farzaneh Karegar *farzaneh.karegar@kau.se*
Karlstad University
Karlstad, Sweden

Daniel Lindergren *dlindegren93@gmail.com*
Karlstad University
Karlstad, Sweden

John Sören Pettersson *john_soren.pettersson @kau.se*
Karlstad University
Karlstad, Sweden

Simone Fischer-Hübner *simone.fischer.huebner@kau.se*
Karlstad University
Karlstad, Sweden

Abstract

Within a project developing cloud technology for identity access management, usability tests of mockups of a mobile app identity provider were conducted to assess users' consciousness of data disclosures in consent forms and flow of authentication data. Results show that using one's fingerprint for giving consent was easy, but most participants had not a correct view of where the fingerprint data is used and what entities would have access to it. Familiarity with ID apps appeared to aggravate misunderstanding. In addition, participants could not well recall details of personal data releases and settings for disclosure options. An evaluation with a confirmation screen slightly improved recall rate. However, some participants voiced a desire to have control over their data and expressed a wish to manually select mandatory information. This can be a way of slowing users down and make them reflect more.

Keywords: Cloud computing, Identity provider, Identity management, Smartphone, Data disclosure, Usability, Privacy

1. Introduction

With the growth of internet people often need to prove their rights to access services not by showing an identity card or membership card, but by digitally authenticate themselves. They also need to provide personal information in that context. Helping service providers to trust users' information and users to trust that their personal information is not used for other purposes than they intend when providing it is an urgent matter in the growing information society. In addition, usability issues arise with growing information and security demands.

CREDENTIAL[1] is an EU-funded Horizon 2020 project that involves developing, testing and presenting cloud-based services for storing, managing and sharing digital identity information and personal data with a higher level of security than existing technology.

The CREDENTIAL *Wallet* is the central component of the tools developed within the project. It offers a set of security and application services providing, among others, authentication and authorization mechanisms combined with novel cryptographic technologies

[1] https://credential.eu/

like proxy re-encryption and malleable signatures. Some specific pilot cases are developed within the project to demonstrate how the CREDENTIAL technology can be deployed in diverse contexts.

In order to display the CREDENTIAL *Wallet* solution, the demonstrated system has to be equipped with user interfaces (UIs). Besides the pilot-specific user interfaces, there must also be some general user interfaces to demonstrate the functionality. The present study covers a set of three user tests (*n* = 3 x 20 participants) made of the core functions of authorization and authentication to which the general user interfaces of the *Wallet* will give access. UI mockups were inserted in interactive prototypes in order to evaluate ordinary Internet users' management and consciousness of data disclosures in consent forms in a mobile app as well as their mental models of the flow of authentication data. The results are of general interests as they demonstrate users' preferences, doubts, and misunderstandings of a type of identity management technology that could possibly be a facilitator in the ever-more digital economy of the world.

This presentation is structured in the following way. Section 2 gives an overview of the project from which this study stems. In Section 3, existing works on protecting users' privacy by improving permission dialogues and users' perceptions of data flow in the context of digital identity providers are described, and we point to lacunas in the body of previous studies. The questions the user tests were addressing are presented in Section 4. Section 5 elaborates on the set-up of the three tests. The results and related discussions, respectively, come in Section 6 and 7. Finally Section 8 includes conclusion and future works.

2. Background to the Present Study: the CREDENTIAL Project

The CREDENTIAL project goal is to enable an information sharing network for cloud-based identity information in which *even the identity provider cannot access the data in plain-text*; hence, the CREDENTIAL technology protects the data owner's right to define the right to access her/his personal data. [11, 15]

CREDENTIAL's basic architecture [16] integrates cryptographic mechanisms for users, the *Wallet*, and data receivers (web shops, cloud services, etc.).

- The User owns data that might be securely stored or shared with other account-holders in the CREDENTIAL *Wallet*. An application in the user's smartphone (and other devices) handles cryptographic operations involving the user's private key, such as signing or generating a re-encryption key.
- The *Wallet* is a cloud-based data storage and sharing service ensuring constant availability on the Internet, scalability, and cost effectiveness. It implements a multi-factor authentication and authorizes access to data stored. When a re-encryption key is available for some specific set of data as specified by the user, these data can be shared with specified receivers even if the user or his/her client application are not available.
- The Data Receiver, who can be either another person or a service provider, access data stored in the CREDENTIAL *Wallet* or authentication assertions issued by the *Wallet*.

The user authenticates at the *Wallet* to get read and write permissions to her *Wallet* account, which are used to upload signed data and other encrypted data. Upon a data request from, e.g., a web site, the user lets the *Wallet* generate a re-encryption key towards this data receiver and defines a "policy" specifying which data may be disclosed. These disclosure rules are stored in the *Wallet*. A receiver will thus receive encrypted data according to the disclosure rules and is able to decrypt the data (and to verify any digital signature on the disclosed data).

