D8.2

Pre-pilot evaluation report

V1.0 2015-09-30

Ali Padyab (LTU), Marita Holst (LTU), Tom Seymoens (iMinds), Laurance Claeys (iMinds) Theodoros Michalareas (VELTI), Noel Catterall (HWC)

This paper reports the results from the early evaluation of the DataBait tool. The evaluation was performed by testing basic functions of the tool together with the intended users of the system within Sweden and Belgium. Two approaches were employed for this evaluation of the DataBait tool. One study was performed in terms of concept evaluation with mock-ups of the system for gathering the evaluation criteria for later phases of the evaluation and second, (some of the) actual features of the system were tested through user interaction. Both technical and non-technical partners were in close collaboration for this delivery. The results show that users appreciated the concepts and the functions of the DataBait tool. However, there are some considerations to be taken into account for future versions of the tool for example password retrieval function, more accurate location concepts, guidelines and some minor visualization improvements. Therefore, this report acts as a formative evaluation approach which will be a feedback to the developers and the USEMP consortium for further improvements.

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	Pilot Studies and Evaluation
Deliverable lead org.	LTU
Deliverable type	Report
Authors	Ali Padyab (LTU), Marita Holst (LTU), Tom Seymoens (iMinds), Laurance Claeys (iMinds), Theodoros Michalareas (VELTI), Noel Catterall (HWC)
Reviewers	Eleftherios Spyromitros-Xioufis (CERTH) Steven Strachan (CEA)
Version	1.0
Status	Final
Dissemination level	PU: Public
Due date	2015-09-30
Delivery date	2015-09-30

Version	Changes
0.1	First outline ToC
0.2	Contributions from involved partners
0.3	After Input from VELTI
0.4	After Input from HWC
0.5	Complete version for internal review
1.0	Final version

Table of Contents

1.Introduction

The aim of this deliverable is to report on the result of formative evaluation of the USEMP tools (called DataBait) from conceptual perspective and its actual test by the end users. Two rounds of tests have been conducted. The first round was performed together with the potential end users of the DataBait in the form of concept evaluation in Sweden. The second round was done through actual testing of the live system in Sweden and Belgium. The current version of the DataBait tool has limited functionality, that's why we call this deliverable "pre-pilot" evaluation report.

The objectives of the piloting are as described in the DOW to:

- To generate realistic data from citizens usage of the USEMP framework
- To carry out pilot tests in iterations; firstly a pre-pilot and thereafter two large-scale pilots
- To evaluate the USEMP framework contribution to citizen empowerment from a multidisciplinary perspective.

This report acts as a feedback to the developers and the whole consortium about the evaluation of general concepts behind the USEMP project as well as issues with the DataBait tool in terms of User Experience, Usability Aspects, Understandability, Navigation and Learnability.

In the first round of the evaluation, we conducted a series of focus groups together with the users and showed them the mock-ups of the DataBait to gather their insights on different functions of the tool. Specifically focused on what is required from a Privacy Enhancing Tool to be adopted in daily end-user use, we brought together test requirements precisely targeted for the DataBait tool. We found that Awareness, Benefit Realization and the message conveyed play an important role in better aligning the tool to the user's expectations. Based on the requirements gathered, we performed a series of workshops and one-to-one interviews to the end users. The combination of two approaches complemented each other in the sense that we were able to collect data from very detailed to general level pieces of information.

The process for user involvement was divided into micro tasks with the clear objective of keeping the users engaged and committed to the process. Related to these micro tasks, the users got incentives. This approach was implemented to make sure that the users were not overburdened with too much work at any one moment.

The structure of the report is as follows. First we will introduce the main objectives of the pre-pilot derived from the DOW together with the contexts where the tests were performed. In chapter 3, we present the main theories used to guide the tests and interpretation of the results. Chapter 4 gives an overview of the DataBait tool and introduces the functions used in this pre-pilot study. We present the methodology and results of concept evaluation phase in chapter 5, accompanied by the test requirements gathered for the DataBait tool. Chapter 6 shortly describes the ethical aspects of the tests. The results of the test are presented in chapter 7 followed by the discussion and feedback to the developers and the USEMP consortium in chapter 8.

2.Main Objective of Pre-Pilot

As described in the DOW the overall aim is to:

> *"In this task an early version of the USEMP platform, named DataBait, will be tested in a small-scale pilot with the Living Labs in the project. In this study, the focus will be to carry out a user-centered formative evaluation of the platform, which will function as input for its continuous development and input to task 8.3. Here all the equipment, routines and requirements are established. The pre-pilots will be carried out by using ICT tools to support the process of user interactions."*

The pre-pilot follows the process outlined in D8.1

2.1. Pre-Pilot Contexts

The pre-pilot was conducted in two different countries, namely Sweden and Belgium, by LTU and iMinds, respectively. Having this pre-pilot in two different geographical and cultural contexts has the benefit that more general conclusions can be drawn out of the resutls, as the tests transcend any local peculiarities. Both partners have access to users and resources through their Living Lab institutions. iMinds, through iMinds Living Labs, has access to a test panel of 20 000 people, which are all profiled with regard to their ICT possession and use. LTU, host of Botnia Living Lab, has a test panel with 6000 dedicated test pilots. In this test panel there is a diversity in the ages of the participants ranging from 18-70 years old.

2.1.1. iMinds[1] Living Labs

To describe the context of iMinds Living Labs we will take a look at how they describe their activities on their website:

The purpose of iMinds Living Labs is to support both the research community and entrepeneurs when developping their products and solutions. They do this by providing the opportunity to test and co-develop with the target audience from the start of the development process. Their aim is to provide the products and services a better preparation to enhance their chance for success on the market. iMinds Living Labs is a leading organization in the European Network of Living Labs – EnoLL (http://www.openlivinglabs.eu). They try to ameliorate and update their tools and methods throughout the various living lab projects.

At the core of any living lab study are the users who help to test and develop future products before they are launched. In this way it is possible to analyze how they are being used in daily life and to see if the target audience can be reached. Anyone can be a test user, depending on the scientific or entrepreneurial necessities.

Today, iMinds have 20.026 individuals assigned to the test panel and they're all willing to participate in our research. The majority are men (56% vs. 42% with 2% unknown). Age-wise, the largest group is situated between 20 and 29. The oldest panel members are two men, aged 94 and 89. Geographically, East- Flanders (where iLab.o HQ is based) is the most represented region in our panel (32%) followed by Antwerp (18%). Only a small sample of 1% lives in Wallonia, which is obvious since a large majority of our research is conducted in Dutch.

[1] Parts of this paragraph are based on the iMinds website: http://www.iminds.be/en/succeed-with-digital-research/proeftuinonderzoek and deliverable D8.1

iMinds Panel Management

For the recruitment of users for the USEMP pre-pilot, we worked closely together with members of iMinds panel management who are responsible for recruiting, motivating and co-ordinating the panel of test users. They look for the right test users according to the goal of the project or certain demographic requirements provided by the researcher. Before the start of a project the panel management team is also responsible for all promotional communication to potential test users. This involves tasks such as providing copy and design for invite e-mails, final check-ups of surveys and potentially targeting new test users by sharing information of the project and call-to-actions on social media.

Their main goal here is to recruit the right people for the project and to get them involved but they're also very committed to showing the world that scientific research is not necessarily dull and that everyone can make a difference in developing a new tool or service.

Secondly, they look after the daily communication with the panel and they're the single point of contact for test-users during the project in case of questions or problems. In general iMinds Panel Management try to build a genuine community feeling among our test users by providing those results and feedback of the projects they have participated in by creating fun info graphics and keeping them informed on important tech and media news on our numerous social media channels.

2.1.2. Botnia[2] Living Lab

Botnia Living Lab was founded in 2000 and is a world-leading environment for user-centric research, development and innovation (RDI), supported by innovative methods, tools and experts. The Living Lab is an effective member of the European Network of Living Labs (www.openlivinglabs.eu) and was one of the founders of the network in 2006. Botnia Living Lab offers an integrated environment of people, infrastructure, tools, processes and services for research, development and test of new and emerging distance-spanning technologies and its applications.

Botnia Living Lab offers research expertise in user involvement and testing, Methods qualified by research for end-user involvement, a database of 6000 creative end-users (individuals) from 18 years of age and older in Sweden, a large partner network including SME´s, public bodies, large industry and other research organizations. Our Key capabilities are, among others, methods for idea-generation with end-users for new solutions, user-interface testing, efficient methods and research expertise for planning and performing user-involvement activities, innovation process management, professional management and performance of large-scale pilots in real life setting including both technology and involvement of the entire value-chain of the actual solution being tested.

[2] Exctracted from D8.1 and http://www.ltu.se/centres/cdt/Vart-erbjudande/Experimentmiljoer/Botnia-Living-Lab-1.111199?l=en

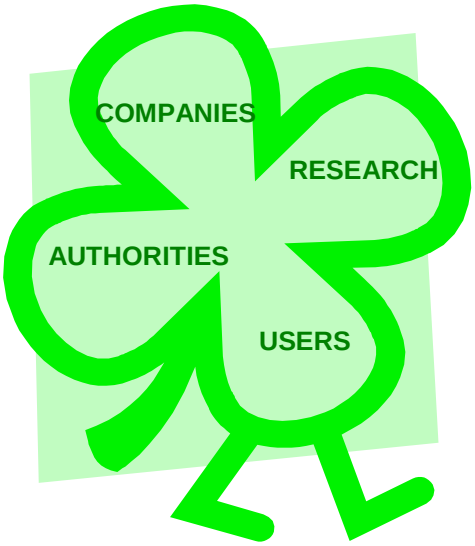


Figure 1 Botnia Living Labs Stakeholders

Botnia Living Lab is hosted by Centre for Distance-spanning Technology (CDT) at Luleå University of Technology (LTU). CDT is a research, design and innovation joint venture between Luleå University of Technology and the IT industry with the main objective to generate sustainable business innovation. Social Informatics is the main research partner of Botnia Living Lab. Social Informatics is a design and innovation oriented research subject that focuses on sustainable life in the digital society. The group have more than 20 years of experience from research and development within the field of user-centred and user driven service innovation. Within this area they have a particular interest in milieus for innovation and design of digital services such as Living Labs and platforms for crowdsourcing.

Test Panel

Over the years, Botnia has built up a community of end-users that it easily can communicate with. In this community, approximately 6000 test pilots are accessible, and the test pilots are represented as individuals who have chosen voluntarily to be part of the Living Lab community. The users in this community are motivated to participate in technology innovation based on their curiosity to try new technical artefacts, and to get the opportunity to influence them. What separates them are demographical and psychosocial factors (Ståhlbröst, 2004). Botnia does not include only users from the database in its projects; if users with some specific character are needed, these are recruited specifically. For example, if the aim of an innovation project is to create mobile services for rural areas, people living in those areas are contacted. The user perspective is that they are involved as partners with the right to exit from the process whenever they choose; they are not bound by any contract.

