D8.3

Pilot case evaluation report and lessons learned – v1

v 1.1 / 2017-02-21

Rob Heyman (iMinds), Ali Padyab (LTU), Paulien Coppens (iMinds), Tom Seymoens (iMinds) Laurence Claeys (iMinds), Jo Pierson (iMinds), Marita Holst (LTU)

Abstract
This deliverable presents the design and findings of the first round of pilots conducted within USEMP. It is a follow up from D8.4 (“Set up of experiments in the living labs”) We aim at evaluating awareness, ease of use, capabilities and practices among the end-users involved in the pilot case of DataBait.

Project acronym	USEMP
Full title	User Empowerment for Enhanced Online Presence Management
Grant agreement number	611596
Funding scheme	Specific Targeted Research Project (STREP)
Work program topic	Objective ICT-2013.1.7 Future Internet Research Experimentation
Project start date	2013-10-01
Project Duration	36 months

Workpackage 8	
Deliverable lead org.	iMinds
Deliverable type	Report
Authors	Rob Heyman (iMinds), Ali Padyab (LTU), Paulien Coppens (iMinds), Tom Seymoens (iMinds) Laurence Claeys (iMinds), Jo Pierson (iMinds), Marita Holst (LTU)
Reviewers	Georgios Petkos (CERTH), Alexandru Ginsca (CEA)
Version	1.1
Status	Final
Dissemination level	PU: Public
Due date	2016-02-29
Delivery date	2016-03-31 (revised 2017-02-21 after reviewers' comments)

Table of Contents

1 Introduction

This deliverable presents the design and findings of the first round of pilots conducted within USEMP. It is a follow up from D8.4 (“Set up of experiments in the living labs”). We aim at evaluating **awareness, ease of use, capabilities and practices** among the end-users involved in the pilot case of DataBait. The deliverable is structured as follows: In the rest of this introductory section we we first frame the objectives and then present the research outset. Subsequently, in section 2 we look into the pilot methodology and living lab tasks and then, on the final section we report on our findings.

1.1 Objectives and research questions

This deliverable is relevant to task 8.4, for which the DoW states the following objectives: “In this task, the pilot case results will be evaluated and documented in two iterations, firstly the pre-pilot and secondly the pilot. An evaluation plan which will be implemented in both Living Labs will be developed focusing on evaluating **awareness, ease of use, capabilities and practices** among the end-users involved in the pilot case. Lessons learned will be shared for the further development of the USEMP platform. For this task we apply the Living Lab method for investigating the prototypes (or proxy technologies) based on an iterative research process.” (USEMP, 2015, p. 35)

The results of task 8.4, the pilot evaluations, will be reported in two deliverables: D8.3 (this document) and D8.5 with the respective deadlines of Month 28 (January 2016) and Month 36 (September 2016). In the first pilot we evaluate identified needs and what should be added in terms of actionable information for users to change their personal information flows if necessary. In the second pilot we evaluate the implemented changes suggested in the first pilot.

The needs we evaluate are build on insights from two sources. First there are the insights of D8.2 (that reported on the evaluation of the pre-pilots). Second, iMinds (2016) summarised a round of feedback from the consortium based on interactions at conferences. Both are outlined below:

D8.2 needs:

- *Increase awareness* about information disclosure and use of personal data disclosed on Facebook for third parties for the following types of information:
    - Location information
    - Image information
- Achieve this in a *user-friendly* manner
- Do this in an *informed way* (DLA)

Other needs identified by the consortium:

- Extra information should be given on how online behaviour can result in *institutional privacy* issues.
- More information should be provided on what revealed *third parties* may do with gathered data

- Means to *control* information disclosure should be communicated to users (link with provision of privacy-enhancing guidelines)

We now translate these objectives into research questions for this deliverable:

Improve the user experience of DataBait as a tool to increase awareness for the following information:

- Images shared on Facebook
- Location information shared on Facebook
- A view of the influence you have over your Facebook friends
- How users are tracked through third party cookies

This results in the following research question:

Where should we further improve the ease of use for the different DataBait tools?

Next we want to measure the effectiveness of DataBait as a Privacy Enhancing Technology (PET) to increase the awareness of users with regard to their information disclosure to third parties. We also want to inform users about what these third parties may or may not do with their personal information.

Is DataBait a PET that can increase user awareness with regard to information disclosed on social media and through cookies?

Lastly, in order to be a useful PET, this tool needs to be able to help users inform themselves so that they can revise their attitude towards this disclosure. In order to achieve the latter, raising awareness alone is not enough, the usefulness of DataBait as a PET also consists of pointing out solutions to particular problematic types of disclosure. While it is difficult to measure how users' behaviour changes while they use this tool, it should be possible to evaluate if the right pointers are provided to aid users if they feel like changing an unsatisfactory situation.

Does DataBait offer users actionable information to revise their online and social media disclosure?

1.2 Living lab outset

As said, we have to report on two pilots. The first pilot consists of the version of DataBait that went live on January 15, 2016. The next pilot consists of the changes we have implemented based on the result of this current deliverable which will be evaluated in D8.5.

The needs identified by iMinds' *Proposed DataBait Updates* (2016) are primarily concerned with making the increase in awareness more useful for users. By this we mean that the increase of awareness of what is disclosed to third parties should not be isolated from knowing what the next steps are. These are assessing what should be shared and what not. But this decision itself is only meaningful if users are presented with a choice and means to act upon this choice: "Means for control taking measures should be handed over to the users to counter resignation" (iMinds, 2016, p. 14)

We therefore propose to put different accents in the two pilots. The first pilot is small scale and exploratory to answer the third research question: “What should be added to DataBait to offer users actionable information with regard to their social media and online information disclosure?” The first pilot emphasizes this question more, but it will also look into the other research questions, that is, it will also look into usability and increase of awareness as researched in D8.2, but with a living lab method.

In the latter pilot, we will evaluate the newly found means to render awareness meaningful to change personal data flows together with the evaluation of other changes to ease of use and increasing awareness of users.

1.2.1 Changes with regard to the previous approach

In deliverable 8.4 we outlined parallel tracks, one being qualitative and the other quantitative. We want to revise this approach since both pilots will be a mixture of both. Instead, we propose to define the two pilots as separate cases with different scopes. We propose a smaller scale first pilot to further explore what users require in terms of information to make DataBait more useful. Then, this addition of information is further evaluated in the last version of the DataBait pilot.

We opt for a small-scale first pilot in order to have face-to-face interaction with respondents and to question them more in-depth about the information they would need to make DataBait a more useful tool. This means asking them to keep notes as to when they feel resigning DataBait or when they are blocked to do anything useful with the information DataBait provides them. By letting them note the issues they encounter with regard to their increase in awareness, we can see what information they require and when. We can then offer extra information and add it to DataBait.

2 Method: Living lab

The general aim of the first living lab pilot is to test the changes implemented by January 15 by our technical partners. Secondly, we want to engage our respondents by asking them what information they require to make the information provided by DataBait actionable. This means asking them through open questions how users felt about the information we provided them and how they wished to proceed further with that information. Next we we will suggest partners to find solutions for the problems respondents made us aware of. This living lab approach is structured as shown in Figure 1. We will further describe this outset in a chronological matter as outlined in Figure 1.

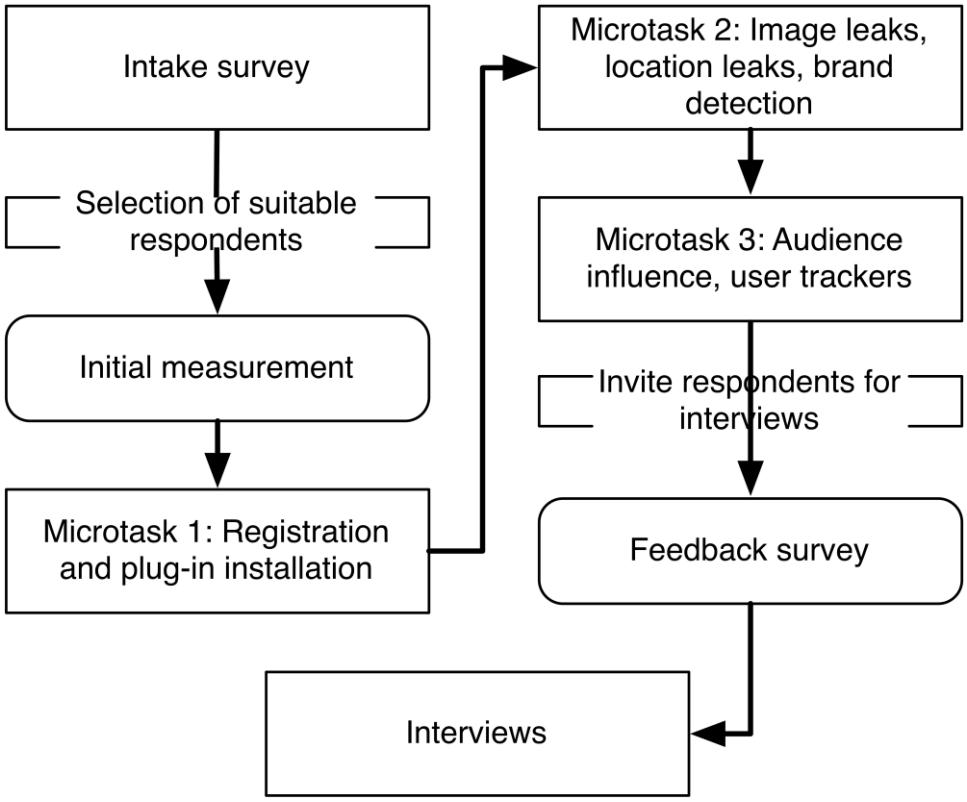

Figure 1: Living lab steps

And this resulted in the following time plan:

Proposed Date	**Action**	**Real timing**
28-Jan	Intake survey	28-Jan
02-Feb	Initial measurement and microtask 1	04-Feb
05-Feb	Microtask 2	11-Feb
08-Feb	Microtask 3	18-Feb
12-Feb	Feedback survey	18-Feb

Table 1 Proposed and real timing

The proposed timing was not followed due to technical malfunctions with the server as outlined in Table 1. Interviews are not mentioned in this table because they were taken on different moments depending on respondents' availability. All interviews were taken after the feedback. The timing of these interviews can be found in the annex[1].

2.1 Participants' selection criteria and intake survey

In this section we define our sampling criteria. We use a purposeful sampling procedure with two sampling criteria: maximum variation and phenomenal variation. Following Pierson and Lievens (2005) we use the principle of maximum variation to select participants: we want the variables age, gender and education to vary as much as possible for analytical purposes. We do this in order to have many different instances[2] and to understand the use of the DataBait tool for people with different disclosure practices.