Within the project, three different pilots in the eGovernment, eHealth and eBusiness domains are developed [11, 16]. Some core functionality will appear to any user of the CREDENTIAL technology; the personal *Wallet* account is important and accessed through an app in the data owner's smartphone (several access devices can be defined). The user interfaces of this app can be used to evaluate not only people's ability to use such an app, but also for important general questions concerning people's understanding of consent-giving to data-disclosure and their appreciation of more privacy-friendly single-sign-on solutions than what is presently offered to the general public. Only a fraction has been utilized as yet, but in order to motivate our study the next section will present other studies within the field to highlight identified problems as well as lacunas in the research. In Section 4 we will elaborate on the precise research questions we address here.

3. Previous Studies

CREDENTIAL *Wallet* supports users with its functionalities in a mobile application acting as an identity provider (IdP) and a data access manager. As mentioned in Section 1, prototyping for a mobile app providing the authentication and authorization functionalities, we aimed at investigating people's understanding of the concept of sign up and sign in to a service provider using a mobile application as an identity provider and their perceptions of the data flow between entities when users share their personal information from the/their *Wallet* and give their consents via authorization dialogues.

Researchers have investigated users' comprehensions and attitudes towards using identity providers (e.g., OpenID or different social network identity providers) on the desktop enumerating some misconceptions and problems regarding understanding the data flow.

The most common in-use identity providers nowadays are social network single-sign-on (SSO) systems like Facebook in first place and Twitter and Google the second [26]. Hence, work on understanding people's information sharing is mainly focused on Facebook acting as an IdP. Besmer's et al. [5] study show that people are not aware of data sharing, its risks and implications using Facebook SSO. Robinson's et al. [20] in line with Egelman's work [10] showed that people have quite good general understanding of the information sharing using an IdP. However, it could be improved by designing better interfaces and help them to understand the exact information being shared. Also, Bauer et al. [4] conclude that people are somewhat aware of the range of attributes passed by the IdP to the service provider but aligned with Egelman's work they emphasized the fact that due to habituation, people do not pay attention to the exact content of the consent forms and have some preconceptions about what is shared.

The number of mobile apps available has exploded over the past few years; many of them request permissions to access private data and resources, like user accounts or location. Users, however, are often unaware of this kind of access even though they must grant the required permissions upon app installation or updating processes. There have been studies on the app permission dialogues which investigate the effects of dialogues on people's decision making and comprehensions and give new suggestions to improve the users' privacy [25, 17]. However, to the best of our knowledge, there is not any previous work on the effectiveness of identity providers' authorization dialogues on the mobile devices to help people understand what they are sharing with which entity, and the users' perceptions of data flow when they use a mobile application as an intermediary to sign up for a service provider.

Considering the problem of sharing information and giving the consent without proper level of knowledge about what is shared in the context of identity providers, some researchers tried to improve front solutions to protect users' privacy by proposing more privacy-friendly designs. For example, Javed and Shehab [12, 13] utilized eye-tracking techniques, and animation to enforce and catch users' attention towards permissions, and Wang et al. [27, 28] proposed new interfaces for Facebook authorization dialogue to eliminate the previous problems of the interfaces and improve the awareness and users' control over their personal data.

Besides lack of attention to and understanding of what is shared, people have incorrect security mental models when they use an IdP. Sun's et al. [24] and Arianezhad's et al. [1] studies show that people believe their identity provider's passwords are shared with service providers. In the CREDENTIAL *Wallet* app, giving consent is done entering a pin code or by a single touch on the Touch ID to scan the fingerprint (only the latter was used in user tests reported below). That is, not only for signing in to the IdP but also for giving consent and confirming the information sharing shown in the authorization dialogue, users should scan their fingerprints (or enter the pin code) as this assures the identity of the user who gives the consent.

This paper reports, thus, on investigations into users' consciousness of data disclosures in consent forms and users' perception of flow of authentication data.

4. Evaluation Goals

Even if a usability study can be thought of as aiming for a swiftly functioning service which users mindlessly can safely use, minimizing task completion times is not the aim in this study, but rather to find what are obstacles for users, especially as concerns their ideas – their mental models – of the solutions developed within the project.