When users are involved in Botnia activities, the aim is to involve them in the whole innovation process that should be grounded in reality. This means that each innovation process and its methods are customised in accordance with the unique requirements for its particular situation. The aim is to involve users in their natural environment by means of technology, with the objective to gain access to users' needs, ideas, and attitudes in their current situation. Due to Botnia's focus on products and services to support a mobile life, the circumstances in which the user involvement processes are conducted become multi-contextual in character. This means that the users can be involved, for example, in their homes, when they walk around the city, when they drive a car, or when they work. Hence, the methods applied in Botnia operations needs to handle this multi-contextuality.

3. Theoretical Foundation

3.1. Genres Theory

The concept of genres of disclosure, as the name implies, corresponds to a class of genres where disclosure is conceptualized as a type of communication. In order to understand the term better it is worthwhile to look more deeply at genres and disclosure.

Genre (from French genre and Latin genus) means “kind” or “sort” and dates back to the ancient Greek as a classification scheme for the literature. Genres transpired in disciplines and paradigms to interpret human interaction (with the world or human-human) and/or products derived from it (e.g. visual arts). A person acquires language in a patterned way through various genres he is exposed to (Caballero, 2008), thus it has shaped our interpersonal abilities in such a way that without it, knowledge of other sorts (e.g. linguistic knowledge) is insufficient for successful interaction (Tomasello, 2010). Genres allow us to recognize different items based on their similarity of content and form. Content refers to motives, logic and themes presented in a communication and Form is a standard unit of communication shaped by linguistic and physical features (Yates & Orlikowski, 1992). For example through genre lens, a movie categorized as western is a type of movie that is clustered to a certain family that share common features. Although movies of the same genre are different from each other, it makes the comparison of each individual movie much easier.

Disclosure from its literal meaning is defined as the act of uncovering secret information.. It can be viewed from two perspectives: self-disclosure and unwanted-disclosure. Self-disclosure involves an individual to willingly provide information about the self to others (Jourard & Lasakow, 1958) that becomes common knowledge existing between people, within groups or between an individual and another party like an organization. Self-disclosure is seen as a regulator for dynamic interaction which is both the product and process of communication encounter (Ioinson & Paine, 2007). Unwanted-disclosure refers to access of a third party to user’s information without the user's consent like various types of hacks leading to privacy leakage.

Combination of genre and disclosure therefore refers to types of disclosure that share the same content and form. Genres of disclosure was first debated by Palen and Dourish (2003, p. 133) in 2003 as “socially-constructed patterns of privacy management,” or “regularly reproduced arrangements of people, technology and practice that yield identifiable and socially meaningful styles of interaction, information, etc”. Central to the concept of genres of disclosure is the adoption of social patterns of expectation and response into recognizable, socially meaningful forms of interaction and information disclosure that genres embody. Social and technical practices will guide and/or affect the social expectations of participants involved in a genre leading to arranging one’s patterns of privacy managements. Genres of disclosure draws attention to the communicative practices involved in a system to insinuate about the expectations of use according to the users; therefore, designing privacy management in a system keeps up with the promise of genre (i.e. expectation of use). For example disclosing credit card information to an online store during check out is a commonly understood type of communication that differs from traditional ways of paying (e.g. with cash in a physical store). It requires a user to reveal some digits, name and last name via a computer mediated device. This genre of disclosure (online credit card payment) raises concerns about the usage of this information in which failure to meet the user’s expectation

will guide the user's privacy managing arrangement to either corporate or defy with that genre (i.e. to accept the norms of the genre and use the service or refuse to use the service). From the system designer's perspective, providing mechanisms that are aligned with the user's expectation of its use will ensure that disclosed content will not be misappropriated and used unpredictably. In USEMP pilot studies we are aiming to use genre theory and its analytical lens as a means to scrutinize the communicative patters of disclosure and to evaluate the USEMP tools based on those genres to see how USEMP platform will be able to transfer the user's regulatory patterns of disclosure and privacy management into the online environment.

3.2. Needs

Understanding users' needs and requirements plays an important role in development of any IT system in order to improve the way people will adopt and enjoy using it. It is therefore commonly known that developers of IT systems must be able to have a holistic understanding of the users' current situation, the actions the users perform and the aim they have with their actions. The starting point in facilitating this users' adaptation of an IT-system is the cumulative task of collecting knowledge gained from the users' goals through enquires that form a set of system requirements. This process which strives to identify user needs is known as "Need-finding" (Preece, Sharp, & Rogers, 2015), which continues during each and every phase of system development and design process (Bergvall-Kåreborn et al., 2008; Bergvall-Kåreborn & Ståhlbröst, 2007; Bergvall-Kåreborn & Ståhlbröst, 2008; Ståhlbröst & Bergvall-Kåreborn, 2007, 2008; Ståhlbröst & Holst, 2006, 2006). But what is a need really?

Starting with the confusion as to what the concept of need stands for and its relation to closely linked concepts, this has been discussed by a number of authors (Bergvall-Kåreborn et al., 2008; Ericson & Ståhlbröst, 2005; Hyysalo, 2003; Oulasvirta, 2005; Vidgen et al., 2004). The main conclusions drawn by these authors are that we need to define and separate more clearly the related concepts, and that we need to shift our focus from requirements to needs due to the previously mentioned benefits of focusing on needs.

Among the authors who do talk about what a need is, though often implicitly, needs are related closely to motivation and "underlying rationalities" (Bergvall-Kåreborn, 2002; Vidgen et al., 2004). Tiitta (2003) talks about "motivational needs" and Mumford (1981) talks about satisfaction. Salovaara (2004), claims that a need is the goal that a user wants to achieve by using a product. Oulasvirta (2004) categorises needs into two types of human needs: motivational needs and action level needs.

- Action level: Action-level needs define what kind of behaviour users are interested in and in what kind of context (Kankainen, 2003).
- Motivational level: Motivational needs rationalise and motivate action in a context and provide a starting point for discovering design opportunities on an individual level. There are two types of motivational needs: basic and quasi.
    - Basic needs: some related to regulating bodily homeostasis (physiological needs), some related to providing psychological nutriments for growth and healthy development (self determination, competence), and some preferring some aspects of the environment rather than other (social needs such as achievement, intimacy, power, and affiliation).

- Quasi needs: these are more ephemeral, situationally induced wants that create tense energy to engage in behaviour capable of reducing built-up tension. They are not full-blown needs in the same sense as basic needs, but they have influence on how we act, think, and feel (Kankainen, 2003).

Both basic and quasi needs are instantiated in a given situation in which users eventually want to perform a certain action that takes them closer to satisfying motivational needs.

The concept “user needs” is often mentioned among authors in systems development as something important to gain knowledge about. However, the identification of needs and the establishment of requirements is not an easy, straightforward process with a defined start and end. Most of the times needs are not obvious and their identification is hardly captured through observation and enquiry summarized into scenarios and examples. Some examples include difficulties for users to articulating and expressing their needs (Holst & Ståhlbröst, 2006; Robertson, 2001), immediate needs are obscured by lack of understanding of the current situations’ needs (Hyysalo, 2003; Salovaara, 2004) and needs could be easily forgotten to be expressed (Kano 1984). Hence scholars like Robertson (2001) argue that some important aspects of user needs are to be neglected, should we expect the users to open up and express their needs.

Users are also sometimes accustomed to their own ways of doing things and are interested in finding their own solutions to the constraints based on what they think is technologically possible. Thus the real needs and requirements could be buried under illusion of what they have solution for. To stimulate the process of gaining insights into users’ situations and their needs, it is useful to give the users something to relate to. When users gain more knowledge and insights into possible solutions, they also expand their needs (Dennis et al., 2002).

Hence in this study we stress on need finding as the heart of analysis and design process by focusing on tools and methodologies that help us look more deeply into ways that the final product could benefit from ways in which users could fulfil their needs and open up needs and requirements.

3.3. Motivators

Users want products or services that improve their quality of life and work and that is what motivates the users to buy and use a specific product or service (Gerstheimer & Lupp, 2004). In other words, the user is mostly interested in his or her individual benefit, hence the possibilities for profitable applications and services and for success in the increasingly competitive market can only be sustained by knowledge of the users’ needs and motives (Gerstheimer & Lupp, 2004).

Everyone experiences motivator by varying degrees. Humans have basic goals that are fulfilled through different intrinsic desires. As a matter of logic, we value that which we desire for its own sake; therefore, the list of “sixteen basic”[3] desires can be reworded as a list of sixteen fundamental values. The experience of a basic goal produces an intrinsically valued feeling called “joys”, and the specific joy is different for each basic goal. Much of what people

[3] Please refer to D8.1 for the complete list

do seems aimed at satisfying these sixteen basic desires (Reiss, 2004). When a basic desire is satisfied, we experience the joy. We feel secure, for example, when we are in an environment with the degree of stability and order we like. We experience love when we spend time with our children and satisfy the desire for family. The satisfaction of each basic desire gives rise to a different joy, so that we go through life trying to experience sixteen different types of intrinsically valued feelings. Soon after we satisfy a basic desire, the joy dissipates and the desire reasserts itself. Therefore, we seek activities that make possible repeated satisfactions of our basic desires (Reiss, 2004, 2005). Because human motivation is fundamentally multifaceted, the sixteen joys cannot be reduced further into super categories such as pleasure versus pain or intrinsic versus extrinsic motivation and the sixteen basic desires are largely unrelated to each other (Reiss, 2005).

In this study, we are focusing on motivators as a tool for the analysis of the data gathered in the interviews/workshops. In presenting the analysis, after introducing the selection of participants and the process for the interviews, each relevant motivator will be introduced including the respective intrinsic feeling. Each motivator appears differently in relation to the situation; hence, our interpretation of the motivation related to the specific situation will be presented. Each motivator has a name of the motive, a motive, and an intrinsic feeling.

3.4. Values

In social science, values have been a central concept for many years according to Schwartz (2012). Values are used to characterise cultural groups, societies, and individuals, to trace change over time, and to explain motivational bases of attitudes and behaviour. To increase the understanding of what motivates people; Schwartz has identified ten basic values. These values represent what is important to us in life and what distinguishes them is the type of goal or motivation it expresses. All of us hold several values, but the degree of importance differs between individuals and that is what characterizes us as individuals (Schwartz 2012; Reiss 2004).

Below is a description of the basic human values and the interpretation of it in the project (presented in D8.1) together with a defined design principle for users of Internet and OSNs.