Secondly, we have our phenomenal variation sampling criteria which refer to those variables required for our research:

- Have a Facebook account
- Have internet access at home
- Be willing to share Facebook data with our DataBait tool
- Understand: Dutch, Swedish or English
- Have to use a browser compatible with our tool (Chrome or Firefox)

The choice for maximum variation criteria is based on previous privacy research where the following variables were tied to different disclosure practices. Recent PEW (Rainie, 2015) and Eurobarometer studies (TNS opinion & social, 2015) reported that the following socio-demographics influence whether or not people have changed their privacy settings, which is a good indicator to show that some users have acted upon something to change their disclosure[3].In particular, it is reported that younger respondents are unsurprisingly more likely to have tried changing their privacy settings: 69% of 15-24 year-olds have done this, compared with only 36% of people aged 55 and over. Individuals with a higher level of education are also more likely to have changed their privacy settings: 64% of people who finished their education aged 20 or over have done this, compared to 36% of respondents who left school aged 15 or under. Among the different occupational groups, managers (68%) are the most likely to have tried to change their privacy settings, whereas retired people (33%) and house persons (47%) are the least likely to have done so. (TNS opinion & social, 2015, p. 93)

The Eurobarometer report did not refer to changes in gender but PEW did: "Women who use SNS are more likely than men to set the highest restrictions (67% vs. 48%)." (Madden, 2012) This was also noted for the action of defriending people on Facebook: "Some 67% of women who maintain a profile say they have deleted people from their network, compared with 58%

[1] See annex, p. 46.
[2] Note that this variation is used for external validity, to make sure that our findings apply to other people with different backgrounds.
[3] As noted in PEW, this does not mean that users who have accessed their privacy settings have more or better privacy: Similarly, as Boyd and Hargittai (2010) have noted: "Familiarity with and decisions to adjust privacy settings are not equivalent to actual privacy protection." See: "Facebook privacy settings: Who cares?" Full article available at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3086/2589.

of men. Likewise, young adults are more active unfrienders when compared with older users." (Madden, 2012)

As a result, we will seek maximum variation for age, gender and education. These seem to be variables that influence the interaction with privacy settings or other means that negotiate boundaries with regard to privacy.

Maximum variance variables:

- Age 50% younger than 24 years, 50% older than 24
- Equal gender distribution
- Ideally a spread of education levels

These variables were then translated in questions that can be found in Annex 3.

2.1.1 Method to gather respondents for the Living labs

We gathered respondents through the respective living labs of each involved partner and made sure that we had a maximum variation for the variables outlined above. Both living labs targeted these variables. The total number of respondents depended on the number of in-depth interviews. We wanted to have at least 4 interviews per living lab and have recruited more living lab participants to ensure we could select 4 respondents.

2.1.2 Invitation

Before we sent out the intake survey, we send out an invitation (Annex 1), which informed respondents that USEMP is a project that aims to increase awareness about:

- What respondents share via Facebook
- What respondents implicitly share via Facebook through logging and inference
- With whom they automatically share data through cookies

The iMinds Living Lab added an incentive of 10 euros for an online shopping site if respondents complete all the tasks. Another incentive of 10 euros was provided for the 8 respondents who participated in the interviews. This resulted in 45 selected participants of which 28 finished the online pilot and 8 participated in the interviews.

The participants from Sweden were selected through the **LTU living lab** database and from a public announcement to the university web site. A total of 18 participants were able to finish the test within a period of three weeks. The participants for the interviews were selected among the interested participants. 4 interested participants were invited to the interviews. All sessions were audio recorded with the interviewee's consent. In return, LTU gave them a gift card of value 150SEK.

2.2 Initial measurement

We identified the following objectives for the initial measurement:

- Find out what PETs they have been using the last year
- Find out how confident they are in using these PETs and other means to manage their information disclosure towards third parties on Facebook (with a focus on the things we can visualise in DataBait)
- Attitudes towards the disclosure of different types of information (disclosure dimensions):
    - Sensitivity
    - Awareness of sharing

  - Disclosure preference
- Are respondents able to understand what they are disclosing to third parties?
  - Is this actionable information or not?
  - What were past decisions with regard to third party disclosure?
- What is the general privacy concern of our respondents?

This leads to the following survey questions.

2.2.1 Global Information Privacy Concern

The following 6 questions were asked to the users. Responses were in seven-point scales anchored with “strongly disagree” and “strongly agree”. The questions are originally found in Smith et al. (1996), but elements have been updated to reflect an online environment by Malhotra et al. (2004):

(1) All things considered, the Internet would cause serious privacy problems.

(2) Compared to others, I am more sensitive about the way online companies handle my personal information.

(3) To me, it is the most important thing to keep my privacy intact from online companies.

(4) I believe other people are too much concerned with online privacy issues.

(5) Compared with other subjects on my mind, personal privacy is very important.

(6) I am concerned about threats to my personal privacy today.

We have chosen this scale because we have used it in the past and want a scale that is general, but short and applied to internet environments.

2.2.2 Are respondents able to understand what they are disclosing to third parties?

This could be cover by the PEW (“Topline,” 2015) survey questions. These questions reflected on the decision making process with regard to disclosing personal information to third parties. It maps the effort required to understand the information provided, whether the information is confusing, how confident people are in understanding what was shared, and whether they had time to act on the provided information or not. Lastly, it probes with an open question to find an example of when respondents thought about disclosure to a third party.

The following questions were first asked, respondents can answer with yes, no or abstain from answering.

At any point, have you felt:

- Discouraged with the amount of effort needed to understand what would be done with your data?
- Confused by the information provided in a privacy policy?
- Confident that you understood what would be done with your data?
- Impatient because you wanted to learn more but needed to make a decision right away?

Next, PEW followed up with an open question that may also help in our case:

- Could you please give us an example of a recent time where you considered whether or not to share information about yourself in return for something? What did you decide and why?

These open-ended questions aided in understanding the trade-offs respondents face when disclosing personal information towards third parties. This is important because it presents a context that usually shows that privacy is either outweighed by a benefit or outweighs a benefit. For instance, some possible answers could be:

- An Internet site was asking for control of my computer's camera. I refused. They would have access to my personal space.
- My car insurance offered discounts for the ability to monitor my driving. I felt it was too invasive.
- Posting of resume online for job search. Needed for visibility to employers, but also risked being available for marketers, etc. My choice was between limited exposure of information and consequently limited [job-hiring] potential vs. full exposure and greater potential. I chose to post. (Rainie, 2015)

2.2.3 Experience with regard to possible solutions

This was covered with the following questions.

Please indicate whether you performed this action last year and how confident you were in performing this task:

- Using an app to block online third party trackers
- Using an ad blocker to stop seeing advertisements
- Visiting an opt-out platform to stop third party behavioural advertising (youronlinechoices.eu)
- Using an advertising preferences menu such as the likes offered by Google or Facebook to change or delete my ad preferences profile
- Deleting content I put on social media
- Deleting content about you someone else put on social media
- Erasing search results from Google that mention my name
- I can change the audience of Facebook content
- I know how to delete my Facebook account
- I have used the do-not-track feature in my browser settings
- I have deleted cookies through my browser

Responses were in the form of yes or no (for the first part, asking whether the task was performed) and a 7 point likert scale ranging from not confident at all to very confident (for the second part, asking about the confidence of the users in performing the task).

The list of solutions is based on the options offered by Facebook, other online PETs and other organisations such as Google and online advertisers.

This question situates prior use of PETs before DataBait. When we ask these questions again during the feedback survey, we can see if more PETs were used and in interviews we can ask if this had anything to do with DataBait. As such, we can research if the information provided by DataBait helps in changing online disclosure to third parties.

2.2.4 Disclosure dimensions

We aimed to investigate the perceptions of users about a) the sensitivity, b) awareness of sharing and c) disclosure preference for different disclosure dimensions. Please note that all identified disclosure dimensions (please see D6.1 and D6.4) were considered with different questions. We have asked the following three questions for each dimension (the listed questions are about the "Profession and income level" dimension, but we had similar questions for the other dimensions):

- How sensitive do you find the information you have to reveal about your profession and income level? (7 point likert (1= not sensitive at all, 7= very sensitive))
- Do you think the information on your Facebook profile reveals this information? Either because you yourself have put it online, or it could be inferred from a combination of posts.
- How important is it for you that this type of information about you remains private? (1= not important, 7= very important)

2.3 Micro tasks

The following micro tasks were presented to the users: registration, installation of the plug-in, image leaks, location leaks, friend influencer and third party tracking. The overall approach is similar to the pre-pilot (D8.2); each feature was evaluated by respondents. We have added questions with regard to how users feel about disclosing information in this specific case and we have added open questions with regard to what users wished to change after seeing their information disclosure.

In each task we asked users to go to the DataBait component we want to analyse. Next we asked them to use it and answer a couple of questions. After the micro tasks were completed, which takes a little bit more than one week, we coupled back their answers in a F2F interviews with a subset of participants.

For each task, we asked users about the usability of the application, and how this new information disclosure information affected them with the same question of PEW but adapted to specific situations. So, for example, after completing the registration process, the following questions were presented to the users:

- How easy or difficult was this task? (1 = very hard, 7 = very easy)
- How confident are you that you fulfilled this task successfully? (1 = I am not confident at all, 7 = I am very confident)
- [Open question] If you experienced issues during the registration process you can describe them here. Feel free to suggest how we could solve this issue in the future.
- First at any point express how much you agree or disagree with the following statements: During the registration process have you felt …
  - Discouraged with the amount of effort needed to understand what would be done with your data?
  - Confused by the information provided in the Data Licence Agreement?
  - Confident that you understood what would be done with your data?

- Impatient because you wanted to learn more but needed to make a decision right away?

- [Open question] Could you please elaborate when you considered whether or not to share information about yourself during the registration process? What did you decide and why?

2.4 Feedback survey

In the feedback survey we asked the same questions included in the initial measurement but we have also added an extra question. A general usability scale, that was also used in D8.2 so that we can compare results and have an indication of the general usability of DataBait.

Indicate your agreement with the statements provided below:

1. I think that I would like to use DataBait frequently
2. I found DataBait unnecessarily complex
3. I thought DataBait was easy to use
4. I think that I would need the support of a technical person to be able to use this tool
5. I found the various functions in DataBait very well integrated
6. I thought there was too much inconsistency in DataBait
7. I imagine that most people would learn to use this tool very quickly
8. I found DataBait very awkward to use
9. I felt very confident using DataBait
10. I needed to learn a lot of things before I could get going with DataBait

2.5 Interviews

As said, we also did in-depth interviews to further understand the quantitative responses in a qualitative way. Here we focused on the following aspects, which have been dealt with during the living labs:

- Initial measurement
- Micro-tasks
- Feedback survey

We asked respondents to elaborate on the answers they provided during the living lab. Respondents were asked to indicate their willingness to participate in an interview during microtask 3. Next, we selected respondents for interviews based on the researchers' and respondents' availability.

3 Reporting

In this part we report on the results in the following order. We start by giving an overview of the respondents' demographic information as well as their attitudes and practices towards disclosure of information. Next, we will review the results of the initial measurement, followed by the rest of the proposed living lab outset as outlined in Figure 1. Note that the interview insights have been embedded in the feedback related to each other living lab step and will not be treated separately.