The research focus was twofold:

- User's consciousness of data disclosures in consent forms
- User's mental models of the flow of authentication data

More specifically, some auxiliary questions were developed before tests but also between them, namely:

i. Do users find this novel solution manageable and attractive?
ii. Does familiarity with an existing authentication application affect attractiveness?
iii. What are users' preferences for selection of mandatory data in consent dialogues?
iv. Do users pay attention to what data they consent to share? (Here: Do users pay enough attention to what data they consent to share to remember these after a few minutes?)
v. Would an extra confirmation page help them to pay attention?
vi. What are users' mental models and preferences for authentication method?

For this first evaluation of the CREDENTIAL UIs, we have made three user tests with users recruited outside of the project [22]. As performance indicators, task completion and duration were used. The SUS scale, i.e., System Usability Scale [7], was used for participants' estimation of how manageable and attractive they found the prototyped functionality. The SUS scale consists of 10 statements such as:

I think that I would like to use this system frequently.
I found the various functions in this system were well integrated.

The respondents comment by a Likert scale from *Strongly disagree* to *Strongly agree*.

In order to address questions of preferences, experience, and understandability, our usability tests were accompanied by also other questionnaires than the SUS scale.

User Test 1 was designed for the above general goals and participants were by convenience limited to people in Sweden familiar with the Swedish Mobile BankID[2] (17 out of 20 actually used the Mobile BankID and three others used Desktop BankID) which has the authentication capability of the CREDENTIAL *Wallet* app and is generally used.

User Test 2 paralleled Test 1: as the participants in Test 1 in general liked the idea of the *Wallet* and had no problems in using the app, a second test was conducted were all participants had to be non-Swedish and unfamiliar with solutions like Swedish BankID. In order to simplify recruitment, the test was again conducted in and around our university, but with exchange students and visiting parents, and guest professors and newly arrived staff.

User Test 3 was also initiated based on the results from the first test. This time the prototype had an additional screen where participants had to re-confirm which data they authorized the *Wallet* to disclose to a certain receiver. Would this make people more likely to remember disclosure options? (We simplified the test to not include a second task present in Test 1 and 2 where participants were to re-enter a web site and simply sign in, i.e. "authenticate".)

5. User Test Design

5.1. Recruitment of Participants

The participants were recruited to create a sample that was evenly distributed in regard to age and sex – we did not analyse gender or age influence, but we balanced the sample to neutralize such effects. We set out to find participants with a wider age range than the ordinary students, which otherwise constitute a varied recruitment ground, why we included both students and

[2] https://www.bankid.com/en/

personnel of our university in the sample (thus, age was balanced between young adults and mature adults). Most participants were pursuing their higher education or they had their higher education degrees already, and none of them were from Computer Science and Information Systems Departments but two in User Test 2 had basic computer science knowledge. For each test, 20 participants were recruited, equally distributed between men and women. For age, see **Table 1**.

Table 1. Age distribution in the three user tests.

Age Group	Test 1	Test 2	Test 3
Age 20-29	11	13	9
Age 30-39	3	2	1
Age 40-49	5	3	7
Age 50-70	1	2	3
Average	32.2	31.0	36.1

As the most widely used identity providers among Internet users are the single-sign-on solutions provided by social networks, we asked the participants about their familiarity and experience in this context. All of the participants of the first and second user study stated that they had seen social login buttons previously, and more than half of these 40 participants expressed that they use the social login buttons on websites like Spotify, Airbnb, SoundCloud and online TV providers. For the third study, this was not interrogated into as that study was meant to see the effect of the confirmation page on participants' ability to remember what data they ticked for sharing.

5.2. Test Procedure

The user tests were conducted using the screens for authentication and authorization (to data sharing) made interactive in Axure[3] (Test 1 and 3) and Ozlab[4] (Test 2) on an Android mobile phone and a prototype of a fictitious website for which the participants were asked to pretend they wanted to sign up. Procedures were standardized to avoid bias of moderators/interviewers [19]. Participants were given a persona with a pre-defined set of personal information showing up in the CREDENTIAL app; by this, participants could feel secure that they were not compromising their own personal details for taking part in the study. Moreover, it allowed full control of what each participant encountered, avouching a standard experience that can be compared between participants [14]. Thus, users were instructed to perform security-related tasks as the primary (and unreal) task of interaction, but the lack of ecological validity is not severely affecting the comparisons between the different tests as the premises remained the same.

Each session took 20-40 minutes and included an introduction and consent signing, a registration task, a questionnaire about this task including SUS grading, an authentication task (for Test 1 and 2) also followed by the questionnaire, and finally a general questionnaire with demographic data and follow-up questions about using fingerprint and data sharing.

One moderator / interviewer and one note keeper were present during each session.