Values	**Defining goal**	**Interpretations of values**	**Design principle for USEMP Tools**
Self-Direction	Independent thought and action. Need of control and mastery	When considering this value in design, it is important to let the user take control over their own data and over what they share through the system	Personal settings, ability to control of their historical data
Stimulation	Excitement, novelty and challenge in life	This value highlights the fact that people needs to be stimulated in the use and they need to be challenged and have the ability to learn new things	Visualization of data brokers and encouragement of learning by means of tooltips
Hedonism	Pleasure or sensuous gratification for oneself	Related to this value it becomes clear that the system needs to support having fun while using it.	Gamification of privacy profile settings
Achievement	Personal success	This value puts emphasis on the	User's ability to make more

	through demonstrating competence according to social standards	importance of seeing individual success and to feel competent.	informed decisions on their disclosed information
Power	Social status and prestige, control or dominance over people and resources	In our interpretation of this value, control and social status came in focus. Here the feeling of being important is in focus.	Being able to view and control traces of personal information leakage
Security	Safety, harmony, and stability of society, or relationship and of self	This value sets focus on avoiding anxiety and on making citizens feel safe while using the system.	Showing and informing possible ways users are encroached upon and indicating protective mechanisms in place to ensure security
Conformity	Restraint of actions, inclinations, and impulses likely to upset or harm others and violate social expectations and norms	Related to this value, the importance of encouraging a particular behaviour and restraining another is highlighted to keep a good and sound community.	The systems must not encourage users to see economic value of their data to share more personal information
Tradition	Respect, commitment, and acceptance of customs and ideas that one's culture or religion provides	This value becomes a bit difficult to translate in this context since the focus of the system is to stimulate innovation and change. However, in relation to systems design, we interpret this as making the use of the system as familiar as possible.	Build the system on well-known and broadly adopted platforms
Benevolence	Preserving and enhancing welfare of those with whom one is in frequent personal contact	Our interpretation of this value encourages us to consider a way to socialise in the system both with new and old friends.	Users need to be able to see their friends and share their settings
Universalism	Understanding, appreciation, tolerance, and protection of welfare of all people and for nature	The system focus on contributing to make users more aware of their personal information value and enhancing Internet privacy, hence this value inspires us to think of how to simulate a focus on the common good and how to improve society through peoples compassion.	Give users the ability to share their experiences for sustainable innovations

Table 1

Value can be sought in experience, based on the view that "value resides not in the product purchased, not in the brand chosen, not in the object possessed, but rather in the consumption experience(s) derived therefrom" (Holbrook, 1999, p.8). In USEMP pilot studies the abovementioned human values will be under microscope when USEMP tools are implemented to scrutinize how different factors can become part of a user experience. In section 7.5 we will discuss more about this model in the light of results.

3.5. System Usability Scale

Created by (Brooke, 1996), the System Usability Scale (SUS) became an accepted and widely adopted way to measure users' subjective assessments of a system's usability

(Orfanou, Tselios, & Katsanos, 2015). It was created out of the need for a tool that could quickly and easily (or "quick and dirty", as it was described by Brooke (1996)) collect users' usability ratings of a technology. Not only straightforward for the researchers, but for the participants of the study as well (Bangor, Kortum, & Miller, 2008). Both Bangor et al. (2008) and Orfanou et al. (2015) distinguish several advantages to the application of the System Usability Scale for usability evaluation:

- The scale is very short, comprising only 10 items which need to be ranked on a 5 point-Likert scale, ranging from strongly disagree to strongly agree, which makes it a quick tool that is also easy to implement.
- Bangor et al. (2008) describe the tool as 'technology agnostic' in that it is flexible enough to assess a wide range of interface technologies
- The result of the testing is a single score between 0 and 100. This makes it a tool that is easy to explain to the variety of people involved in the development of new products that do not need to have much experience in usability.
- Tullis et al. (2004) found that the tool performs well with even a small sample size (in their test between 12 and 14 respondents). So the results can be reliable even with a limited number of test users.
- Not unimportant and also partly explaining its widespread use is that the tool is freely available to use, which makes it cost effective.

Figure 2 shows an example of the usability scale.

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

Figure 2 The System Usability Scale (Bangor et al., 2009, p.116)

3.6. Stakeholders and USEMP tool

Stakeholders play an important role in the the ecosystem where USEMP tool can evolve in the future. D3.5 studied different stakeholders in the ecosystem identifying business opportunities and challenges that can be seen to arise when the user has the means to assert more control. Here we will enumerate the identified stakeholders and use the result found from the previous studies in D3.5 to pinpoint existing value network as well as provide some recommendations on how the USEMP tool needs to be incorporated in this study and future pilot studies to make more sense in terms of integration into the value network.

Users

The user is the actor which all other actors' activities ultimately revolve. However, the user is passive, providing personal data and User Generated Content (UGC), eyeballs/attention in return to impression, and money when buying products or consuming services. He/she interacts in the first place with content provided by the publisher (mobile application or online content) and related ads or with an Online Social Networks (OSN) (same attention as for publisher, but also UGC). Then, with an advertiser or a company that is offering a product/service for sale: visiting the advertiser in case ad is interesting, proceeding through the sales funnel and ideally reaching its end, the purchase. The advertiser is also collecting information. Notably, data intensive operations seem to take place "far away" from users. Within the pre-pilot and pilot studies this stakeholder is in direct communication with the pilot team and the DataBait tool. We continue to collaborate with the end-users to facilitate the benefit realization of the USEMP tool and position the tool in the user's habit of daily usage.

Publisher

Publishers are in the business of providing inventory for advertisers in the form of online content or mobile applications. This groups online content providers (e.g. newspaper web sites) and mobile apps, because they are also selling ad space and building their business models on advertising; either in a "free" business model where the online content or application is offered for free to the end-consumer in exchange for advertisements revenue or in a freemium model where there is a "free" version of the service related to advertisement and additional premium features that the consumer has to pay for to access). Publishers are in direct contact with the user, providing content and receiving 'eyeballs' in return. Mobile applications are especially important today, as they add the dimension of location services. Here, location data can be gathered and re-used for targeted and context-specific ads, thus presenting another source for personal data leakage. The publisher's view on the USEMP tool will be the focus of the next round of pilot studies where we will study in more detail how it will affect their business.

Advertiser

The counterpart to the publisher is the advertiser, who provides the advertisements to be displayed on/into the publisher's content. The advertiser may or may not act through an advertising agency that helps to generate and place the ad. Advertisers share data with demand-side platforms (DSPs) and supply-side platforms, ad-exchanges and ad-networks, on target groups and users who have previously been noted as visitors/buyers. They pay the publisher of the ad and the intermediaries for facilitating the process. Due to the complexity of the value network and its deviating value chains, advertisers need sufficient information: about which actor is performing which role, and how successfully they do so. This is also a

reason for the significance of user IDs, as they are necessary for conversion attribution. Just like publishers, the USEMP tool needs to be mature enough to be presented to different stakeholders in order to receive an accurate feedback. Due to the fact that the tool is currently focused on basic concepts, these stakeholders will be incorporated in the next round of study.

Ad Network

The ad network is the oldest aggregator of publishers and their impressions. It buys remnant inventory publishers had difficulties to sell, and sells in CPM[4]. This actor has come about in the nineties, as the number of online publishers increased. As a result advertisers or advertising agencies had difficulties planning and buying from the increasing number of publishers. Ad networks buy remnant ad inventory from publishers. Then the inventory is categorized and sold through the ad network. Perhaps this stakeholder is the most relevant one to get their insights on the USEMP tool.

OSN

OSNs are different from other actors. First, they might perform roles that are very similar to other actors in the value network; to DSP and ad exchange (profiling, targeting) or to publisher (showing ads). Second, they evolve from a single publisher into an ad exchange where additional publishers are added to the impressions inventory they serve ads to; Facebook, Twitter, LinkedIn etc. have first focused on user growth and later on monetisation through data. Social media have an advantage in data in terms of kinds of data and means to coerce consent form data subjects to gather and monetise their data. Contacting OSN such as Facebook has not been successful in the past, however the pilot studies could benefit from collaboration with OSNs to facilitate transparency and raise more awareness among users.

The report D3.5 gives directions on where the USEMP tool should fit into the value network because the ecosystem as a whole suffers from lack of trust and relatively low data quality. Applying the central tool and complying with the rules and standards it sets will then have positive aspects on the customer relationship, as well as on data veracity. In the current report we have been including the end-users within the design and development of the USEMP tool. However, we believe that in order for the tool to better integrate into the market, close collaboration and feedback from other stakeholders is vital. Our aim in D8.3 would be to involve other stakeholders when there is a more complete version of the product ready. As we will present in the section 4, there are a few features of the tool available at this stage of the project which does not make it a good representation of the tool to be assessed by different stakeholders in terms of threats and opportunities. Some features of the tool are not as well developed as others are still lacking. For this reason, we opted for postponing the assessment of the tool from other stakeholders to the pilot studies.

[4] CPM or cost per mille is the cost of a 1000 delivered ad impressions. Other performance based compensation modes are e.g. cost per engagement (CPE) and cost per view (CPV).

4. DataBait Tool

4.1. DataBait Tool Introduction

In the rest of this section we provide an overview of the DataBait tools that have been implemented, tested, deployed, and used for the purposes of the pre-pilot. The tools are described in detail in the following sections of this chapter.

In order to support the pre-pilot activities a number of features have been developed, tested and deployed to the live environment of DataBait platform which is available under the DNS domain **DataBait.hwcomms.com**. These features can be grouped in the following categories

- **DataBait pre-pilot backend services:** a new set of services have been developed to support the pre-pilot in order to collect all the required data from OSN and WEB (currently Facebook is supported from OSNs, as well as collecting data from WEB browsing) and to process the necessary privacy and value indicators so that these can be accessed by the end users that participate to the pre-pilot
- **DataBait pre-pilot frontend/webapp:** a new set of features have been developed in order to gather the collected indicators from DataBait and to make them available to the pre-pilot OSN users. Also, a number of improvements have been implemented in the DataBait webapp flow of information to make the tool more appealing & comprehensible to the end-users and to satisfy a set of legal requirements
- **DataBait pre-pilot webtrackers chrome plugin:** a new set of features that have been developed at the chrome plugin that allows OSN & web users to block 3rd party tracking and to communicate related piece of information to the DataBait pre-pilot backend services

The described features have been developed and tested in two separate environments before they were deployed into the live environment prepared to support the pre-pilot activities: a) features have been developed and tested initially in mock development environments from the implementation team (CEA, CERTH, VELTI, HWC) and b) features have been deployed and have been tested end-to-end internally in a **staging** environment operated for that purpose only and maintained under the DNS name: DataBaittest.hwcomms.com.

In order to support the necessary tests related to Facebook functionality in the staging environment the testing facilities from Facebook.com have been used (using a test Facebook application that can collect the necessary data and support test operations in DataBait staging environment.

4.2. What will be tested in Pre-Pilot - DataBait workflow

This section will outline the workflow of the backend systems which occur when a pre-pilot participant signs up to use the DataBait service. The predominant elements available for testing during the pre-pilot phase are as follows:

- DataBait account creation and sign-up.

- Social media account linking and data retrieval services.
- Concept detection imagery analytics.
- Location detection textual analytics.
- Visualisation elements and user interaction.

On creation of a DataBait account, the system will then instruct the user to link their Social Media account (during the pre-pilot phase only Facebook is supported, however the architecture is designed to be extensible to support additional networks). On linking the backend system will issue an instruction to fetch the users profile data, posts and images subject to the permissions the user has authorised for the DataBait application.