The interviews and open questions have been coded openly per research question for each living lab step. The interviews were not transcribed, researchers made notes during the interviews.

3.1 Respondent overview

Here we add the distribution of our participants' gender, **age**, **education** and **use of Facebook**. Also we mention how respondents feel about sharing data with third parties in general. These results come from the intake and initial measurement survey and we look at the Belgian and Swedish participants independently

3.1.1 Demographic information of Belgian participants

We now present the distribution of demographic data of the Belgian participants. T As shown in the figures below, we have tried to spread our sample between genders and age categories, but the most frequent participant was male, in his twenties and a daily user of Facebook.

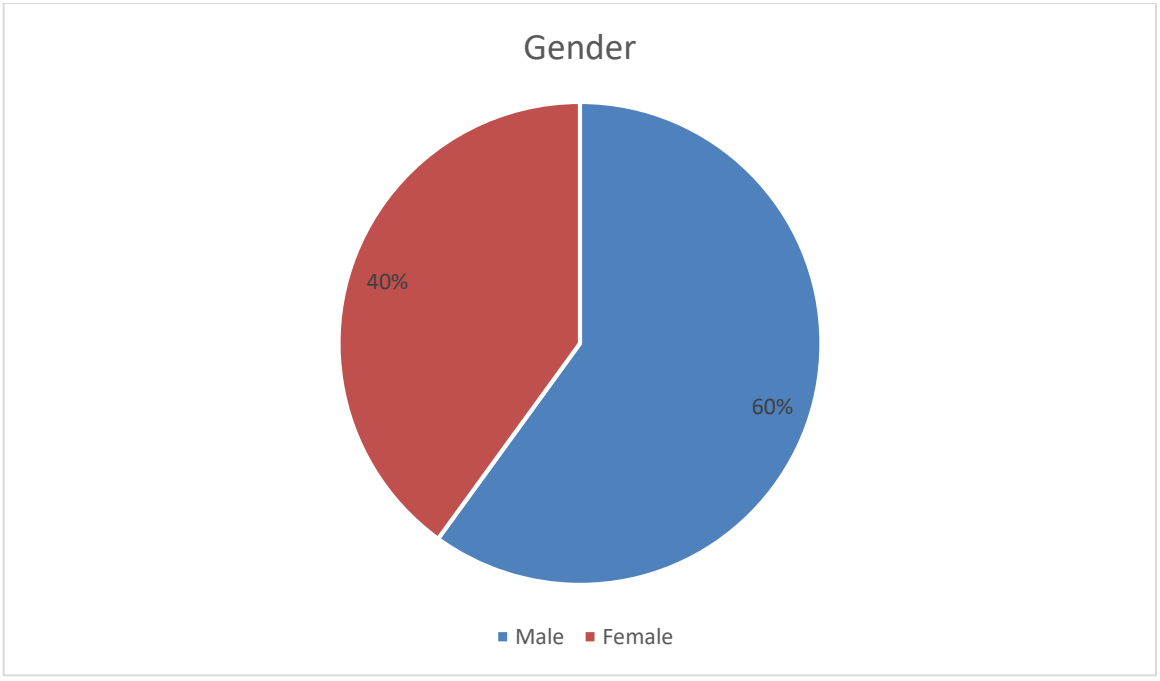

Figure 2: Gender (%, iMinds n=45)

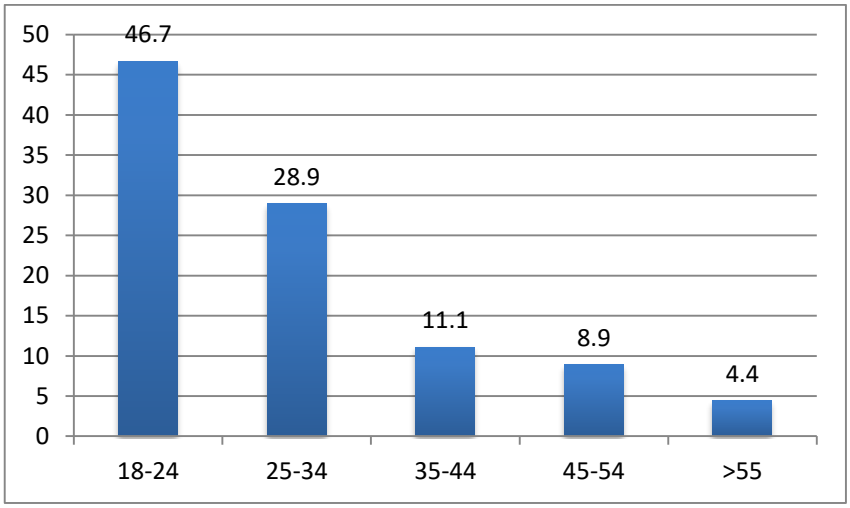

Figure 3: Age category (% ,iMinds n=45)

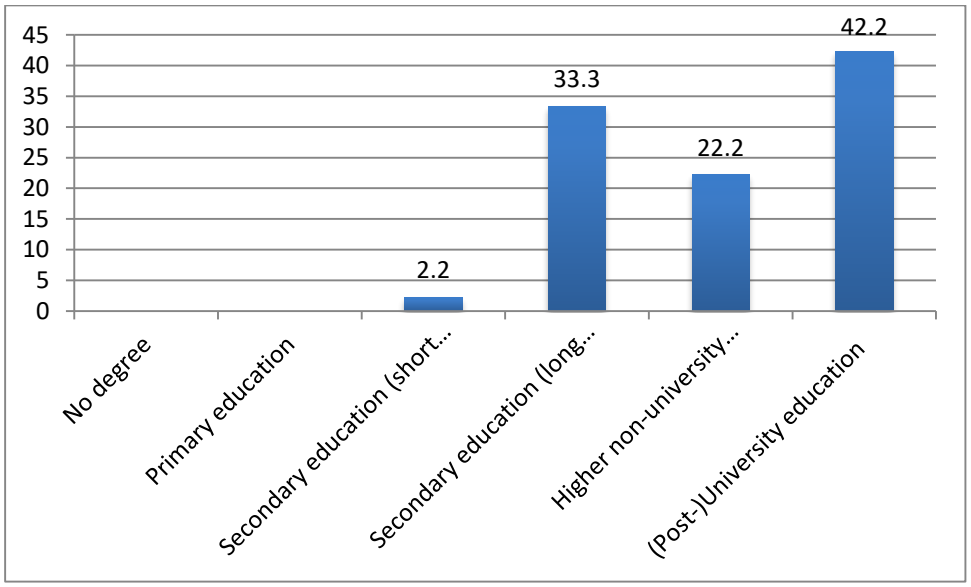

Figure 4: Education (%, iMinds n=45)

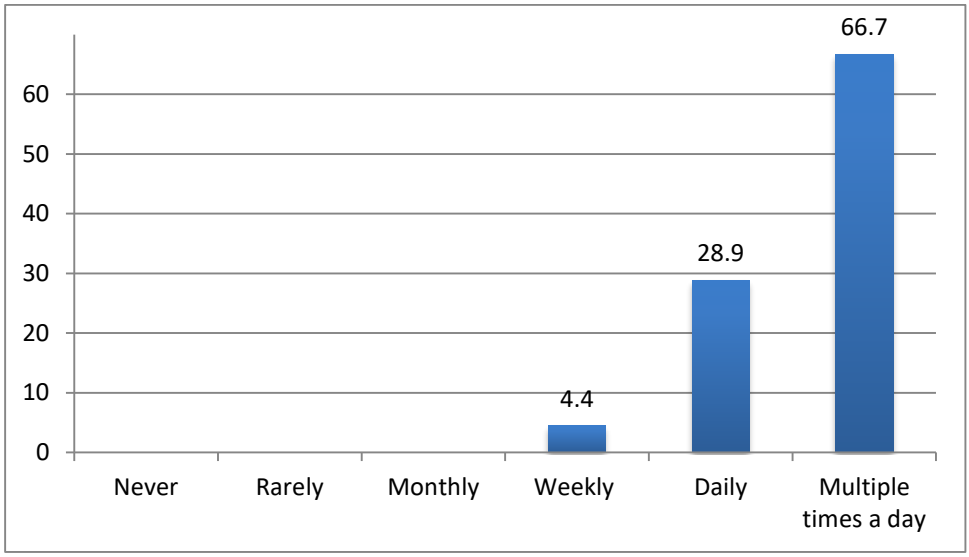

Figure 5: Facebook use (%, iMinds n=45)

3.1.2 Demographic information of Swedish participants

Intake	Initial Measurement	Micro Task 1	Micro Task 2	Micro Task 3	Feedback
240	44	28	22	18	18

Table 2 LTU participants evolution

The Swedish participants were chosen from the poll of total 240 interested people who signed up for the test. In total 61 people were invited to the test in which 44 participated in the initial measurement and 18 people were able to successfully finish all the questionaires as outlined in Table 2.

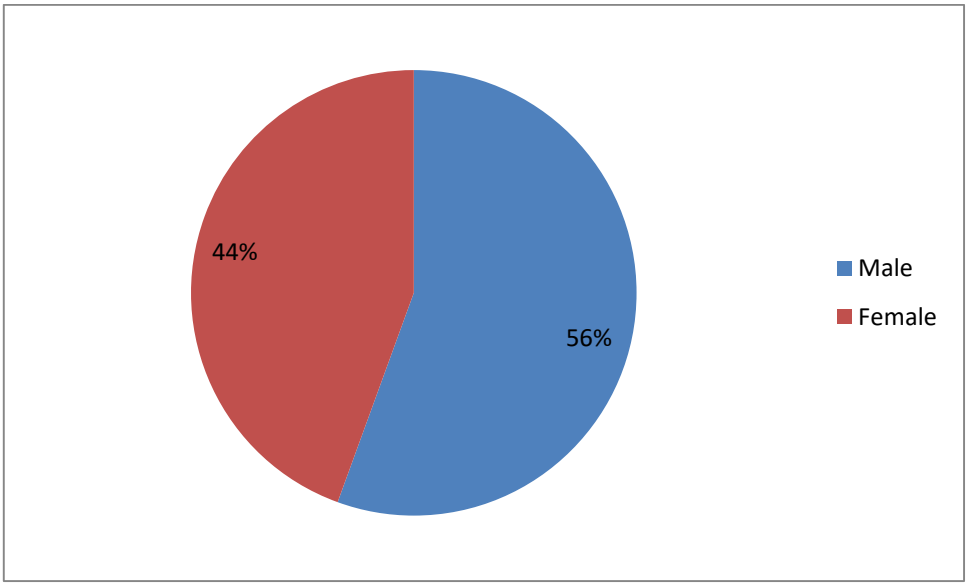

Figure 6 :Gender (%, LTU n=18)

In the LTU living lab all participants used Facebook multiple times a day.