Besides the SUS form the questionnaires were aimed at investigating if participants paid attention to what personal information they shared with the website from the CREDENTIAL account during the authorization task. We asked about 14 different personal information types including the mandatory and optional information and fingerprint pattern. Moreover, we investigated to what extent the participants noticed the informative links in the UI to receive more information about what happens if they share their data with the website from their CREDENTIAL app. Also, we were interested to see to what extent the participants' mental models and the way they think are compatible with what is happening nowadays when they use an IdP to share their data with other parties.

[3] https://www.axure.com/

[4] https://www.kau.se/en/ozlab/

5.3. Description of Interactive Tasks

For reproductions of the user interfaces, see [8] where data from User Test 1 is reported in detail (the same user interfaces were used in User Test 2 and 3, while an additional re-confirmation page were used in the app mockup in the last test).

Task 1: Participants were asked to sign up (register) to the (fictitious) web site “PhotoHex” using the CREDENTIAL button instead of the Facebook button and manual alternative shown on the same web page. They entered the user name of their persona. Then they had to authenticate to the *Wallet* app on the mobile phone, and received the request from PhotoHex for their name, email address and birthday, and also (optional) profile photo and interest data (in the app mockup “Photography, Feminism, and 23 more” were shown). After ticking at least the mandatory information, they could accept the request by using their fingerprint.

Task 2: Participants in User Test 1 and 2 were furthermore told to imagine re-entering the PhotoHex web site some hours later. Now, they had to sign in (log in) to their PhotoHex account using the CREDENTIAL *Wallet* app. Instead of a data request, the app now showed an authentication request, which was made with a fingerprint click again, just as in the consent in Task 1.

Both tasks ended with the app showing a success screen informing the user to return to the webpage that initiated the request.

6. Results

6.1. Ease of Use and User Experience Metrics

Almost all 60 participants managed to complete the tasks given to them. In each of the first two user tests, two and three persons had problem in Task1 to find out how to continue; in the third test there was a “Continue” button which no one missed. As for completion time, the tests with two tasks showed a clear decrease in completion time between the two tasks (on average from 100s to 57s in Test 1; from 62s to 44s in Test 2). During the study some participants mentioned that it is always hard when one is using an app for the first time. This show that users are expecting some initial efforts in order to be soon rewarded by a more fluent and quicker use.

Also the SUS values in **Table 2** reflect this: compared to two of the grading scales presented in [3] that are based on a high number of answers from different studies, the CREDENTIAL *Wallet* app in its mockup dress was perceived as “Good” which definitively is within the “Acceptable” range. However, other responses showed participants confusing data disclosed with other data. In fact, it was because already User Test 1 demonstrated these good SUS values, that the evaluation went further to not only balancing BankID users/non-users (Test 2) but also to see the effects of a separate confirmation screen where selected data were listed (Test 3).

Table 2. SUS scores after each task.

Test	Test 1		Test 2		Test 3
Task	**Task 1**	**Task 2**	**Task 1**	**Task 2**	**Task 1**
Aver.	75.88	81.98	78.19	85.86	78.25
St.dev.	11.74	11.69	13.98	12.29	18.24
Min.	47.50	57.50	50.00	47.50	35
Max.	95.13	97.50	100.0	100.0	100

6.2. Expressed Preference for Data Selection

When asked in the questionnaire after the first task whether they preferred to have to manually tick checkboxes (“i.e., an interface as in the current task”) or to have all the mandatory information selected by default, a majority of the participants in the three tests wanted to select the mandatory information manually (15 in Test 1, and 13 in each of Test 2 and 3). This puts the time and SUS values in perspective: speed is not the top priority for many users.

6.3. Recall of Consents to Data Disclosure

In line with other studies (Section 3) we found that people do not pay sufficiently attention to details of data releases and settings for disclosure policies.

Only two of the 60 participants could correctly identify in a list of fourteen items what information was shared from the CREDENTIAL app with the website. In the UI prototype of the app, the mandatory information were full name, birthday and email address. All participants accepted the request in the authorization task so all of them shared the information with the website (PhotoHex in the study). However, in Test 1 and 2, a few participants ticked "no" when asked to state if they shared the full name, email address, and birthday. Interestingly, some did not seem to notice that when they shared their birthdays, they implicitly shared their ages: some participants did not have consistent answers for their birthdates and ages. A handful in each test ticked "no" for age but many, especially in Test 2, showed uncertainty and chose "not sure". In Test 1, two participants, who were not sure if the birth date was shared, ticked that the age was shared.

For optional information, some selected to share all the optional information when doing the authorization tasks but not all of them remembered properly what they shared. In Test 2, 12 people shared "interest" but only 3 remembered having done so. Also, some of the people who did *not* share the optional information did not remember correctly if they shared a photo or their interests. Photo sharing was more remembered than sharing interests. For instance, in Test 1, regarding the photo, just two participants answered incorrectly but for sharing the interests five participants were not successful to correctly answer the questions. As an example of test results, **Table 3** shows the numbers of participants in User Test 1 who ticked YES, NO or NOT SURE for each piece of information.