Textual data and imagery data is fetched in parallel, and is able to process multiple users simultaneously up to available network bandwidth, if this nears saturation point, future requests are queued on a first come, first served basis. From time to time, an update will be issued to ensure data held on the system is kept up to date, provided the user remains active on the DataBait platform.

Textual data is processed in-situ with the location detection module as it is streamed from the social network into DataBait servers, and is then stored with the additional metadata within the backend systems.

Imagery data is first collated and stored independently of textual data, and is processed in a batch once all imagery data has been retrieved, using the concept detection modules.

Textual data is available as soon as the data enters the backend system, and can be presented to a DataBait user right away. Imagery data takes longer to fetch and process, and therefore has a noticeable latency before results are able to be displayed to the user.

On completion of data processing, the visualisation and front-end user interaction elements are then free to query the back-end systems and report findings to the DataBait user. These user interaction elements can be seen in the next section.

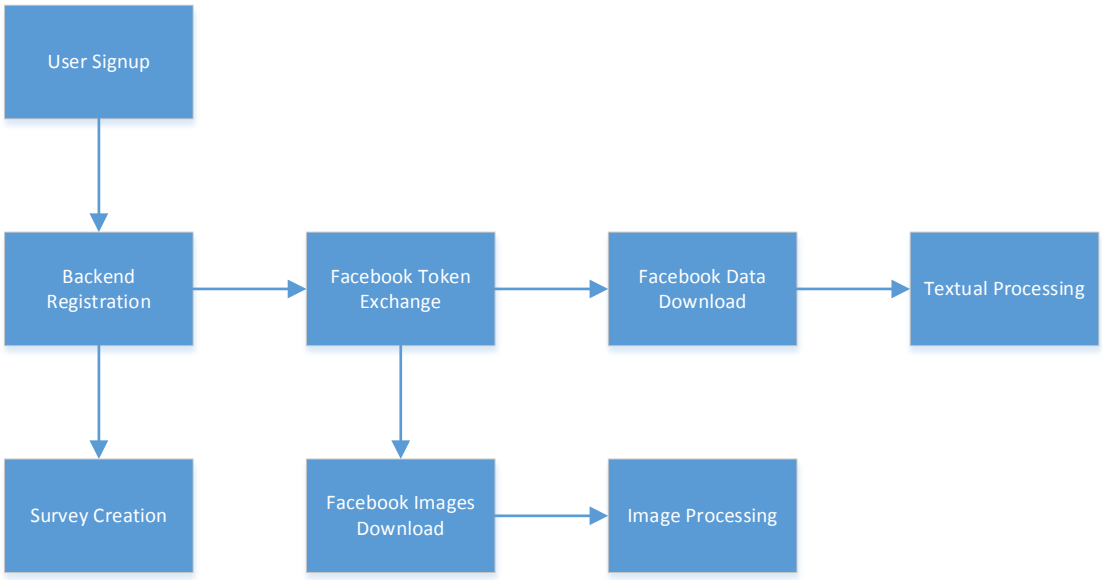


Figure 3 DataBait System Work Flow

4.3. Pre-pilot DataBait components

In this section we describe in detail the components/features that have been implemented, tested and deployed to support the pre-pilot activities. We have grouped the pre-pilot features in terms of the tools (webapp or chrome plugin) & and the data sources (Facebook or web online behavior) that are provided to the OSN & web user.

4.3.1 WEBapp tools to support registration & DLA

For the pre-pilot purposes a number of improvements have been implemented as part of the DataBait webapp available at: http://DataBait.hwcomms.com to support: a) the registration flow from pre-pilot users, b) access of the pre-pilot users to informational material explaining DLA (Data Licensing agreement) and how users' data are processed by project partners, c) facility for users to opt out their data from the pre-pilot if they wish it after they have registered. More specifically:

a) for the registration flow a check has been implemented requesting pre-pilot users to validate if they are older than the 13th year of age (see Figure 5 below)
b) for the registration flow a new DLA review flow has been implemented for users so that they can view all the chapters of USEMP DLA before they sign up for the pre-pilot (no skip option, see Figure 6)
c) a new information section titled : "what, why, how" or DataBait-at-a-glance section that explains DataBait features to the end user including a new set of infographics, text and video
d) update of the UI design to better support mobile and tablet devices
e) update of the opt out/withdraw data option
f) update of all the necessary info for the navigation in the webapp tool in three languages (English, Swedish, Dutch/Flemish)
g) update of all the web analytics configuration so that data can be collected during the pre-pilot

We provide some indicative screenshots from the pilot environment for the registration & DLA related features.

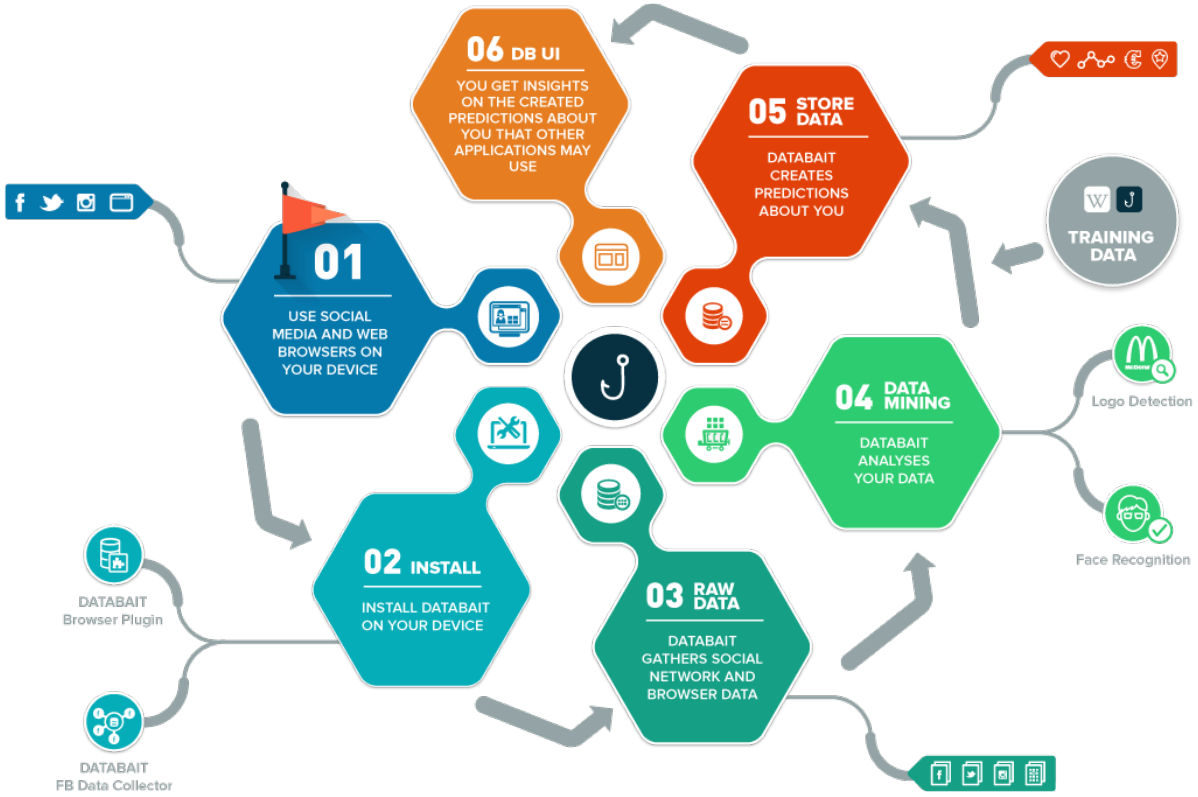


Figure 4 Infographic explaining the process of DataBait for evaluating users' data implemented as part of the DataBait-at-a-glance section of the pre-pilot