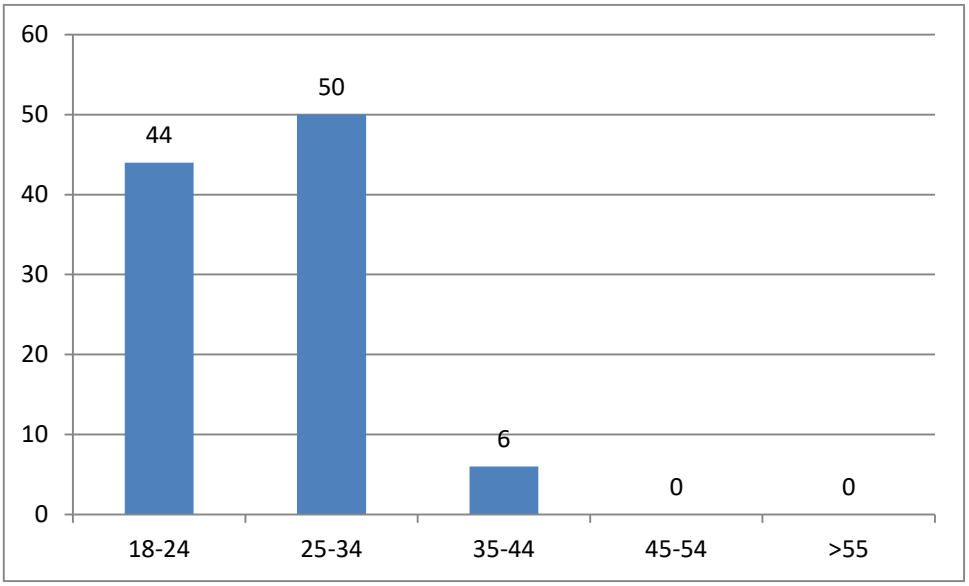

Figure 7: Age category (%,LTU n=18)

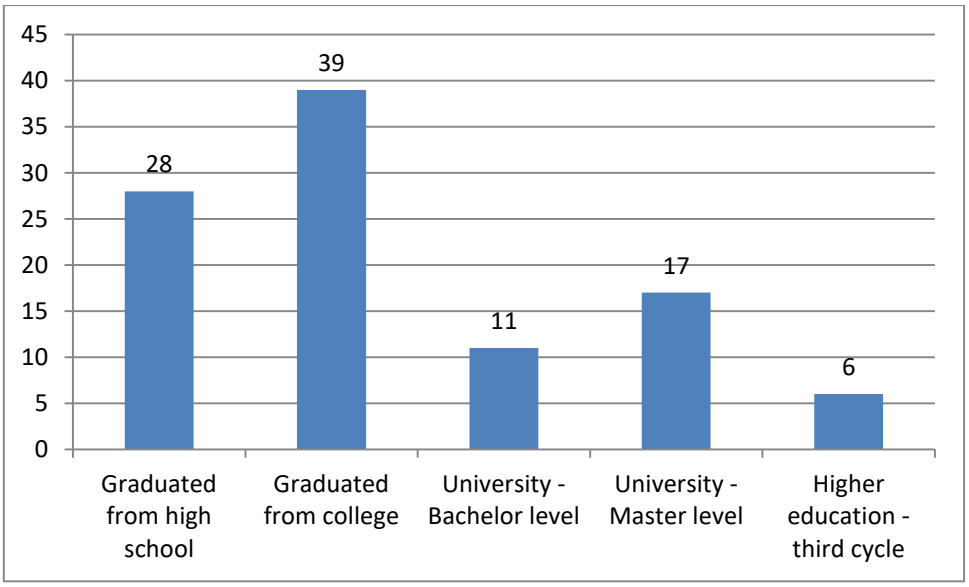

Figure 8: Education (%, LTU n=18)

3.1.3 Initial measurement

In Table 3, we summarize how our respondents felt in general about disclosing personal information to third parties. On average, respondents agree they feel discouraged by the effort required to understand what happens with their data (4,40), which is most clearly found in how confused they are with privacy policies in general (4,96) what resulted in neutral to low confidence in their understanding of what happens with their data (3,36). While this may be partly due to the effort and confusing content of the privacy policy, it may also be caused by their impatience to make a decision right away (4,47).

	iMinds (mean)	LTU (mean)
Discouraged with the amount of effort needed to understand what would be done with your data?	4,40	4,75
Confused by the information provided in a privacy policy?	4,96	5,06
Confident that you understood what would be done with your data?	3,36	2,38
Impatient because you wanted to learn more but needed to make a decision right away?	4,47	4,56

Table 3: Disclosing information to third parties in general "At any point, have you felt …?" (1= strongly disagree, 7= strongly agree) (iMinds n=47, LTU n=18)

Our respondents also had to provide an example where they decided to disclose or keep data form a third party. In these cases, respondents reflected on general rules with regard to disclosing information online, decisions to share data and decisions not to share data.

It appears that generally, respondents prefer not to disclose personal information to third parties. If there is no choice, but there is no other way to obtain the service or end a transaction, then respondents choose to disclose the data as coarse as possible. For instance, a response that we got was the following:

Hanne, 29: “Often, websites ask for a lot of data during registration and this is usually also your address, which isn’t always safe and not that important if it does not involve webshops or websites that might send you mail concerning payments. Luckily, I can fill in my town only instead of my full address in most cases.”

When respondents did disclose information, they did so for the following reasons:

- Because they had to in order to receive a benefit
- Trust was in proportion with what was asked
- It was clear why they had to disclose particular data

Respondents who withheld information did so for the following reasons:

- The benefits were too low
- Too much information was asked
- They distrusted the third party

The overview presented in Table 4 summarizes respondents’ use of PETs and settings with regard to third party data disclosure. Here we can see that respondents were most often using settings to change audiences (80,9%) or delete Facebook content (63,8%), followed by using an ad blocker to block ads from appearing (72,3%). With regard to third party tracking, respondents only blocked ads, but only a fraction blocked third party trackers (29,8%), and only one tenth visited an opt-out platform. This means that respondents will block advertising but not the tracking thereof. When we asked about this behavior during the interviews, this was mainly caused by them being unaware of tools such as Ghostery or because they were afraid this would cause websites to behave differently.

	iMinds	LTU
Using an app to block online third party trackers	29,8	18,7
Using an ad blocker to stop seeing advertisements	72,3	43,7
Visiting an opt-out platform to stop third party behavioural advertising (youronlinechoices.eu)	10,6	12,5
Using an advertising preferences menu such as the likes offered by Google or Facebook to change or delete my ad preferences profile	25,5	31,2
Deleting content I put on social media	63,8	56,2
Deleting content about you someone else put on social media	46,8	43,7
Erasing search results from Google that mention my name	8,5	18,7
I can change the audience of Facebook content	80,9	81,2
I have used the do-not-track feature in my browser settings	29,8	37,5
I have deleted cookies through my browser	59,6	62,5
None of the above	0	0

Table 4: Which actions did you perform in the last year? (%, iMinds n=47; LTU n=18)

3.2 Registration

Respondents had to self-report how difficult they found the registration process and how successful they thought they performed the task on a 7-point Likert scale, ranging from 1 (very difficult) to 7 (very easy) and from 1 (very uncertain) to 7(very certain). Results are shown in 5 and Table 6. Respondents found the registration process quite easy (mean 5,16) and were quite confident in how they performed (mean 5,43) for the iMinds living lab. These results are not comparable to LTU's findings, where respondents found the registration process slighlty easier (mean 5,31) and this was also reflected in their certainty of succeeding in this task.

		iMinds	LTU
How easy or difficult was the registration process?	Very difficult	6,8	0
	Difficult	4,5	0
	Slightly difficult	9,1	6,2
	Neutral	11,4	25
	Slightly easy	20,5	25
	Easy	9,1	18,7
	Very easy	38,6	25
	Mean score	5,16	5,31

Table 5: "How easy or difficult was the registration process?" (%, n = 44 iMinds, LTU n=18)

		iMinds	LTU
How certain are you that you fulfilled the registration process successfully?	Very uncertain	4,5	6,2
	Uncertain	6,8	6,2
	Slightly uncertain	6,8	6,2
	Neutral	6,8	0
	Slightly certain	15,9	25
	Certain	15,9	18,7
	Very certain	43,2	37,5
	Mean	5,43	5,38

Table 6: "How certain are you that you fulfilled the registration process successfully?" (%, n = 44 iMinds, LTU n=18)

Despite respondents' high ease of use and confidence, many respondents reported that they did not read the DLA or found it too long or difficult to read. This means that these participants self-assessed that they did well, although they were not fully informed about what data they were disclosing to what purpose.

When we look at the problems encountered during the registration process, we can divide these among problems with Facebook registration and challenges related to the DLA. With regard to the Facebook registration, many respondents reported they did not receive a confirmation message with regard to a successful registration. This made them wait or doubt they could use DataBait.

Suggestion: Add a clear confirmation message

Just a few remarks were found with regard to loading icons that keep on loading. In this case, respondents lost time and were unaware of what caused this problem.

Suggestion: Add a status bar for loading times (1-100%) or make sure error messages are displayed when something takes too long to load or does not work.

The DLA was mentioned more. It contains a lot of jargon, making it difficult to read and secondly, it was still too long. As Max told us during the interview and in the survey:

Max, (age 26): "During registration you get to see a lot of texts that are made by lawyers or similar people. So, much information that does not mean much to people. It doesn't make people believe it like that. I think they just show me the privacy policy in pieces on a screen and chunking it up does not make it more readable."

Here Hanne, Bart and Tom agreed, chunking up text as it was done for the DLA, does not increase its readability. This resulted in partial or diagonal readings at best.

Suggestion: Have a summary of the DLA about what is shared with whom to what purpose.

Have this summary in as much laymen's terms as possible.

Another respondent suggested showing an example of the collected data

Note that increasing the readability of the DLA may not increase understanding thereof as many respondents in the survey did not read any part of the DLA because they trusted one of the following: the involved living labs, scientific research or the specific research centers.

We asked respondents to report on how they felt about disclosing their data to DataBait as summarized in the table below:

	Discouraged with the amount of effort needed to understand what would be done with my data		**Confused by the information provided in the Data Licensse Agreement**		**Confident that I understood what would be done with my data**		**Impatient because I wanted to learn more but needed to make a decision right away**	
	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU
Mean	3,4	3,8	3,7	3,7	4,0	3,9	3,5	3,5

Table 7: "How much do you agree with the following statements? During the registration process, I felt ..." (%, n=44 iMinds, LTU n=18)

These results are very close to the 'neither agree or disagree' option for the aforementioned statements in both living labs. When we compare these to Table 3, we can see that respondents were less discouraged by the amount of effort to understand what was done with their data and less confused by our DLA than other privacy policies. They were also more confident in understanding what happened and less impatient to make a decision.

Next we inquired with an open question about their thougths while disclosing their Facebook data to DataBait. Here respondents based their decision on the safety of their Facebook profile, their trust in the project, research center or science or things they read in the privacy statement. Respondents who based their decision on the safety of their Facebook profile either referred to their settings making sure that all their content is secured or they said that they had nothing to hide. In the first case, respondents mistakenly thought that we would get the same visibility as a particular audience view, which is wrong.