In the list of information types we included items that were not on the authorization screen of the app, such as hometown, credit card number and educational background. Interestingly, some people incorrectly stated that some of this information was shared with the website. Among all the information being asked that was shared or not, people were most successful at remembering about sharing the photo and were most unsuccessful at understanding if they shared the fingerprint with the website.

Would people fare better if a separate confirmation page was shown in order to re-inforce the impression of the selected data? User Test 3 was designed for this comparison. Before calculating ratios of successful mentioning of the correct pieces of information, it is important to realise that the ratios are dependent how many alternatives our questionnaire contained. To better report the awareness of participants of personal data being shared we measured the relevance of responses using *precision* and *recall*. Ronen et al. (2013) used these measurements to report people's awareness of data exposures when using a federated identity provider to sign in to websites. Ronen and co-workers defined the precision (P) as the ratio between the number of data types a participant named correctly and the number of data types the participant named, and defined the Recall (R) as the ratio between the number of data types a participant named correctly and the number of data types that were actually transferred to the website.

In order to better see to what degree of certainty people were, we let them tick either "yes", "no", "not sure" for *every* data type. The self-expressed uncertainty averaged to 21%, 30% and 22% in the three tests, i.e. rather high which is interesting and speaks for solutions like the Data Track found in [14]. The option to tick "not sure" instead of what sometimes might have been "yes" made us calculate precision in two ways. Method One was as Ronen by interpreting "not sure" as "no"; thus only "yes" answers were counted. Method Two calculated for precision all correct "yes" *and* "no" on top of *all* "yes" and "no" per person.

In **Table 4**, the related average, standard deviation, minimum and maximum for recall and precision are reported with bracket around Method One values. "Age" was removed from the table as some respondents might have thought of explicit consent rather than what "Birthday"

Table 3. Distribution of YES, NO and NOT SURE answers in User Test 1.

Information	YES		NO		NOT SURE	Sum
	Correct	Incorrect	Correct	Incorrect		
Your post address	N/A	4	9	N/A	7	20
Your full name*	15	N/A	N/A	4	1	20
Your email address*	17	N/A	N/A	1	2	20
Your educational background	N/A	0	16	N/A	3	19†
Your fingerprint pattern	N/A	17	2	N/A	1	20
Your home town	N/A	1	11	N/A	8	20
The city in which you live	N/A	1	11	N/A	8	20
Your birthday*	14	N/A	N/A	3	3	20
A photo of yours**	6	2	12	0	0	20
Your phone number	N/A	2	12	N/A	5	19†
Your interests**	4	2	10	3	1	20
Your credit card number	N/A	1	16	N/A	3	20
Your age (* by implication)	14	N/A	N/A	3	3	20
Your mobile device model	N/A	1	10	N/A	9	20

* Mandatory information ("Birthday" included year number in the mockup).

** These were presented in the UI as optional choices. Six participants chose to share both photos and interests and one participant shared only the interests. The remaining thirteen selected just the mandatory data.

† One participant forgot to answer.

N/A Not applicable.

implies. (Two participants in Test 1 did not answer to all the questions for data being shared; their responses are excluded from these calculations.) Although the average numbers are not very low, the distribution figures show that many had a rather limited awareness of the data types being transferred. The average recall rate for the 38 participants included in **Table 4** for Test 1 and 2 is 0.74, while Test 3, with its extra confirmation page, had a notably higher Recall rate. Recall rates did not correlate with age (R^2 = 0.006 counted on all 58 participants).

The prototype contained the very common multi-layered approach to present the necessary data for informed consent (for instance, a multi-layered approach is argued by the Art. 29 Working Party [2]). The only participant who actually opened the full information did not show any remembrance of this when answering the questionnaire afterward.

Table 4. Precision and recall values excluding "Age" (Method One values in brackets)

Test	Test 1		Test 2		Test 3	
	Precision	Recall	Precision	Recall	Precision	Recall
Average	0.80 (.68)	0.79	0.75 (.72)	0.70	0.80 (.68)	0.86
St.dev.	0.14	0.22	0.18	0.17	0.14	0.18
Min.	0.46	0.33	0.33	0.40	0.58	0.50
Max.	1.00	1.00	1.00	1.00	1.00	1.00

6.4. Mental Model and Preference for Authentication via an App

In User Test 1, most participants seemed to understand that they used the CREDENTIAL *Wallet* mockup to authenticate, but this could depend on their previous experience with a similar solution as explained in Section 4. That would bias the conclusion about the acceptability of this technology if not other subjects, for instance from other countries, are included. The User Test 2, however, showed that inexperience of this type of authentication did not affect the SUS values (cf. **Table 2**). User Test 3 merely confirmed previous results.