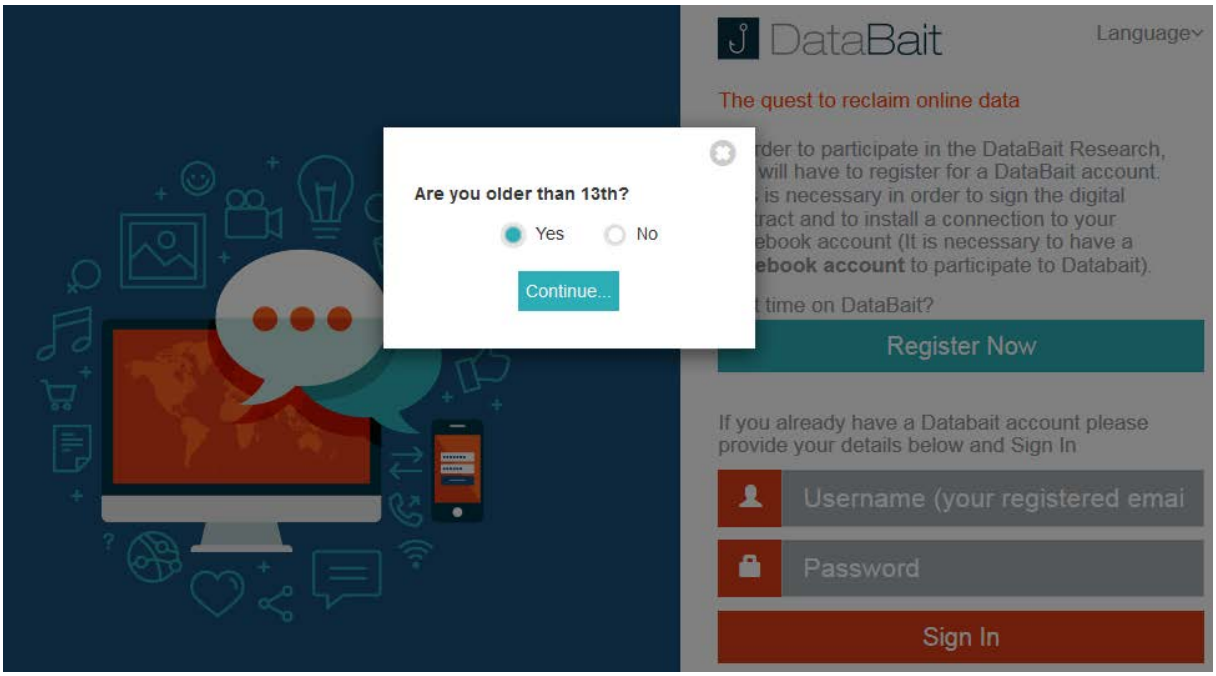


Figure 5 Control Flow for age of consent (13th) during pre-pilot registration

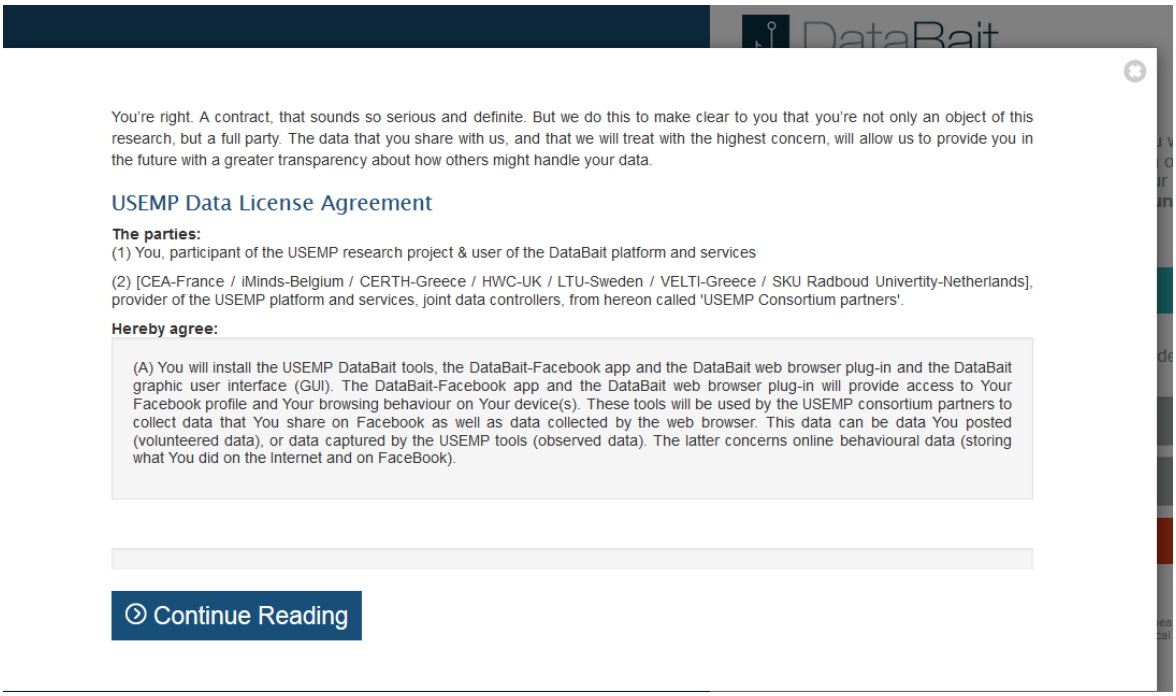


Figure 6 New DLA review process (excluding skip options) during registration for the pre-pilot implementation

4.3.2 WEBapp tools based on Facebook data

For the pre-pilot purposes a number of features have been implemented as part of the DataBait webapp available at: http://DataBait.hwcomms.com so that end users are provided information related to the data that they have shared with OSNs (Facebook in this case).

Overview section: Once the DataBait pre-pilot users have signed in to the DataBait webapp they are presented with a summary of info they have shared on Facebook focusing on privacy related historical data shared by the user with Facebook: the latest video the user has shared & the latest place a user has been tagged in Facebook (see also Figure 7 below). The text and design has been updated for the pre-pilot purposes.

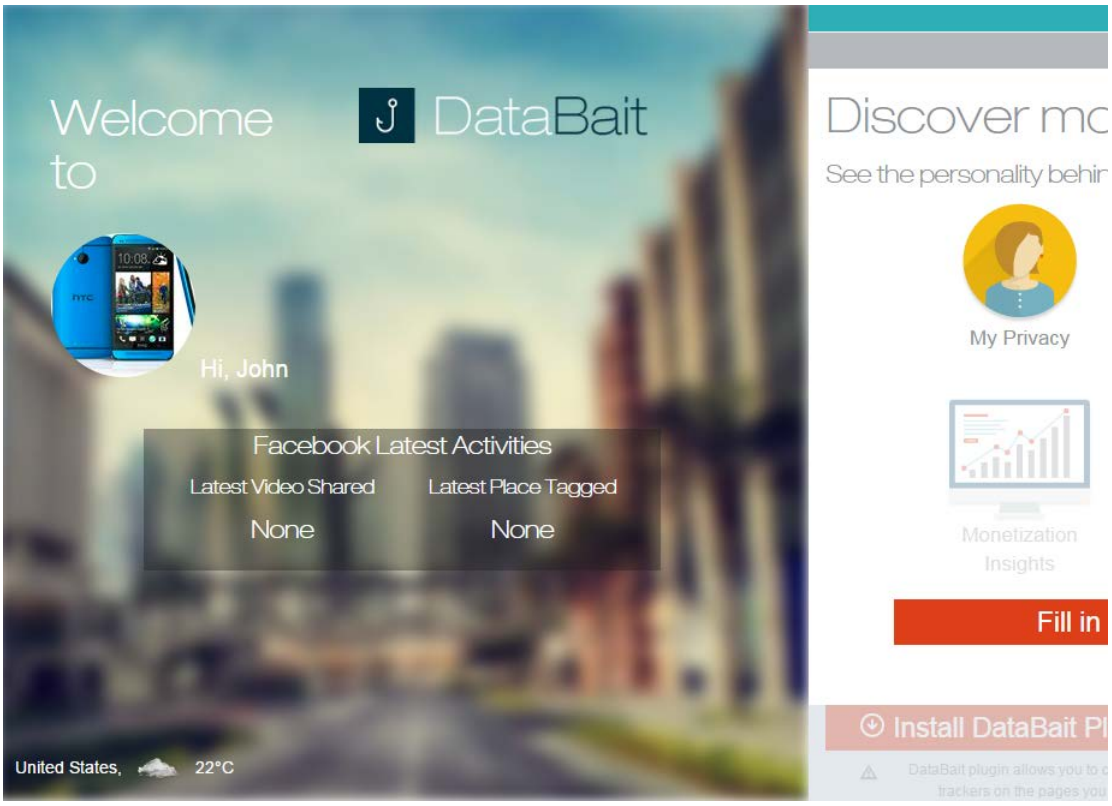


Figure 7 Screenshot of the pre-pilot DataBait webapp page after the user has signed in, at the left part of the page there is a summary of information related to the user's latest Facebook activities

My Privacy section: A number of tools have been integrated from the DataBait backend services. KPI (key performance indicators) values provided by these services are provided to the pre-pilot users based on the processing of their Facebook data, related to user privacy and value of data shared. The following is a summary of the KPIs and respective tools integrated under the My Privacy section:

Image Leaks/Visual Concept Mining Tool: In this subsection of MyPrivacy the user is provided with the list of concepts that can be inferred from the images the user has shared with others in Facebook. These visual concepts can act both as an indicator of privacy (what algorithms can infer for a Facebook user based on the images he/she has shared) and value (detected appealing visual concepts can be used for targeted marketing & advertising from interested brands). The visual concepts can be automatically inferred by DataBait from the images that you posted on Facebook. DataBait predicts tags from a set of over 17,000 visual concepts. The concepts are visualized using tag cloud visualization in which the tag cloud shows the identified concepts with a size proportional to their frequency in the posted online social network images. If the user selects a concept the images where this concept has been detected are shown along with a measure of confidence for the detection from the corresponding algorithm. A screenshot of this function is illustrated in figure 15.

Location Prediction: In this subsection of MyPrivacy the user is provided with the list of locations that can be inferred from the posts the user has shared with others in Facebook. The locations detected are the result of an automatic location estimation algorithm that processes the text content of user Facebook posts and tries to predict the location where these posts refer to or the location indicated in Facebook by the

user. The tag cloud shows the identified locations at the city level with a size proportional to their frequency. Tags are coloured with different colours to indicate how the location was detected (Facebook explicit information or inferred information) If the user selects a location the posts where this location has been detected are shown along with a measure of confidence for the detection from the corresponding algorithm.

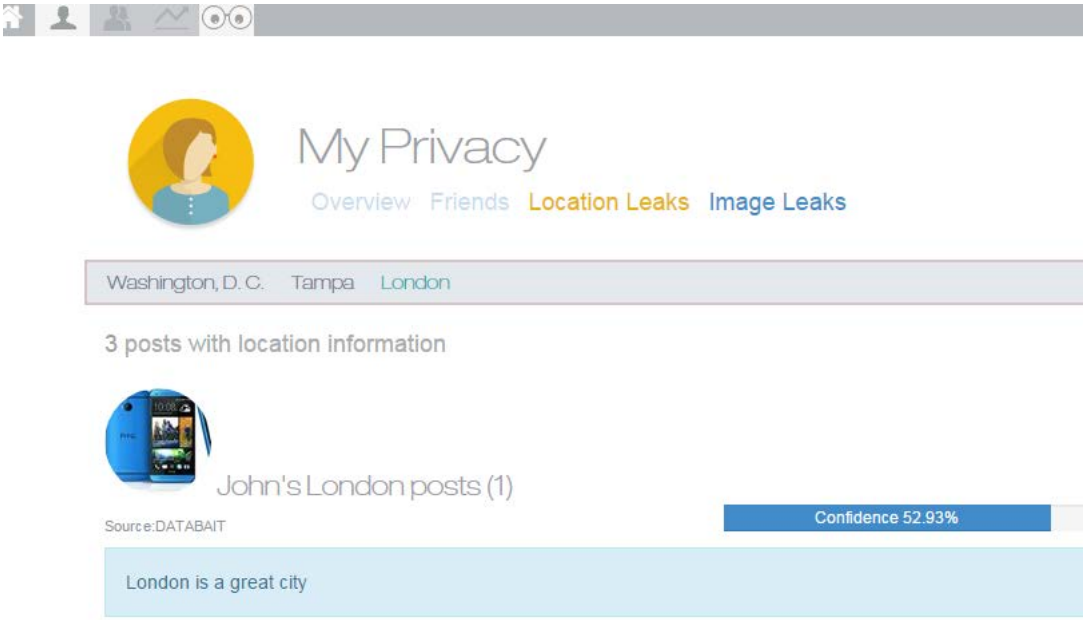


Figure 8 Screenshot of the MyPrivacy->Location Leaks page after the user has selected a specific location (London) to view additional information related to it

4.3.3 WEBapp tools for User trackers – based on online/web data

For the pre-pilot purposes a number of features have been implemented as part of the DataBait webapp available at: http://DataBait.hwcomms.com so that end users are provided information related to the data that are tracked from their behavior while they browse web content online. This section of the webapp requires that the user has successfully installed the DataBait browser plugin (see section below).

Some information of interest is displayed to the pilot user who has installed the DataBait browser plugin:

- set of tracker services for each visit site in two different views: a) graph view where sites and trackers are represented as nodes in a graph, b) list view where sites and trackers are represented as members of a list
- information on the amount of time each user has spent on each site
- information on which of these tracker services the end-user has selected to block

In addition to the information presented, the pre-pilot users have the ability to select a tracker and flag it as to be blocked one specific site. This information is stored in pre-pilot user profile and utilized by the DataBait browser plugin to enforce blocking (see also next section on the operations supported by the browser plugin).