As said before, many respondents did not read the DLA at all or only partly and decided that regardless of what was being gathered or what the purposes were, they trusted the party that contacted them and therefore agreed to disclose this information. Lastly, for iMinds 6 out of 36 responses referred to actual content in the DLA and based their decisions on the fact that we treat the data in a secure manner. This implies that we can be sure that at least a minority scans the DLA diagonally.

3.3 Installation of the plugin

In this task we asked users to install the DataBait plugin[4] in their supported web browsers. The aim of the plugin was twofold. First, to see who is tracking the users' browsing behaviour and secondly, to gather this information within the DataBait tool. Therefore, installation of the plugin is important to show the trackers that gather information about the users' web browsing behaviour. In the USEMP project the goal is to make the plugin available within Firefox and Google Chrome add-on market, however the process requires the plugin to be evaluated by Google and Mozilla and since the development of the DataBait plugin is still ongoing, we have asked the users to install it manually, meaning that users need to be guided through the process (there are particular steps that need to be followed for the installation). End users deemed the process cumbersome as it led to frustration. Users were used to "one-click" installations.

Suggestion: It is important to publish the plugin through Mozilla or Google Web store so that the process becomes easier.

The installation of the plugin also revealed errors that need to be fixed. These issues are enlisted below:

- Firefox version 40+ does not allow installation of untrusted add-ons. In order to work around this, users need to make changes within their Firefox browsers. These changes were difficult and led to frustration. Participants were worried about their security because Firefox warned them about possible issues with untrusted apps.
    - **Suggestion**: This message could be alleviated by having the add-on reviewed by Mozilla.
- A bug was found within the DataBait plugin which prevented the main DataBait site to open. The plugin blocks Facebook trackers. Since the DataBait website requires the retrieval of Facebook data from the backend (see D7.1) the blocking caused the DataBait page to freeze.
    - **Suggestion**: We recommend this bug to be fixed through removal of Facebook from the list of trackers within DataBait.
- Manual installation of the plugin within Chrome web browser was difficult for respondents because they had to set their browsers to developer mode.
    - **Suggestion**: Here it is also important to make users feel safer by giving them a whitelisted app.
- The plugin also had to be downloaded through a download link and next it had to be unzipped. This was troublesome for participants without an unzipping application.
    - **Suggestion**: Next time we will provide a direct link to the installation file.

To conclude, this process was not very intuitive and needs to be improved as outlined above.

3.4 Image leaks

Respondents had to self-report how easy they found the 'Image leaks' feature and how successful they thought they were in using it on a 7-point likert scale, either ranging from 1 (very difficult) to 7 (very easy) and from 1 (very uncertain) to 7(very certain). Respondents found Image leaks quite easy (average 5,71) and were quite confident in how well they performed the task (average 5,9) for the iMinds living lab (Table 8).

[4] In this report the terms plugin and add-on have been used interchangeably.

	How easy or difficult was this task?		**How certain are you that you carried out the image leaks task successfully?**	
	iMinds	LTU	iMinds	LTU
Very difficult	9,1	6,25	6,1	0
Difficult	3	0	0	12,5
Slightly difficult	0	25	0	18,75
Neutral	3	12,5	6,1	12,5
Slightly easy	9,1	12,5	9,1	12,5
Easy	24,2	12,5	30,3	6,25
Very easy	45,5	31,25	42,4	37,5
Mean	5,71	4,88	5,9	4,94

Table 8: Image leaks ease of use (%, n = 31 iMinds, LTU n=18)

	How accurate is the information offered for your complete photo collection?		**How accurate is the information offered for one photo?**	
	iMinds	LTU	iMinds	LTU
Totally not accurate	18,2	6,2	15,2	0
Not accurate	6,1	25	21,2	12,5
Slightly inaccurate	18,2	0	9,1	25
Neutral	12,1	25	9,1	37,5
Slightly accurate	24,2	12,5	24,2	6,2
Accurate	12,1	12,5	12,1	0
Totally accurate	3	6,2	3	6,2

N/A [5]			12,5	12,5
Mean	3,71	3,86	3,58	3,71

Table 9: Image leaks accuracy (%, n=31 iMinds, LTU n=18)

In Table 9Table 9 we can see that respondents are neutral[6] about the accuracy of the concepts linked to the images. But when we count the number of respondents who found the concepts tied to their collection inaccurate or totally inaccurate, this is almost 1 in four (24,3%) and this is 36,4% for particular pictures.

With regard to **usability issues**, users commented on the following. The service did not work at all at certain moments, loading times were quite long and the frame for the tag cloud was too small to see all concepts. The server did not work, which meant that DataBait did not show any concepts. This was discussed with USEMP's technical partners and was caused by security updates on the backend.

We also discussed the long loading times for image leaks, but this could not be helped as all the pictures have to be transferred from Facebook before they are processed.

Suggestion: Have a notification that warns users about this long loading time. Instead of just warning them about the waiting time, the space reserved for the concept tag cloud should contain instructions on what to do next if users find concepts they do not find relevant or too sensitive.

Lastly, some concepts were shown partly or not at all because the frame that enclosed the concept cloud was too small.

Suggestion: Have a larger frame or have the concept cloud fit within the frame.

Did image leaks increase awareness? This requires a nuanced answer; it increased awareness about this particular image recognition algorithm. Respondents were impressed by the fact that an algorithm could guess this well. But the accuracy was not high enough for participants to feel as if the algorithm really knew what it was looking at. This is illustrated in the following answer that we got to the question, "have you considered changing information showed by image leaks?":

"No, I did not consider this because it isn't worth to correct the mostly wrong conclusions (of image leaks). These conclusions are so far from the truth that it is rather funny."

Suggestion: Improve the accuracy of concepts.

[5] Due to unavailability of the service, 5 testers were not able to see their results for LTU

[6] Note that part of this neutrality may be caused by the malfunctioning server that went down during the weekend this microtask took place. Only a minority was able to use this function as intended.

Tom (age 46) also expressed how impressed he was by the progress these *Facebook* algorithms had made. He supposed that the concepts were based on Facebook's algorithms instead of DataBait's.

Suggestion: Have a more clear notification while the concept cloud is loading that explains how the concept cloud is created.

Lastly, some respondents did imply that they would like to change their pictures after seeing the concept cloud. In these cases, DataBait brought up old pictures that were forgotten by the respondent.

Suggestion: In order to render DataBait more relevant to end-users it may be interesting to let them prioritize a subset of pictures based on the age of these pictures.

	Discouraged with the amount of effort needed to understand what would be done with my data		**Confused by the information provided in the Photo insights**		**Confident that I understood what would be done with my data**		**Impatient because I wanted to learn more but needed to make a decision right away**		**Discouraged with the amount of effort needed to understand what how I should change this information**	
	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU
Mean	3,06	3,69	3,72	3,5	4,16	3,88	3,13	3,25	3,03	3,63

Table 10: "How much do you agree with the following statements? During the Photos insights process, I felt ..." (n=31 iMinds, LTU n=18)

Table 10 summarizes the results with regard to the PEW test. Here we can see that most respondents stayed close to the neutral answer options; they did not agree or disagree with most statements. This is most probably so due to the server errors. As such, it is better to rely on the qualitative feedback we mentioned above.

3.5 Location leaks

For this function we asked users to review the places revealed through DataBait and give us feedback based on the usability, accuracy and affordances aspects.

In terms of accuracy of the places there was no agreemente between the Swedish users. 40% found the overall precision of the places as not accurate to all to somehow not accurate, while 60% were neutral to believing very accurate prediction of places. The distribution of opinions on the accuracy of a single place is almost the same of above with 45% believing the accuracy of a single place to be not accurate at all or somehow not accurate and 65% being neutral to being very accurate. The Belgian panel was milder as only 24,2% found the accuracy of a single place to be inaccurate at all or somehow not accurate. This is also reflected in the higher means compared to LTU. Table 11 summarizes the results of the Swedish and Belgian participants regarding the accuracy of the locations.

	How accurate is the information offered for your all the locations?		**How accurate is the information offered for one location?**	
	iMinds	LTU	iMinds	LTU
Totally not accurate	9,1	9	3	18
Not accurate	9,1	18	12,1	5
Slightly inaccurate	9,1	14	9,1	23
Neutral	24,2	18	18,2	5
Slightly accurate	18,2	9	24,2	9
Accurate	15,2	23	12,1	23
Totally accurate	9,1	9	15,2	14
Mean	4,23	3,94	4,55	3,87

Table 11: Location leaks accuracy (%, n=31 iMinds, LTU n=18)

The quantitative track showed that most users were able to understand the use of their location leaks by the OSN provider. They were however not sure of how this information could be of value for themselves.

Suggestion: DataBait should highlight the impact of locations on personal privacy. We believe that this could be done more effectively through integration of the “Scoring

Framework" (D6.1 and D6.4) where we would be able to point out to revelations of the places into their "demographics" dimension.

The advantage of the function is that users are able to make more informed decisions based on the information provided. Still they have the option to see if they need to take out content based on their own privacy preferences. The challenge for the future of the DataBait tool in terms of location leaks is to encourage users to change their historical location data. None of our participants were eager to do this due to different reasons. Among these reasons was they believe that the leaked locations are not personal enough or not pressing enough. For example, a participant in the LTU questionnaire commented that:

"I decided not to change information about myself. Partly because I do not share as much information in general, and partly because this is only revealed little." (survey question)

We further investigated this in the interviews and the main reason was the amount of work required to go back to Facebook and find the content influencing the location disclosure. Hence users found this manner of managing their privacy too cumbersome and wanted something that is easier. This combined with lack of immediate threat from a location per se did not lead to an action. **We would therefore propose that DataBait needs to assist users to get more insights about the immediate threat of location leaks and how this disclosure could be alleviated through e.g. limiting Facebook audience or deleting the post through in application's guides and popups.**

Also following up from the pre-pilot studies, there was a step forward towards the visualization of the locations as recommend in D8.2; also locations with low confidence levels were filtered which showed improvement w.r.t. feature's goals. Some performance issues were encountered by the participants which led to long loading times.

	Discouraged with the amount of effort needed to understand what would be done with my data		**Confused by the information provided in the location insights**		**Confident that I understood what would be done with my data**		**Impatient because I wanted to learn more but needed to make a decision right away**		**Discouraged with the amount of effort needed to understand what how I should change this information**	
	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU
Mean	2,87	3,56	3,48	3,19	4,19	3,81	2,94	2,69	2,87	3,31

Table 12 "How much do you agree with the following statements? During the Location leaks process, I felt ..." (n=31 iMinds, LTU n=18)

When compared to Image leaks, Location leaks are more accurate and were able to make more sense of the information provided according to the PEW statements.

3.6 Brand detection

Table 13 illustrates that Belgian respondents found brand detection easy to use, while Swedish participants had difficulties. One main reason was that during the perriod of this task the system was quite slow and at some points the service was not available. Most of the participants were not able to see their result and this added some confusion about whether they have completed the task successfully or not. The number differs between the two studies since the service was available during the time of the iMinds' test, which was different than LTU's.