Nonetheless, some participants were worried about the security and thereby also the privacy because they could not really figure out how it worked. Moreover, the single point of failure, as it is called in the literature, that is if the one and only IdP malfunctions, the user is at loss. This could be a cause to the worries but in fact only a handful participants mentioned this: either that hacking gives wide access to users' info or that dropping the phone inhibits users' access to the CREDENTIAL app.

Using the fingerprint for giving consent (at the same screen requesting for the personal information in authorization dialogue, for example) was from a simple usability perspective not a problem, but from the post-task questions it is obvious that most of the participants did not know that the pattern is processed only locally. **Table 3** shows that in User Test 1, almost all participants (17 out of 20) thought the fingerprint pattern was forwarded to the web service requesting some data. In Test 2, seven people thought so. The mistaken view of authentication data disclosure is reflected in the answers to the last questionnaire in the user tests which included a question "Where do you think that your fingerprint is stored and processed?" **Table 5** lists the number of respondents that selected one or several of three alternatives.

Table 5. Perceived storage and processing location for fingerprint.

Where is the fingerprint stored?	Test 1	Test 2	Test 3
Mobile device (correct answer)	4	9	5
CREDENTIAL app	11	6	8
Mobile and app	2	2	5
App and website	2	0	0
Mobile, app, website	1	2	1
Website	0	1	1

On a question "What do you think about using your fingerprint to agree to share your personal data?", four participants in User Test 1 and thirteen in Test 2 expressed that they did not like to use their fingerprints and said they prefer passwords. One explained that it was not pleasant to give the consent with the fingerprint because it might be saved and used in unwanted ways. Others thought it could be hacked.

7. Discussion

Like our study, Bauer et al. [4] demonstrated that participants' precise understanding of what is sent is not significantly affected by the consent dialogues (instead, as that study indicates, it is affected by their privacy concern level and that they have some preconceptions about what is going to be sent). Egelman's study [10] also showed that participants did not read the authorization dialogues and they did not pay attention to the details of the dialogues during the test.

Interestingly, the extra step of being confronted with a confirmation screen, which simply repeated what had already been agreed to, made the recall value rise in Test 3 as well as the minimum values. The recall rate of 0.86 can be compared with 0.74 which is the average for the 38 persons from Test 1 and 2 included in **Table 4**. As neither of these two tests included the confirmation screen, it seems that there are reasons to evaluate the impact of confirmation screens also in future studies, and possibly leave it as the default option when requests concern several data types, even if users may be able to switch it off.

As mentioned in Section 6.2, when asked in the questionnaire after the first task, a majority of the participants in the three tests wanted to select the mandatory information manually. This can be a way of slowing users down and thus make them reflect more. We do not mean that people should feel obstructed, but the design should make people more actively see or choos the data to be disclosed and the conditions for the disclosure. Nevertheless, the results in Section 6.3 shows that in spite of the selection that our test participants had to do, they had in fact rather vague ideas of what they had “consented” to share with PhotoHex. This principle could be pursued further (as we will do) to see if more explicit action of selections, such as drag-and-drop, would instill a better impression on users’ short-term memory.

Even if the data in **Table 5** is not consistent with the numbers on the “fingerprint” row of **Table 3** (which covers Test 1, but similar results were obtained in the other tests), it is anyhow obvious that many people do not understand how “local” this authentication is. The results concerning the fingerprint are in agreement with Sun’s et al. [24] and Arianezhad’s et al. [1] studies which show people have incorrect security mental models when using IdPs to sign up. But more investigations are needed to see if the problem is related to using the fingerprint for giving consent or in general it relates to the lack of knowledge about how fingerprint works on the mobile phones. There are a few works on users’ experiences, attitudes, and adoption decisions scanning the fingerprint to unlock their mobile devices [6]. There seem to be no works on users’ understanding and perceptions of fingerprint when it is used in the context of identity providers to give consent or on users’ opinions about who can access their fingerprint patterns when it is used to unlock an app on their phones.

The Swedish respondents in Test 1 and 3 seem to be more ill-informed than the foreign visitors to Sweden in Test 2.[5] This was not really why we conducted the tests, but the results raise intriguing questions. Can the cause be the Mobile BankID? Do the users trust the app that the banks trust? Whatever the cause for the misunderstanding, the implication can be seen in answers to other questions. For example, the data briefly accounted in last paragraph of Section 6.4 point to a hesitance to embrace the fingerprint method. There is an immediate remedy: in order to follow the Android standard, the UI should contain alternatives, for instance a pin code. Of course, this can be elaborated further: will users think that the pin code is sent to web services?