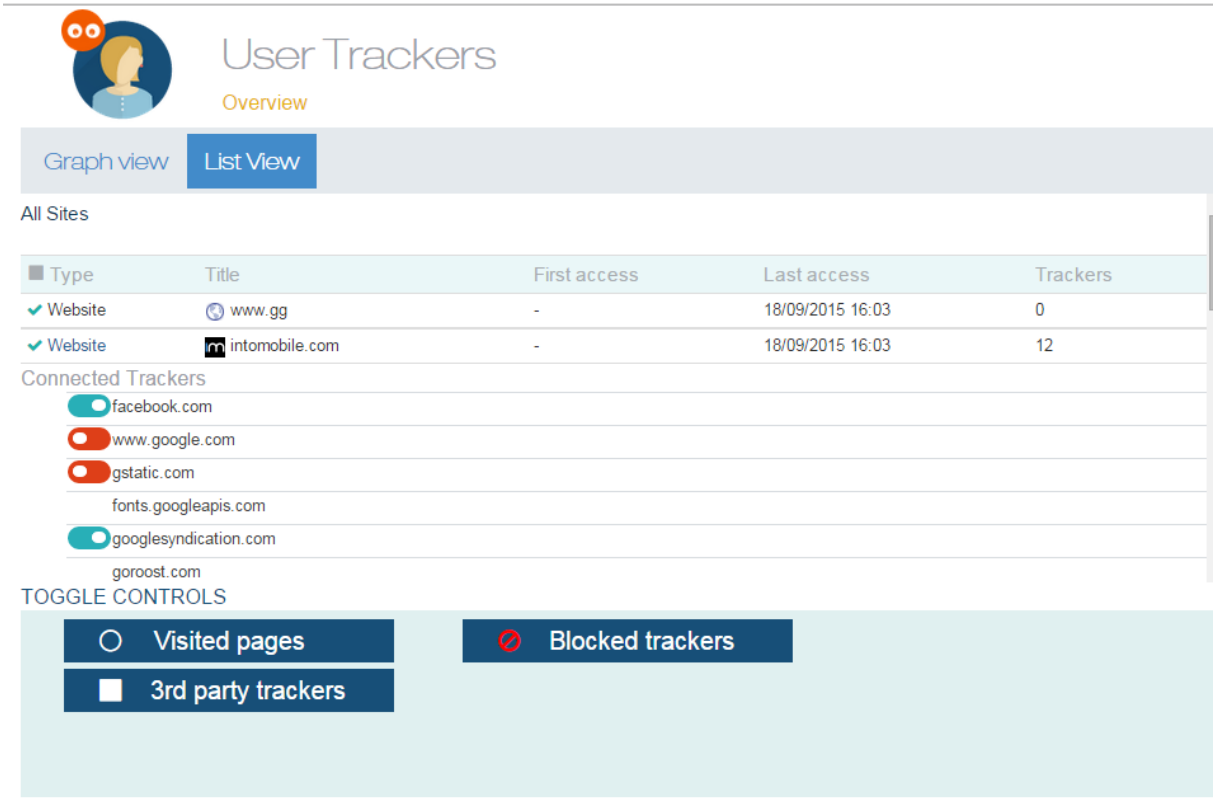


Figure 9 Screenshot of the List View for User Trackers subsection of the pre-pilot DataBait webapp; the list of connected trackers for website intomobile.com are depicted along with user selection to block two of these trackers

4.3.4 Chrome web-trackers plugin

For the pre-pilot purposes, the DataBait browser plugin for Google chrome browser has been implemented. The pre-pilot users have to go through the process of installing the DataBait browser plugin as a developer extension in their web browser for the related functionality to become active. The following features have been implemented for the browser plugin for the pre-pilot:

- pre-pilot users can view in a small add-on view the user trackers for the web site they are currently browsing (see Figure 10 below)
- pre-pilot users can select to block any of these trackers on the current page
- pre-pilot users can view any updates that they are applying on which trackers should be blocked from the related DataBait webapp tool for user trackers
- DataBait browser plugin can apply blocking of trackers in selected pages (either through the plugin itself or via the webapp tool for user trackers)
- DataBait browser plugin stores in DataBait backend the information of which trackers are blocked and which pages are visited for further analysis (privacy & value scoring)

Before using the DataBait browser plugin it is necessary for the pre-pilot user to sign in with his/her pre-pilot credentials. The plugin has been tested extensively prior to the pre-pilot deployment in different operating systems (MacOSX, Linux, Windows 7, Windows 8) and versions of the Chrome web browser.

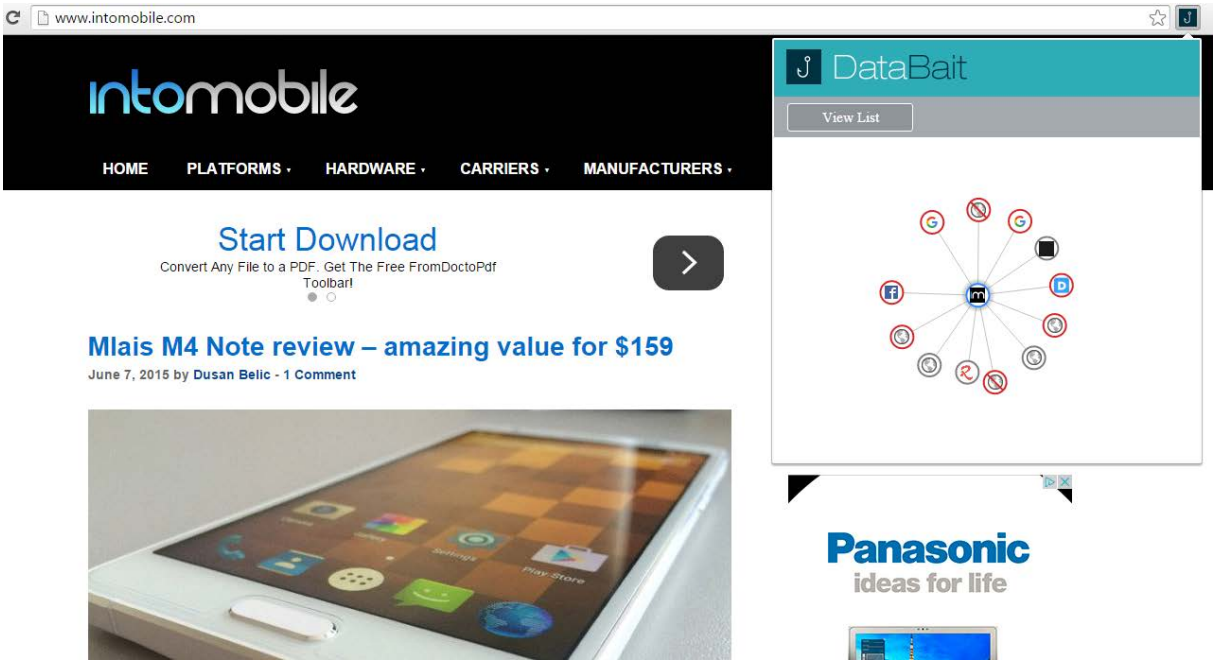


Figure 10 Screenshot of the DataBait browser plugin in operation; the address bar icon is highlighted and the plugin view is opened listing the trackers for the visited site (http://www.intomobile.com), two (2) of which the user has selected to block (indicated with an overlaid stop sign)

4.4. DataBait deployment architecture

The DataBait system is deployed in a two tier configuration with initial deployments being pushed to a test configuration allowing for new features to be tested and debugged prior to moving them over to the live system for use within pilot events. Due to the nature of hardware requirements, some elements are shared across both the live and test systems. This is a requirement as the imagery analysis module requires use of GPU hardware only available on a single server. Processing for both live and test servers for imagery concepts therefore take place on the same machine, albeit with data segregation for information running on live and test systems.

All connections to DataBait live and test, are over a secure HTTPS connection which is terminated at a reverse proxy. This proxy then delegates connections to the associated backend server. No direct connection to backend storage systems is possible from the internet side of the system.

In addition, DataBait specific account information is kept segregated from textual social media information, which is again segregated from imagery data. This segregation allows for data to be kept close to those machines processing the data and limits exposure of such machines to the internet meaning all access is restricted provisioned DataBait system calls, managed by a security framework.

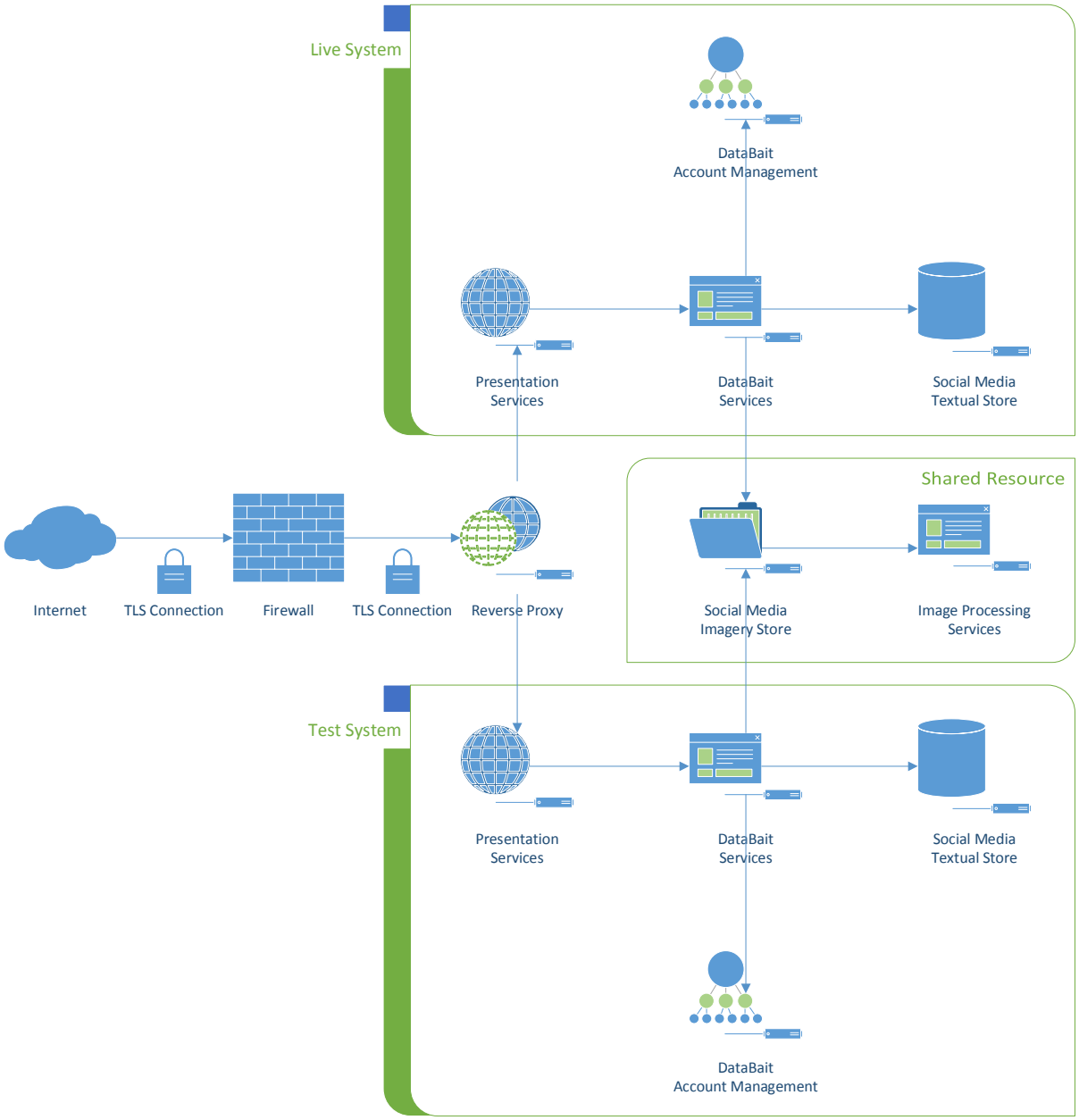


Figure 11 Server Deployment Diagram

5. Pre-Pilot Method

5.1. Planning

To support knowledge sharing and to get a coherent view of the context in which the pilots are implemented, it is important that the pilots are described in some detail. In this phase the focus is to understand the basis for the user studies and the purpose of them.