	How easy or difficult was this task?		**Was it easy to find this feature?**	
	iMinds	LTU	iMinds	LTU
Very difficult	3	18,75	9,1	12,5
Difficult	0	6,25	0	12,5
Slightly difficult	0	6,25	0	6,2
Neutral	9,1	6,25	18,2	6,2
Slightly easy	6,1	12,5	9,1	6,2
Easy	24,2	6,25	18,2	12,5
Very easy	51,5	18,75	39,4	12,5
N/A		25		31,2
Mean	5,68	4,08	6,13	4,00

Table 13: Brand detection ease of use (%, n=31 iMinds, n=80 LTU)

	How accurate is the information offered for your all the locations?		**How accurate is the information offered for one location?**	
	iMinds	LTU	iMinds	LTU
Totally not accurate	27,3	18,7	27,3	12,5
Not accurate	18,2	6,2	12,1	12,5
Slightly inaccurate	27,3	6,2	6,1	6,2
Neutral	12,1	6,2	27,3	6,2
Slightly	6,1	12,5	9,1	6,2

accurate				
Accurate	3	6,2	9,1	12,5
Totally accurate		18,7	3	12,5
N/A		25		31,2
Mean	3,1	4,08	3,19	4,00

Table 14: Brand detection accuracy (%, n=31 iMinds, LTU n=18)

While Belgian respondents were pleased with the feature's usability, this cannot be said for its accuracy: more than 45,5% found the information offered on brands inaccurate. This is also directly reflected in their opinions about how they feel about being confronted with this information.

	Discouraged with the amount of effort needed to understand what would be done with my data		**Confused by the information provided in the location insights**		**Confident that I understood what would be done with my data**		**Impatient because I wanted to learn more but needed to make a decision right away**		**Discouraged with the amount of effort needed to understand what how I should change this information**	
	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU
Mean	3,06	3,25	3,61	3,06	3,77	3,5	3	2,56	3,06	3,13

Table 15: Brand insights statements (4) ..." (n=31 iMinds, LTU n=18)

Table 15 clearly shows that respondents had little effort to understand what was done with their data, were neutral about being confused with the data presented and felt neutral about being confident that they knew what happened with their data. They slightly disagreed about being impatient to make a decision or discouraged about the amount of effort required to change something.

Only a very small amount of participants were able to comment on what they thought while being presented with the data of brand detection. This was caused by server malfunctions. Those who did respond, said there were only a few concepts related to brands and that those that were found, were most often inaccurate.

During the interviews, respondents pointed out that the low number of brands being detected and the inaccurate brands associated with their pictures, led them to believe that this feature did not work well. As a result, it could not be used to assess their privacy with regard to brands. Eric (age: 26) for example, could speculate the purpose of the function based on his interpretation and guides but not from the detected brands: "I did not have any detected brand but this feature, I think was to detect the brands that I use from my Facebook, I suppose! … the use could be like if I get commercials".

In order to render this feature relevant to users, the number of brand concepts and the accuracy needs to be increased.

3.7 Most Influenced Friends

	How easy or difficult was this task?		Was it easy to find this feature?	
	iMinds	LTU	iMinds	LTU
Very difficult	0,0	0,0	0,0	0,0
Difficult	3,4	0,0	3,4	0,0
Slightly difficult	3,4	17,6	3,4	23,5
Neutral	3,4	5,8	0,0	0,0
Slightly easy	13,8	5,8	10,3	0,0
Easy	24,1	41,1	24,1	23,5
Very easy	51,7	29,4	58,6	52,9
Mean	6,07	5,59	6,24	5,82

Table 16 Friends influence ease of use (%, n=28 iMind, n=18 LTU)

	How accurate was the most influenced friends graph?	
	iMinds	LTU
Very inaccurate	0,0	0
inaccurate	6,9	17,6
Slightly inaccurate	3,4	11,7
Neutral	13,8	23,5
Slightly accurate	34,5	23,5
Accurate	20,7	11,7
Very accurate	17,2	11,7
Mean	5,14	4,35

Table 17 Friend influence and detailed influence accuracy (%, n=28 iMind, LTU n=18)

	Discouraged with the amount of effort needed to understand what would be done with my data		Confused by the information provided in the Audience influence tool		Confident that I understood what would be done with my data		Impatient because I wanted to learn more but needed to make a decision right away		Discouraged with the amount of effort needed to understand what how I should change this information	
	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU
Mean	2,46	3,35	2,68	3,12	4,29	3,82	2,71	2,82	2,53	3,82

Table 18 Friend influence statements (5) ..." (%, n=28 iMinds, LTU n=18)

This feature helps to improve their understanding of the interactions with their friends and impact over one's privacy. As indicated by the results shown in Table 16, in terms of understanding of what the function does there was no significant issue. Participants were able to grasp the idea behind the function both through the visualizations and guides in the DataBait application. The results did not show any significant impact over users in terms of changing their privacy behavior after seeing the result of their interactions with friends. We correlate this with 64,7% of the answers revolving around being unsure about what could be done with their information. This was mainly due to the fact that users did not see an impact over their privacy based on their current most influenced friends. We will push this matter in the future pilot to see if users could find reasons and cases for changing their behavior.

Some usability issues were enumerated from the participants. First, there was a visualization matter, in particular. some of the friends' bubbles in the graph were almost outside the window. Second, there was a performance issue; in particular there was a long loading time, although a message appeared to inform about this issue.

Tom (age 46) said that the detailed friends interaction displayed two bugs, it seemed to show the same interaction for different bubbles and if he clicked on one of these bubbles, some of the links were not working. It was difficult to prove that each bubble contained another message, or that some bubbles contained the same message. Stan (age 22) mentioned that when he clicked on some of the bubbles, the links were not there, or the post was deleted or seemed to be deactivated a long time ago, but Facebook has kept the log.

Suggestion: If these detailed interactions would not move, users would have less difficulties checking them.

Suggestion: Remove dead links or explain what can be done if a user decides to find an update on Facebook but got a dead link.

In the open survey, a few responses referred to the fact that there were too many bubbles and that many of these bubbles contained very old interactions. Respondents seemed to imply that these were no longer relevant.

Suggestion: Have a means to limit or change the period of displayed interactions.

3.8 Statistical data

Swedish and Belgian participants were confused about this function. The main reason for this was that the server was not fetching data from the backend, therefore the information

was not presented. However on the right side of the screen there is a panel which shows the friends who are also using the DataBait tool. The majority of the participants were mixing these two together and had different interpretations about the panels. The appearance of the two panels (1 and 2 in Figure 9) in the same page created the impression that the graph is either related to actual Facebook friend's list or friends who also happen to use DataBait . As a result, the information in panel 2 is based on the friends using DataBait, as presented in panel 1. One participants in the interview raised this issue:

Kenny (age: 27), "Actually now that I see the screen I think I have missed in the questionaire, but isn't it based on the list of friends you have in Facebook? "

The fact that Kenny (and many more) thought he had missed the stats was because he had no friends that uses DataBait, the page was just white with no information so he thought that the page is not loading correctly.

This mixture of DataBait friends and statistical data in the same page led to confusion and misinterpretation of the stats. We suggest that panel 1 in the statistical page is moved to a new tab.

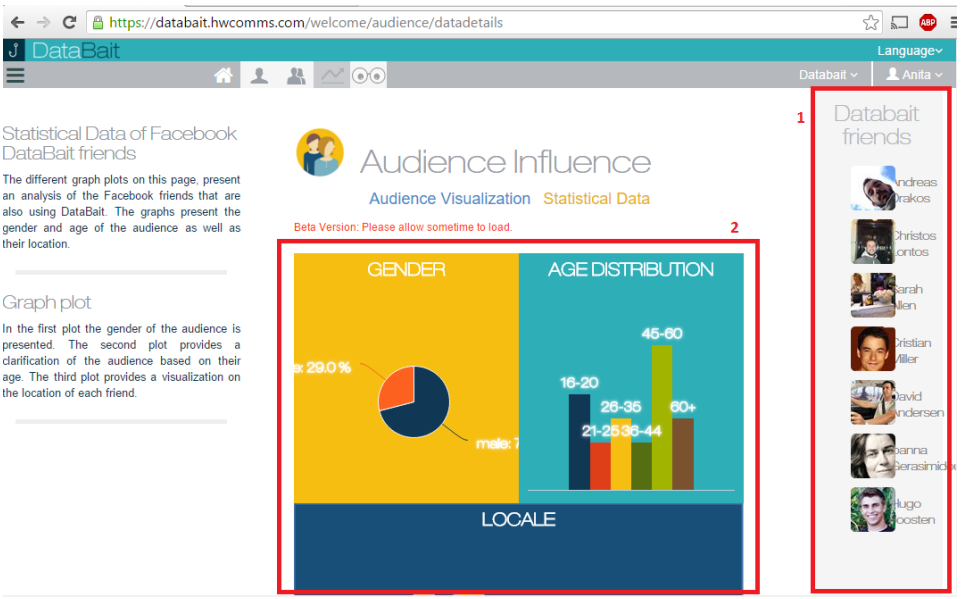

Figure 9: Audience influence

Suggestion: If users have no DataBait friends, this feature should show a message directly. This message should then inform users of the fact that this feature only works if other friends join DataBait.

3.9 3rd party tracking

This part was different from the other tools of DataBait because it showed the information other parties had gathered about a respondent, instead of what information a respondent had shared on Facebook. As such, the general feeling of respondents was amazement or shock with regard to the number of trackers per site, but for most respondents this plugin had no function, but to increase their amazement. While most participants were worried or shocked by the large numbers of trackers, they were unsure how to proceed.

	How easy or difficult was this task?		Was it easy to find this feature?	
	iMinds	LTU	iMinds	LTU
Very difficult	3,4	0	3,4	0
Difficult	3,4	6,6	3,4	6,6
Slightly difficult	3,4	0	3,4	6,6
Neutral	3,4	13,3	17,2	6,6
Slightly easy	13,8	26,6	27,6	26,6
Easy	27,6	13,3	41,4	20
Very easy	41,4	40	96,6	33,3
Mean	5,79	5,60	5,82	5,47

Table 19 Trackers (%, n=25 iMinds, LTU n=15)[7]

As mentioned during the feedback of the installation process, some respondents had difficulties installing the plugin. As a result a minority also had issues with viewing the results of this plugin.

	How accurate is the information offered for your trackers?
	iMinds
Totally not accurate	0,00
Not accurate	3,57
Slightly inaccurate	3,57
Neutral	21,43
Slightly accurate	25,00
Accurate	32,14
Totally accurate	14,29
Mean	5,21

Table 20 Trackers accuracy (%, n=25 iMinds)[8]

Respondents were quite convinced about the accuracy of this tool, only 7% believed that this tool was not accurate or slightly inaccurate. These are quite high scores compared to the other visualizations. This may be attributed to the fact that these visualizations do not rely on inferences but recorded data.