While comprehension is most important, ease of use is of interest as it will affect adoption [9]. Ruoti et al. [23] have studied the usability of seven authentication systems including federated single sign-on systems like Google OAuth 2.0 and Facebook Connect. In their study, they defined three groups of authentication techniques and the method which won in each group was also compared and tested for SUS score with the winners of the two other groups. In the final analysis, Google OAuth was the winner between different technologies with the SUS score of 75. Our average SUS scores for the CREDENTIAL app, both for authentication and authorization task, align with Ruoti’s study.

8. Conclusions: Way Forward for Identity Access Management GUI Design

In sum, further evaluation is required concerning how people would understand what data are needed, how to select these, and what data are actually sent. Our results suggest that a confirmation screen can be a default option in authorization (consent) dialogues.

The desire expressed by many to select the mandatory information manually can be explored as a way of slowing users down, and make them reflect more.

Also, user evaluation should be wanted of more compound data disclosure from cloud-based wallets. Projections as in the CREDENTIAL project [8] to use secure data sharing also for authorization requests concerning file-based data (documents such as signed documents and

[5] Despite the non-random sampling, attempting an hypothesis test comparing proportions π of persons answering correctly with H_0: $\pi_{\text{BankID}} - \pi_{\text{no}} = 0$ and H_1: $\pi_{\text{BankID}} - \pi_{\text{no}} < 0$, one gets at a significance level $p < 0.05$ that the proportion of non-BankID users answering only “Mobile Device” is larger than the proportion of BankID users doing that. [18]

health records), will need careful design to avoid that users select wrong files, especially if data are not directly visible in an app for the identity access management.

Fears of single point of failure and other misconceptions are clearly a problem. The comparison of BankID users with others indicates that trust is trust in an actor (bank), not in the system, which can possibly lead to a skewed mental model of what should be trusted. This calls for educating the general public, not only to design user interfaces.

Acknowledgements

This work was made within the H2020 project CREDENTIAL (grant agreement No 653454). The authors want to thank Caroline Kayser and Manuel Gawert who conducted the data collection for the second user test (Farzaneh Karegar and Daniel Lindegren made the original prototypes and conducted the first user test, and Lindegren conducted the third user test), Charlotte Bäccman who reviewed a project report and helped with an initial set of participants, and Dan Larsson who discussed comparing proportions.

References

1. Arianezhad, M., Jean Camp, L., Kelley, T., Stebila, D.: Comparative Eye Tracking of Experts and Novices in Web Single Sign-On. In: Proceedings of the Third ACM Conference on Data and Application Security and Privacy, pp. 105-116. ACM, (2013)
2. Art. 29 Data Protection Working Party: Opinion 10/2004 on More Harmonised Information Provisions. (November 25th, 2004). 11987/04/EN WP 100. European Commission (2004)
3. Bangor, A., Kortum, P., Miller, J.: Determining what individual SUS scores mean: adding an adjective rating scale. Journal of usability studies, 4(3), 114-123 (2009)
4. Bauer, L., Bravo-Lillo, C., Fragkaki, E., Melicher, W.: A Comparison of Users' Perceptions of and Willingness to Use Google, Facebook, and Google+ Single-Sign-On Functionality. In: Proceedings of the 2013 ACM Workshop on Digital Identity Management, pp. 25-36. ACM (2013)
5. Besmer, A., Lipford, A. H.: Users' (Mis)conceptions of Social Applications. In: Proceedings of Graphics Interface 2010, pp. 63-70. Canadian Information Processing Society (2010)
6. Bhagavatula, C., Ur, B., Iacovino, K., Kywe, S. M., Cranor, L. F., & Savvides, M.: Biometric Authentication on iPhone and Android: Usability, Perceptions, and Influences on Adoption. In: Proc. USEC 2015 (2015)
7. Brooke, J.: SUS: a "quick and dirty‟ usability scale. In P.W.Jordan, B. Thomas, B.A. Weerdmeester, and I.L. McClelland (Eds.) Usability Evaluation in Industr, pp. 189-194. London: Taylor and Francis (1996)
8. D3.1 UI prototypes V1. Deliverable from the project CREDENTIAL (2017). Available at: credential.eu/publications/deliverables/d3-1-ui-prototypes-v1/
9. Davis, F. D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13(3), 319-339 (1989)
10. Egelman, S.: My Profile is My Password, Verify Me!: the Privacy/Convenience Tradeoff of Facebook Connect. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2369-2378. ACM (2013)
11. Hörandner, F., Krenn, S., Migliavacca, A., Thiemer, F., Zwattendorfer, B.: CREDENTIAL: A Framework for Privacy-Preserving Cloud-Based Data Sharing. In Availability, Reliability and Trust – ARES 2016. *DOI 10.1109/ARES.2016.79*. (2016)
12. Javed, Y., Shehab. M.: Investigating the Animation of Application Permission Dialogs: A Case Study of Facebook. In: International Workshop on Data Privacy Management, pp. 146-162. Springer International Publishing (2016)