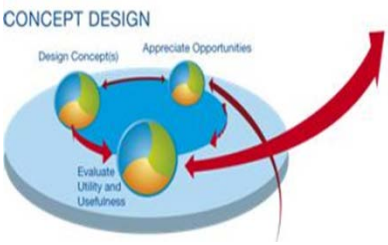
Each phase needs to start with planning and with discussion and answers to these areas:

- Purpose, questions, methods
- Type of users, number of users etcetera
- Time period and number of interactions
- Motivations and incentives
- Technical equipment needed
- Competences and other recourses needed
- Ethical considerations
- Context (social/technical/organisational/physical)

And as also discussed Key Principles of Living Labs operations as well as the 10 I's create the basis for design and are implemented in all activities.

Working in an iterative and interactive way always affects the planning. Therefore the first plan is not in too much detail in terms of methods and time slots for different activities. This is outlined in more detail for each phase depending on the results in previous phases.

5.2. USEMP Cycle one – Pre-Pilot – Concept Design



Moving into the first cycle we focus on utility and usefulness. In USEMP we will do this cycle in two iterations. First we worked internally with needfinding and design of concepts during the start-up of the project. Results from this have been outlined in D2.1, D3.1 and D4.1 as well as D7.1. The second iteration will be external with real users and is outlined in the following. The time-plan for the second iteration is to:

- Planning pre-pilot in detail (November/December 2014)
- Running Pre-pilot (August 2015)
- Analysing results. Summarizing and feeding back to developers. Pre-pilot evaluation report April. (August-September 2015)

The planning is guided by chapter 5.2.2 cycle 1 and this includes;

- Detailed plans for the appreciating opportunities process.
- Detailed plans for the design of concepts process.
- Detailed plans for the evaluation of concepts which since we are also working in cycle two include appreciation of opportunities for the next phase (cycle).

Hence, each step of the process needs to be designed. In the first cycle we will work in all three phases; appreciate opportunities, design concepts and evaluate concepts. A draft process could be as described in the following sections.

5.2.1. Appreciate Opportunities with users

In the first cycle and first phase need-finding will be conducted through user-pool brainstorming sessions where the users will tell their needs, problems and give ideas and suggestions on privacy enhancing tools, this phase also includes different identified stakeholders, such as developers and providers of OSN. The sessions will be conducted in both Living labs, and with the best practice Living Lab methods. The user-pool suggestions (and solutions) will be documented, compared and packaged. In the next phase these concepts will be elaborated and tested further. Consequently we need to;

- Make plans for the process for this phase
- Prepare material
- Decide which users should be involved
- Recruit people for the groups to interact with
- Perform the workshops and document results
- Summarize the collected data and report to whole project group

This step is the process of generating and understanding users' needs in situations where people carry out, for them, meaningful activities with the objective to improve the situation as a whole in relation to privacy in OSN. It is important to separate between requirements, which are related to a solution or artefact, and user needs that are subjectively experienced, and context dependent.

5.2.2. Design Concepts (by project partners)

The design phase is also the most innovative phase in the concept design cycle since this is where all collected data is clustered in different ways and viewed from different perspectives with the aim to construct concepts that represent users' needs.

First versions of the DataBait concept have been developed in the first iteration of this phase. In this second iteration we will use the designs as stimuli at the same time as they are evaluated and new ideas are generated. We need to:

- Plan the process for this phase
- Decide the number of concepts to develop
- Decide who should develop the concepts (all partners?)
- Decide how to design the concepts (films, narratives, mock-ups, etc)
- Designs concepts based on the results in phase 1
- Document designs
- Present designs to all partners in the decided form

5.2.3. Evaluate Concepts

Methodology

In this section we present the methodology used at the concept evaluation stage. Following the FormIT (Bergvall-Kareborn, Holst, & Stahlbrost, 2009) methodology presented in D8.1, the first phase in the methodology deals with the artifact at the conceptual level where researchers and designers interact with each other to design conceptually through iterative Appreciate Opportunities, Design Concept(s) and Evaluate Utility. In order to build the DataBait tools on solid grounds of user acceptance, LTU conducted three focus group sessions with users. Our main motivation for this method was to be able to explore one specific theme, in this case strengths/weaknesses of DataBait tool in relation to daily Internet and social media use. The objective of those qualitative studies was to set the basic pre-pilot evaluation criteria and help us design experiments accordingly. In a setting that stimulated participants respond to and act on each other's views, new insights and perspectives were expected to emerge more easily than in an individual interview setting. Such a strategy would foster more discussion, allowing new ideas, challenges, problems, solutions and etc to cultivate. Also, a discussion in a group would challenge the participants to explain and verbalize their views very accurately (Bryman, 2012) in order to get their point across, allowing us to capture the sensitivities and nuances of what they really meant.

In concept evaluation study we employed a formative evaluation approach trying to understand why an artifact works or not and what factors are affecting those. This approach gives value to the life cycle of an artifact by investigating through a qualitative method, with exploration of how well a product works and what experience attributes are associated with it to address those issues in the next round of design & implementation of the DataBait tool. In this approach more "open questions" were used which served towards understanding of all aspect of the tool use and adoption (Goodwin & Cooper, 2009).

Concept evaluation design

Three focus group interviews were conducted in English within 25th February and 4-5 March 2015. The focus groups were in Luleå and LTU's main campus with participants of different backgrounds. In total 12 people participated to the focus groups along with researchers of LTU working in USEMP project as the moderators of the discussions.

To recruit respondents, we advertised a call to participation in LTU's student web portal. We deemed to not limit ourselves to student participants. Belanger, Hiller, & Smith (2002) argue that privacy research is being heavily reliant on student-based samples and this need to be

alleviated. For this reason we balanced the participants by inviting the researchers and other people not related to academia. We informed all participants that all information gathered during the discussions is analyzed based on the themes and individual responses would be anonymized if to be quoted. The duration of a session was on average 90 minutes and took place in one time slot of afternoon.

To leave sufficient room for discussion, we opted for semi-structured focus groups. A script was prepared, mentioning the major topics and some key questions we could ask in order to spark the discussion. The focus group interviews were structured as follows:

1. Short introduction of the facilitator, practicalities of the session, explaining the reason for recording of the session, acquiring their consent with recording and the general outline of the focus group
2. A chance for the participants to introduce themselves and to explain the motivations behind the participation
3. Introduction to the USEMP project, consortium and roles of LTU within the project
4. General discussion about the social media use, sort of information users disclose, privacy concerns with respect to social media and Internet browsing
5. We briefly introduced the DataBait tool from the registration, linking to Facebook account and the main interface. All presentations were only screenshots of the tool with no active interaction.
6. Introducing the main features of pre-pilot with mock-ups of:
    a. multimedia function
    b. location leaks
    c. trackers

    Each function presentation was then followed by questions targeted on insights for values, motivations and barriers to use.
7. Closing general level discussion focused on reflecting upon the values of DataBait tool with all functions together

The full script of concept evaluation focus groups can be found in the Appendix at the end of this report.

Table 2 shows information about the sessions:

Focus Group	**Date**	**Participants**	**Duration**	**Location**
1	25/02/2015	5	1 H 47 M	Luleå
2	04/03/2015	4	1 H 47 M	Luleå
3	05/03/2015	3	1 H 18 M	Luleå

Table 2

Analysis

To analyze the transcribed data material from the focus group interviews, a qualitative content analysis was used. Graneheim & Lundman (2004) present qualitative content analysis as following:

- Repeated reading of the collected data will be made to get an overview and sense of the material.
- Identifying meaning units (minimum identifiable unit with the same meaning, e.g. words, phrases), which responds to the aim.
- Condense meaningful units into codes with maintaining content.
- Grouping codes into categories and subcategories, alternative themes and sub-themes
- The categories/themes should represent the content and answer the aim, without omitting any essential or include any irrelevant.

The three workshops were transcribed from the audio recordings captured during the sessions. The transcriptions of the records were done by one of the employees in LTU. The transcripts were double checked by one the authors of this report. The analysis of the transcripts was conducted through NVivo 10 for Windows. NVivo makes this process more automated to give text meanings in the form of nodes or concepts and classify the interviews based on different classifiers. The association between nodes creates higher level categories which facilitates detection of trends across the collection.

Result of Concept Evaluation

The analysis of the focus groups revealed several important core concepts. Awareness showed to be an influencing factor when it comes to online privacy. Another factor which has impact on awareness is the benefit realization. These two factors need to correlate together and communicated in the DataBait tool to convey the right message to its users in order to enhance the usability of the tool. Each of these core concepts are discussed in following sections.

Awareness

We asked users about the normal social media and internet usages in order to indirectly and directly capture their awareness of the ways they disclose information. The indirect way was to ask them to openly explain one normal day of their online behavior in terms of interactions through the social media and internet. During the course of a normal social media usage there are different communicative actions which are established. Users have personal motivations and external influencing factors that force them to use social media. Therefore their usage is not totally optional and for this reason their personal information is inevitably disclosed. Among personal motivational factors that can be enumerated are the willingness to reach a wider audience in order to promote themselves for example with different political activities, keeping in touch with the families, friends, acquaintances, keeping track of their events like their friend's birthdays, to keep themselves updated about what is happening in surroundings and etc. However there are other trends of users who despite their lack of motivational considerations (e.g. awareness of institutional privacy, not willing to disclose their personal information, bothered by other's post), were forced to use the social media because of their connections in offline world. For example we could see that using Facebook group as part of study groups are very popular which requires students to have an account in

order to have all the necessary information to study. Users who don't like social media but like to be part of these study groups sometimes create fake accounts or provide false information. However their communication with their colleagues is based on mutual communication just like outside social media. Disclosing various types of information is evident in these types of social media usages. We asked them about the kinds of information they think they are disclosing in their everyday usage of social media. The main motivation for us was to capture the level of their awareness towards privacy issues. Our analysis showed that most users think of privacy as only the basic personal information they disclose voluntarily like name, age, relationship status. The awareness towards the observed and inferred data sources is extremely low among the users. Some users have good knowledge regarding the dangers towards their privacy through media, documentaries and TV but they lack enough information on how social media sites could exploit their personal information. They only think that they need to explicitly write or upload something about themselves but they could not express different forms of observed and inferred techniques which could have adverse effects on their privacy. This situation of awareness was the same when users interact and surf over the internet with various inflations. Most users were aware of voluntary data magnitudes however a few were aware of the observed and inferred data. For example their knowledge of trackers and how trackers gather their information and use them were ambiguous to the participants. This lack of awareness sometimes comes from misconception about the type of data disclosed or could be related to the cultural/geographical differences. Misconception about the data revealed refers to the sensitivity of the data with respect to its direct effect to the user. For example, explicit mention some sensitive data that could harm the user directly like revealing bank account login information or passwords. Other stream of this disregard could be related to people with different cultures. We could see that all of our Swedish participants carry this disregard with various degrees mainly due to the openness culture within Sweden. They believed that Nordic countries could be regarded safe with respect to protection of personal information and dangers to their privacy.