[7] Three paarticipants were not able to install the plugin due to Firefox restrictions
[8] Due to a miscommunication this question was not included in the LTU living lab.

	Discouraged with the amount of effort needed to understand what would be done with my data		Confused by the information provided in the tracker influence tool		Confident that I understood what would be done with my data		Impatient because I wanted to learn more but needed to make a decision right away		Discouraged with the amount of effort needed to understand how I should change this information	
	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU
Strongly disagree	25,00	20	21,43	26,6	7,14	13,3	25,00	33,3	17,86	26,6
Disagree	35,71	13,3	28,57	26,6	17,86	40	25,00	13,3	28,57	26,6
Somewhat disagree	7,14	20	10,71	0	7,14	26,6	17,86	20	17,86	6,6
Neither agree or disagree	17,86	13,3	25,00	13,3	32,14	0	21,43	13,3	21,43	6,6
Somewhat agree	7,14	13,3	0,00	13,3	21,43	6,6	10,71	13,3	7,14	6,6
Agree	3,57	6,6	10,71	6,6	10,71	0	0,00	0	7,14	6,6
Strongly agree	3,57	13,3	3,57	13,3	3,57	13,3	0,00	6,6	0,00	20
Mean	2,71	3,60	3,00	3,33	3,89	3,00	2,68	2,87	2,93	3,4

Table 21 Trackers statements (5) (%, n=25 iMinds, LTU n=18)

Participants were not discouraged to understand what happened with their data (2,97), but they were also not really confident about what happened with their data. When we couple this to the open question, respondents mentioned that they could understand at a glance that they were being tracked by many trackers. But, they were unable to understand what these trackers were doing. If we compare this to the interviews, we can understand why this led to very little changes with regard to blocking third parties or implementing other privacy management changes. Participants like Tom, Max and Bart were eager to decrease the number of trackers, but since they were not aware which trackers were needed to keep the site running and which were not, they did not change anything because it might cause some of the websites they visit to malfunction.

Suggestion: Categorize trackers according to their activities and inform users how blocking them would affect their browsing experience.

Bart refrained from changing anything because he did not want to change each tracker. He felt there should be an option to change access trackers have in general: if tracker A has 10 tracking instances, Bart would like to be able to disable all these instances on all websites.

Suggestion: When trackers are blocked, make sure these preferences are applied to all websites.

3.10 Feedback survey

The aim of the feedback survey was to see if there has been any change of the users' attitudes towards privacy concerns. To measure any change of attitude, we asked the same questions as in the initial questionnaire regarding what they thought of personal and institutional privacy. Although change of attitude cannot solely be related to the use of DataBait, this matter was pursued in the interviews. The questionnaire also had a part of system usability scale (SUS) questions, giving a global view of subjective assessments of usability.

The result from SUS in Sweden and Belgium shows that participants were positive towards the tool as a whole. The tool was deemed to be easy to use, self-explanatory with consistent language and concepts. Table 22 shows the result and distribution of answers for each question for Swedish and Belgian respondents.

Question	1 (Strongly Disagree)		2		3		4		5 (Strongly Agree)		Rating Average	
	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU	iMinds	LTU	**iMinds**	**LTU**
I think that I would like to use this product frequently.	17,2	17%	17,2	50%	24,1	17%	37,9	11%	0,0	6%	**2,86**	**2,39**
I found the product unnecessarily complex.	31,0	39%	34,5	33%	6,9	28%	13,8	0%	10,3	0%	**2,36**	**1,89**
I thought the product was easy to use.	6,9	0%	10,3	6%	17,2	17%	34,5	56%	27,6	22%	**3,68**	**3,95**
I think that I would need the support of a technical person to be able to use this product.	37,9	50%	37,9	28%	10,3	22%	6,9	0%	3,4	0%	**1,96**	**1,73**
I found the various functions in the product were well integrated.	6,9	0%	24,1	17%	20,7	44%	41,4	28%	3,4	11%	**3,11**	**3,34**
I thought there was too much inconsistency in this product.	13,8	28%	48,3	39%	10,3	33%	20,7	0%	3,4	0%	**2,5**	**2,06**
I imagine that most people would learn to use this product very quickly.	10,3	0%	20,7	0%	17,2	33%	41,4	50%	6,9	17%	**3,14**	**3,84**
I found the product very awkward to use.	17,2	39%	37,9	50%	10,3	11%	20,7	0%	10,3	0%	**2,68**	**1,73**
I felt very confident using the product.	3,4	0%	20,7	6%	24,1	33%	34,5	39%	13,8	22%	**3,36**	**3,78**
I needed to learn a lot of things before I could get going with this product.	44,8	39%	37,9	44%	6,9	0%	3,4	11%	3,4	6%	**1,79**	**2**

Table 22: System usability scale (%,iMinds n=28, LTU n=18)

One only vigilant concern regarding the results presented in Table 22 is the fact that users were not eager to use the DataBait in the future. This is signified by the low mean level 2,39 for LTU and 2,86 for iMinds. After further investigating the matter in the in-depth interviews we came into the conclusion that users are seeing the PETs as a deluxe application which could only act as “just for fun” means for average users. The same opinion was reflected in the Belgian interviews. Respondents felt that the tool could only satiate their curiosity once. After this, the tool lost its value because participants knew that little new information would be present the next time they visit in the near future.

Apart from the research nature of this pilot study, one reason participants took part was their sense of curiosity. This was something that subsided after using the tool which could be related to deduction of their DataBait result as being something normative. We were able to see this theme repetitively within the interviews indicating that participants were acting within the norms of their privacy preferences and DataBait was just another stamp over approval of their disclosure behavior.

Suggestion: The USEMP project needs to take into account and create a mechanism of which this curiosity is sparked regularly.

3.11 General recommendations

Any information provided directly about DataBait (e.g. explanations of visualizations) is not read by the users. It is shown on the left hand side as grey text and far enough away from where the visualization is which renders it easily ignorable. It is only when this information is shown or referred to during the interviews that it is noticed.

Suggestion: Since loading does not happen instantaniously, it may help if we move the information on the side to the center during loading times.

Pictures, locations, and text are all perceived as UGC and therefore treated as content that will only be seen by other users. Institutions such as banks, apps, governments, are not considered. This has some implications. First of all, it means that users only need audience settings and a visualization of the current settings. Secondly, we are unable to increase awareness with regard to what would happen if third parties such as banks, apps or governments learn or infer things from this information.

Suggestion: Notify users through possible scenarios of any threats if these parties get access. It would be interesting to open up the conversation and get them to think beyond their regular Facebook audience.

Many respondents were still unaware about trackers as they have not really thought about the inference possibilities these trackers have. When Bart (age 38) was told that we are not able to uncover what these parties can infer, he told me that he had never thought about the fact that these websites might actually infer or compose profiles. For him, cookies were black boxes storing data. What data is collected by trackers or what happens with this data had not entered his mind because he was still in the process of learning about the number of trackers.

Suggestion: In line with the previous suggestion, we need to add more information about what these trackers might infer.

General remarks

When Hanne (age 29) and Bart (age 38) told us about how they did not worry about their privacy, we asked them whether this was the case because they had changed their privacy

settings (as they mentioned in the survey). As a result, they said they need not worry about their privacy because they had changed their settings. So the question if somebody is worried, occupied or thinking about his or her privacy, may be answered negatively because people have already changed their settings in the past. In this case, we may need to add a new option: “I do not worry about my privacy because I believe my settings are good enough”.

The PEW statements helped only for general feelings towards gauging disclosure decisions to third parties. The open question that asked what users thought while we presented them with new information did help. We should use these answers as answer possibilities in the next pilot.

Conclusion

The research questions of this deliverable were:

1. Where should we further improve the ease of use of the different DataBait tools?
2. Is DataBait a PET that can increase user awareness with regard to information disclosed on social media and through cookies?
3. Does DataBait offer users actionable information to revise their online and social media disclosure?

But these questions cannot be answered without incorporating the needs DataBait tried to address for its users. We therefore evaluated whether the pilot was able to address the following needs:

Increase awareness about information disclosure and use of personal data disclosed on Facebook for third parties for the following types of information:

- *Location information*
- *Image information*
- *Friends interactions*
- *Trackers*

This should be achieved in a user-friendly manner and in an informed way.

We will first address the needs identified in USEMP and how these were fulfilled during the DataBait pilot. Next we answer our research questions.

When referring to the first three points, location, images and friends information, we can see that DataBait was able to increase awareness in an intelligible way. DataBait presented respondents with information they recalled having shared on Facebook in the past. Our participants had no need to change this information because they deemed it safe due to the audience settings they had applied to it. In this regard, DataBait is only able to increase awareness for social privacy issues but not issues related to third parties. Since this is part of the needs identified, we have to conclude that we did not succeed entirely. The latter cannot be said about trackers. The new addition of the plugin and its associated visualisations have succesfully increased awareness with regard to personal data disclosure to third parties.

Were these needs fulfilled in a usable and informed manner? Let us address user-friendliness first. The SUS scale indicated that the usability of the tool is acceptable. This would most likely have been higher if we did not experience server issues during the pilot.

Were our respondents informed? This question is more difficult to answer because our respondents had a living lab bias. Respondents trusted us and therefore neglected to read the DLA. On the other hand, many respondents also mentioned that they were not planning to ever read this document, because even if it is chunked up, it is still too long and difficult to read. Here we have to make a trade-off, how much easier can we make this decision without dumbing the informaiton down to an uninformed decision? For our respondents, presenting more information with regard to what we gather and how we store it would be a priority.

How can we further improve the ease of use of the DataBait tools? If we want to further improve DataBait, we should add more means to (pre)select the information DataBait provides. The large amount of information disclosed with regard to image detection, friends interactions, trackers and locations renders the perceived amount of work insurmountable. A selection or pre-selection will render the provided information easier to digest. If we further determine ways to prioritise what should be shown first, we can become even more relevant to end-users. This also solves a part of the second research question, is DataBait a PET that increases awareness?

We have answered this question positively, but is this the right kind of awareness? No, users have become aware of social privacy issues for location, images and interactions on the one hand, and respondents have become aware of trackers and thus third parties on the other hand. Our participants were unable to see threats from third parties for their content. So, in order to improve this dimension of awareness, we will have to include examples or cases of what can happen if third parties learn about this information. These will not be added to frighten users, but to help them understand what DataBait concepts mean for different third parties.

Lastly, we wanted to see whether the DataBait pilot offered enough actionable information to change personal information disclosure. Little to no participants have deleted or altered information although they implied that they would like to have changed what certain trackers could track. While we offered ways to block trackers in DataBait, this was not done because respondents were afraid of negative consequences that might hamper their viewing experience of particular websites. Another reason to refrain from blocking trackers was the large amount of trackers. So, in order to solve this issue, we will have to refer to other tools that are able to group trackers and delete them collectively. One of these solutions would be Ghostery.

To conclude, we as a consortium will have to make a choice with regard to what information we want users of DataBait to see first. Here we will have to select our tips of the ice bergs in such a way that they better explain what third parties may do. This means that we make certain political choices with regard to threats we think we need to expose. But the risk of this paternalistic/ideosyncratic position is still better than the risk of giving too much information at once.