13. Javed, Y., Shehab. M.: Look Before You Authorize: Using Eye-Tracking To Enforce User Attention Towards Application Permissions. Proceedings on Privacy Enhancing Technologies Vol. 2017(2), pp. 23-37 (2017)
14. Karegar, F., Pulls, T., Fischer-Hübner, S.: Visualizing Exports of Personal Data by excercising the right of data portability in the Data Track – Are people ready for this? In Lehman et al. (eds.): Privacy and Identity Management. Facing up to Next Steps, pp. 164.181. Springer (2016)
15. Karegar, F., Striecks, Ch., Krenn, S., Hörandner, F., Lorünser, T., Fischer-Hübner, S.: Opportunities and Challenges of CREDENTIAL. Towards a Metadata-Privacy Respecting Identity Provider. A. Lehmann et al. (Eds.): Privacy and Identity 2016, IFIP AICT 498, pp. 76–91. Springer (2016)
16. Kostopoulos, A., Sfakianakis, E., Chochliouros, I., Pettersson, J.S., Krenn, S., Tesfay, W., Migliavacca, A., Hörandner, F.: Towards the Adoption of Secure Cloud Identity Services. Accepted for SECPID / ARES 2017 (2017)
17. Liccardi, I., Pato, J., Weitzner, D. J., Abelson, H., De Roure, D.: No Technical Understanding Required: Helping Users Make Informed Choices about Access to Their Personal Data. In: Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, pp. 140-150. ICST ((2014)
18. Lind, D., Marchal, W., Wathen, S.: Statistical Techniques in Business & Economics, 17th ed., pp. 550ff on "Two-sample tests about proportions." McGraw-Hill (2017)
19. Onwuegbuzie, A. J., Leech, N. L.: Validity and Qualitative Research: An Oxymoron? Quality & Quantity 41(2), 233-249 (2007)
20. Robinson, N., Bonneau. J.: Cognitive Disconnect: Understanding Facebook Connect Login Permissions. In: Proceedings of the Second ACM Conference on Online Social Networks, pp. 247-258. ACM (2014)
21. Ronen, S., Riva, O., Johnson, M., Thompson, D.: Taking data exposure into account: How does it affect the choice of sign-in accounts? Proc. of the SIGCHI Conference on Human Factors in Computing Systems CHI '13, ACM, pp. 3423–3426 (2013)
22. Rubin, J., Chisnell, D.: Handbook of usability testing: how to plan, design and conduct effective tests. John Wiley & Sons (2008)
23. Ruoti, S., Roberts, B. Seamons, K.: Authentication melee: A usability analysis of seven web authentication systems, Proceedings of the 24th International Conference on World Wide Web (Republic and Canton of Geneva, Switzerland), WWW '15, International World Wide Web Conferences Steering Committee, pp. 916–926 (2015)
24. Sun, S.-T., Pospisil, E., Muslukhov, I., Dindar, N., Hawkey, K., Beznosov, K.: What Makes Users Refuse Web Single Sign-On?: An Empirical Investigation of OpenID. In: Proceedings of the Seventh Symposium on Usable Privacy and Security, pp. 4:1-20. ACM (2011)
25. Van Kleek, Max, Ilaria Liccardi, Reuben Binns, Jun Zhao, Daniel J. Weitzner, and Nigel Shadbolt.: Better the Devil You Know: Exposing the Data Sharing Practices of Smartphone Apps. Forthcoming in: CHI'17. ACM (2017)
26. Vapen, A., Carlsson, N., Mahanti, A., Shahmehri, N.: Information Sharing and User Privacy in the Third-party Identity Management Landscape. In: IFIP International Information Security Conference, pp. 174-188. Springer International (2015)
27. Wang, N., Grossklags, J., Xu, H.: An Online Experiment of Privacy Authorization Dialogues for Social Applications. In: Proceedings of the 2013 Conference on Computer Supported Cooperative Work, pp. 261-272. ACM (2013)
28. Wang, N., Xu, H., and Grossklags, J.: Third-party Apps on Facebook: Privacy and the Illusion of Control. In: Proceedings of the 5th ACM Symposium on Computer human Interaction for Management of Information Technology, p. 4. ACM (2011)