4 Bibliography

iMinds. (2016). *Proposed updates to DataBait* (p. 14).

Madden, M. (2012, February 24). Privacy management on social media sites. Retrieved from http://www.pewinternet.org/2012/02/24/privacy-management-on-social-media-sites/

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, *15*(4), 336–355. http://doi.org/10.1287/isre.1040.0032

Pierson, J., & Lievens, B. (2005). Configuring living labs for a "thick"understanding of innovation. In *Ethnographic Praxis in Industry Conference Proceedings* (Vol. 2005, pp. 114–127). Wiley Online Library. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1559-8918.2005.tb00012.x/abstract

Rainie, L. (2015, December 30). Americans conflicted about sharing personal information with companies. Retrieved from http://www.pewresearch.org/fact-tank/2015/12/30/americans-conflicted-about-sharing-personal-information-with-companies/

TNS opinion & social. (2015). *Special Eurobarometer 431 Data Protection* (Special Eurobarometer No. 431) (p. 220). Brussels: Directorate-General for Justice and Consumers. Retrieved from http://ec.europa.eu/public_opinion/archives/ebs/ebs_431_en.pdf

Topline: Sharing personal information with companies. (2015, December 29). Retrieved from http://www.pewinternet.org/2015/12/29/americans-conflicted-sharing-personal-information-topline/

USEMP. (2015). *Annex 1 - Description of work*.

5 Annexes

5.1 Annex 1: living lab invitation

"Dear…,

Do you ever wonder what happens with your Facebook or browsing data? And more importantly what other parties may deduce from your data? We are in the process of creating a tool to help those who want to know more about how their data is used by Facebook and third parties.

Within USEMP, we try to create a tool to help you understand what happens with your data and help you on your way to change or delete parts of your personal information if you feel they should not pass through the hands of third parties.

If you would like to participate, in making sure this tool provides you and other users with relevant information and features, please participate. Your feedback and suggestions will be implemented during the course of this living lab. What is more, you will learn about your own online privacy and means to change aspects you no longer like.

We will ask you to perform tasks such as filling in short surveys, use our tool and report on what you did with the tool. Next we will conclude this project with an interview at a location of your choice that will last for one hour. The period that we need your input runs from the beginning of February to the 3th week of February.

Best regards,"

5.2 Annex 2: Detailed living lab tasks

Initial measurement and Micro task 1: registration

Registering with DATABAIT

Objectives

Get respondents to log in on DATABAIT (This requires an admin to add their user ID to the DB app in Facebook[9])

Ask about the usability of the process, their attitude towards sharing this information with DATABAIT and lastly, when they actively thought about disclosing information.

Install the DATABAIT plugin in Firefox or Chrome

Respondent instructions

Welcome to your first assignment. During this assignment we will ask you to do three things for us. Fill in a short survey, register and install a plug-in.

Visit the survey here and fill in the questions until you are told otherwise.

Next, visit https://databait.hwcomms.com and register by reading the instructions on this website. Also, open this survey and fill in the questions after completing this process. In case you encounter problems, you can mail them or add them in the survey at the remarks section.

[9] https://developers.facebook.com/apps/1547194038870780/dashboard/.

Remember that the DATABAIT tool only works with Chrome and Firefox.

After completing the registration process, please fill in the following questionnaire:

How easy or difficult was this task? (1 = very hard, 7 = very easy)

How confident are you that you fulfilled this task successfully? (1 = I am not confident at all, 7 = I am very confident)

If you experienced issues during the registration process you can describe them here. Feel free to suggest how we could solve this issue in the future.

First at any point express how much you agree or disagree with the following statements: During the registration process have you felt …

Discouraged with the amount of effort needed to understand what would be done with your data?

Confused by the information provided in the Data Licence Agreement?

Confident that you understood what would be done with your data?

Impatient because you wanted to learn more but needed to make a decision right away?

Open question:

Could you please elaborate when you considered whether or not to share information about yourself during the registration process? What did you decide and why?

Installing the DataBait plugin

In order to tell you who has been following you through cookies online, we need you to install a plugin. Can we ask you to install this plugin on your most used computer at home? In order to present you with who is tracking you, we need you to use the browser you installed the plugin on.

(In case you have difficulties installing this plug-in, please contact (address of direct line of help for Databait issues (iMinds: Tom or Paulien; LTU: Ali))

If you decided to use Firefox, follow the instructions below:

The plug-in is in beta and this requires a change in Firefox

Enter ‘about:config’ in the address bar

Use the search field on the right hand side and type in ‘xpinstall.signatures.required’

It mentions a value ‘true’, double click it or edit it so that it says ‘false’

Log in on Databait

Click on ‘install Databait Plugin’

You should now be asked by Firefox to install this plugin and follow its instructions

If you decided to use Chrome, follow the instructions below:

To install in chrome,

Unzip this file anywhere (but remember where, for example desktop)

Go to extensions

Check developer mode (top right)

Select “load unpacked extension”

Navigate and select X\usemp-databait-plugin\firefox\content\databait.safariextension\opera\chrome

The plugin should be working now

Lastly, tell us if you encountered any difficulties during the process?

After the micro task

After each micro task, we will summarise the feedback provided by all respondents in order to highlight what parts of DATABAIT require improvements. We will invite respondents to suggest improvements via mail by the time of the next micro task or from as soon as they see the summary. This means we will have to make it a recurrent question at the start of each micro task.

Micro task 2: Image leaks, location leaks and brand detection in images

Image leaks

Objectives

Have respondents test the image leaks function to understand usability and how aware they became about this information. Secondly, ask them what they require to act upon this information and if enough information was provided.

Instructions

Before we get started with the new assignment, we ask you to think about changes or comments you may have after seeing our feedback. If you wish to change something or suggest anything, please do so here:

Please go to the DataBait app, click on “My Disclosures” and go to Photo Insights. Next use this function to see what your Facebook pictures reveal about you. Try to stay on this page and see what could be found on this page. After that fill out this questionnaire:

How easy or difficult was this task? (1 = very hard, 7 = very easy)

How confident are you that you fulfilled this task successfully? (1 = I am not confident at all, 7 = I am very confident)

How easy was it to find this function? (1 = very hard, 7 = very easy)

How difficult was for you to understand the language and concepts (words) used in the image leaks? (1 = very hard, 7 = very easy)

In general, how accurate are the information provided for your whole image collection? (1 =not accurate at all, 7 = very accurate)

In general, how accurate are the information provided for a single image? (1 =not accurate at all, 7 = very accurate, Uncertain)

If you encountered any problems or found points we can improve, please let us know below: First at any point express how much you agree or disagree with the following statements: During the registration process have you felt … (7 point likert: 1: I completely agree – 7: completely disagree)

Discouraged with the amount of effort needed to understand what was done with your data

Confused by the information provided by image leaks

Confident that you understood what was done with your data

Impatient because you wanted to learn more but needed to make a decision right away

Discouraged with the amount of effort needed to understand how to change this information

Open question:

Could you please elaborate when you considered whether or not to change information about yourself shown in the Photo Insights tool. What did you decide and why?

Location leaks

Analogous to image leaks

Brand detection in images

Analogous to image leaks

After the micro task

After each micro task, we will summarise the feedback provided by all respondents in order to highlight what parts of DATABAIT require improvements. We will invite respondents to suggest improvements on the blog by the time of the next micro task or from as soon as they see the summary.

Please tell us if you have, from the first questionnaire until so far, experienced any issue or would like to give any suggestion that could help us:

Micro task 3: Friends influencer and 3rd party tracking

These are analogous to the example provided in image leaks

Audience Influence

From DataBait's main screen click on Audience Influence and wait for sometime for the result to load. Read the information about the page on the left side and answer the following questions:

[same as image leaks]

Detailed Interactions

Within the Audience Influence page click on the next tab “Detailed Interactions”. Read the information on the left side of the page. Try to click on different circles and see what happens. Try to play around with the bubbles and then answer the following questions:

[same as image leaks]

Statistical Data

Under Audience Influence page click on Statistical Data. Look at the information provided and then answer the following:

[same as image leaks]

User Trackers

From the main DataBait screen click on Trackers icon.

visit cnn.com

- open any news item
- on the right side there a share to social media option
- open DataBait plugin
- list all trackers
- find “gigya tracker” (share in social media) and disable it
- re-load page (automatically)
- social media share option is not shown

Feedback survey

Here we will ask questions similar to the initial measurement survey to be able to measure an increase in awareness. In case we lack the time to roll out these surveys, the feedback survey will be part of the F2F interviews (see below).

Ending interviews

We will end this LL method with F2F interviews. In these interviews we will couple back to the open question answers and delve deeper into each tool.

The interview guide will consist of :

A short introduction of the researcher, a summary of what is to be expected during the interview, the fact that this interview will be recorded and if they are transcribed respondents names will be replaced by a pseudonym.

We will discuss each tool by addressing the following questions in this order;

Ask about their internet use

Ask about their Facebook use

Where do they access it?

How often?

Why?

Ask about their use of PETs and privacy settings?

Refer to the answers of the survey

When and why have they used these tools?

Did they use any during the pilot?

Run over each tested tool and ask the following while keeping the answers of the surveys nearby

How easy did you find part x to use?

What did you think about the information provided by DATABAIT?

How did you feel about the information DATABAIT told you disclosed?

What information was clear?

Did you do something after receiving this information?

What information was missing?

Ask about DATABAIT in general

What have you learned new from DATABAIT?

Will that change the way you use FB in the future? i.e. your disclosure will change? In what way?

Were you expecting to see the information you saw in DATABAIT?

What should be changed?

What do you want to see more?

Will you use DATABAIT in the future?

What is still missing?

5.3 Annex 3: Intake survey

- Do you have a Facebook account? (If not, end the survey)

Internet access and browser

- Do you have access to internet at home? (If not, end of survey)
- Can you use one of the following browsers for a month? (If something else, end of survey)
    - Chrome
    - Firefox

Facebook usage

How often do you use Facebook?

- Never
- Very sparsely
- Monthly
- Weekly
- Daily
- Multiple times a day

Socio-demographic

What is your gender?

- Male
- Female
- Other

What is your birth year?

What is your education level?

Can you provide us with your Facebook ID?[10]

5.4 Annex 4: Interviewed respondents

Pseudonym	Age	Date of interview	interviewer
Max	26	02-Mar-16	Rob
Hanne	29	08-Mar-16	Rob
Bart	38	08-Mar-16	Rob
Tom	46	07-Mar-16	Rob
Matilda	24	17-Mar-16	Paulien
Peter	28	17-Mar-16	Paulien
Donna	26	15-Mar-16	Paulien
Marc	22	15-Mar-16	Paulien
Eric	26	29-Feb-16	Ali
Stan	22	1-March-16	Ali
Kenny	27	4-March-16	Ali
Kyle	26	4-March-16	Marita

[10] The latter is required to register users to DataBait on Facebook because this applicaiton is still in development.