> *"…I don't feel that way when I am in the Nordic countries, maybe I feel like, but more safer than if I was for example, oh what would I say in some places like in Latin America … "*

Non-Swedish participants were different in their awareness of dangers towards the privacy some due to different reason like the educations (e.g. two students in Information Security were more informed) or country of origin (where interception of data by government is recognized) and we could note that learning factors play an important role to raise one's awareness towards potential vulnerabilities on information privacy.

Benefits realization

The DataBait tool was presented in the workshops with the introduction of the aims and objectives of the project accompanied by the tool's main functionalities which are to be tested during the pre-pilot. Each functionality was presented using the mocks of the future tool and the expected results. The participants were allowed to freely discuss about the features and ask questions about the functions. Therefore we created a milieu for the lively discussion and to capture their concerns and how the tool could serve them in different scenarios. For example we could observe that most of the participants were curious about the features and at some points were shocked by the level of the tool's sophistication.

Discussions of the tool's functions allowed us to better understand the nuances of a privacy enhancing tool adaptation and usage. Several important trends emerged during the

discussions on the tool's features and future usability adoption issues. One surge of issues evolved around what are the gained benefits. The benefits were mostly related to the unconscious disclosure of personal information that might have an impact on one's public image. Again the role of (un)awareness was dominant. Although some participants were aware of the institutional privacy and general dangers towards leaking of personal information, their behaviors were different than of their intentions (privacy paradox). By looking at the different functions, they could clearly see this paradox and react on their actual behavior. Therefore matching of behaviors and intentions was one important finding that was expressed by the participants.

Users displayed some elements of scepticism towards the institutional privacy but their lack of awareness outweighed taking necessary actions that could alleviate their cautions demeanour. The users saw the benefits in this awareness awakening through manipulation of informed disclosure. The benefits were also associated to each and every participant's disclosed information at various levels. Users disclose information at various levels that are determined by each and every user's beliefs, cultures, economical values gained, political outreach and etc. DataBait's personality trait function showed to be beneficial in this sense since users can be sensitive to different subjects.

Multimedia and location leaks functionalities could draw user's attention on various levels of disclosure both those revealed intentionality and those that are unintentional. From intentional point of view users find this helpful with respect to the values of the contents to the Social Media owners. So what made them more aware of their shared content was the ability to see the profits of their contents from the social media owner's perspective; to see what could be gained from the contents and how those could be inferred. Even though they are aware of their shared contents, their perception of the contents secondary usage was limited so that social media owner's bad intentions could hide in the user's low institutional privacy awareness.

Unintentionally revealed sensitive information interpreted by DataBait could help the participants learn more about the adverse effects of their actions and seek to possible solutions e.g. deleting photo/location leak or limiting audience. The part of the benefit realization presented in relation to the concerns over historical information privacy. The presence of personal information from the period of social media profile creation can take many years which can give a comprehensive amount of data to observe and infer either by the other normal end users like friends, friends of friends, public or even the third parties interested in such profiles. Most of the participants mentioned that it can sometimes be a cumbersome task to go through all of their personal data and review them again. It can sometimes be a daunting task since the privacy concerns also varies over time, so it might need periodic review of the historic content which lack of an automated mechanism could lead to frustration and therefore leading to leaving the profile as it is with all possible breaches. The connection between those concerns over historical data and automated mechanisms (like the ones in DataBait) showed to create a positive attitude led by raising awareness directing towards consequently a progressive behavior.

> *it is mainly for work or school, because I don't like the people can find so many things, there was much information, I mean there was information from back from anything 2007, 2008 on Facebook and people scroll back and see whatever dumb things , I just brought or some stupid pictures , so I deleted everything.*

It is important to note the perceived benefits and relate them to the privacy concerns. DataBait concept evaluation showed that the tool can give benefits to its users through alleviation of various concerns. Those are benefits over managing historical data, varied levels of disclosure for each user, creating more awareness towards consequences over the perceived benefits, social media owners' perspective over what could be drawn and helping to create more informed decisions over visibility of social media contents. These factors need to be incorporated in order to educate, raise awareness and help manage online social media content accordingly.

Conveying the right message

Abovementioned emerged concepts had direct relation to the way information needs to be presented to the users and therefore user experience of the DataBait tool. There were several encounters within the workshops that triggered user's curiosity about the application which were mainly due to revelation of what could be drawn from personal information and content. We found that users are willing to be more educated through the tool about adverse effects of their sharing habits and a sense of dread could raise their effective awareness. By effective awareness we mean that the tool could trigger the potential action, which might be overlooked over perceived benefits, into a de facto action. Here the idea is that the users are more intrigued when they see dangers more explicitly. This has then led the users to perceive such privacy tools to be more effective. Simplicity showed to have an impact on how the users are willing to adopt a tool. Most of the users agreed that the tool needs to have a simple to use settings with self-explanatory features. For example after seeing concepts related to a photo, users need to know what this means and what can they do about it.

> *I would say if you can get for example what could happen if you don't use it, like I think information towards what is going on, most people pay attention to that really, that information you, the information that Facebook wants you to share, it is like a purpose it is not they want you to have fun like if they want the location, it is not just for fun like most people think it is, it is like very commercial purpose, you know they would sell the information to companies in the neighborhood in the area and everything and I mean when I know what you use and like this then I would use it but, I don't think most people would just think "Oh, it is another way just to track me." Maybe the benefits of it should be here and the benefits of not having it, maybe people could be interested to use it.*

Therefore there are different ways that the DataBait tool communicates messages to its users; guidelines that educates users about different features, the experience provoked, vulnerabilities presented and meaning of different vulnerabilities.

5.3. Test requirements

The results drawn from the concept evaluation study revealed several important perspectives for the design of the pre-pilot. The DataBait tool needs to enhance the user's understanding of privacy issues in the social media and over the internet through raising privacy related awareness. Apart from this there are other factors that need to be incorporated like usability issues, level of precision, user experience and learnability of the tool.

Following test requirements need to be fulfilled after the pre-pilot:

- DataBait needs to inform users about the locations and photos that have been revealed unconsciously (if there are any)
- DataBait needs to inspire users about the privacy consequences of their previous location and photo shares
- Users need to understand what each function does through in-tool guides and tutorials
- The tool must be easy to use
- Raised end user awareness through information revealed from location leaks, photo leaks and trackers
- Users should be able to run the tool smoothly
- The tool should help towards informed users' behavior in their online sharing habits

Based on the requirements we have set, the research questions for the pre-pilot as follows:

1. Users need to be able to see whether their previous shared contents have privacy consequences
    a. Their photos
        i. What is drawn from the Facebook photos?
        ii. Whether concepts are relevant; if they're conceptualized correctly?
        iii. If DataBait tool is able to show it effectively and meaningfully?
        iv. Are users able to understand how institutional privacy works based on the concepts?
    b. Their locations
        i. What is drawn from the location leaks from user's Facebook profile?
        ii. Whether concepts are relevant; if they're conceptualized correctly?
        iii. If DataBait tool is able to show it effectively, is it meaningful?
    c. Are users able to understand how location and photo concepts could be exploited by OSN?
2. Trackers
    a. What are trackers? If users are able to understand the purpose of the trackers?
    b. Type of language and information used to educate users
    c. User should be able to manage the trackers; is it intuitive to have a blacklist/whitelist of trackers? Why (not)?
    d. Does tracker functionality affect the user's privacy concerns?

5.4. Testing of Live System

Methodology

This section presents the method used to perform the tests together with the actual users of the DataBait tool. The requirements mentioned in the previous section acted as an input to the second part of the test. Starting from requirements, it was demanded to opt for a qualitative study because this would give us the chance to closely observe the users while using the system. Teague, De Jesus, & Ueno (2001) assert that users verbalize what they are experiencing and therefore it can give better understanding to users' reactions to a site, product, or application. Among different usability testing approaches together with users we preferred workshops and interviews. A workshop gives the participants the opportunity to the participants to comment, elaborate, criticize or contradict other participants which will spark a lively discussion which lifts up the nuances of the subject under study. On the other hand, the individual in-depth interviews help to gain deeper insights into an individual user's attitudes, desires and capabilities when working with the tool. This helps to overcome the group dynamic that might occur when in group workshops.

The workshops were held in Luleå, Sweden. Invitations were sent to the Botnia Living Lab, students of LTU and employees within LTU. After receiving interest from the participants and analyzing their profile, a balanced mixture of participants in terms of occupation, sex, and age were invited to the workshops. All the equipment to run the workshop like computers with internet access and a projector to demonstrate the system were installed in the designated rooms. Total 15 people attended the workshops in three sessions. The order of the workshops was as follows:

1. Introduction of participants in the group
2. Introduction of project
3. Presentation of DataBait tool (aims and objectives)
4. Usability testing on registration process including DLA signing
5. Answering an online questionnaire for registration
6. Discussion on photo and location sharing practices in Facebook
7. Usability testing on location and photo leaks functions in DataBait
8. Answering an online questionnaire regarding location and photo leaks functions in DataBait
9. Discussion of different features and lessons learned
10. End of workshop

As mentioned above, two methods of investigation were used for evaluation. A quantitative questionnaire and qualitative group discussion or interview. The questionnaire was used during the workshop sessions and users were asked to answer certain questions based on the features that were being tested. The questionnaire consisted of 19 questions in total with questions regarding the demographics of participants, DLA signing process (D3.1) and its importance, overall registration process, Facebook data linkage to DataBait, location leaks function, image leaks function, tracker plugin and information button. The full version of the questionnaire can be found in the Appendix.

The qualitative study was designed based on the output results of the concept evaluation study (section 5.2.3). We designed the overall flow of tasks in a way that users would be able to compare their privacy preferences with their actual behaviors to capture the ability of the DataBait tool in raising user's awareness about the consequences of their Facebook profile. For this aim, we first tried to capture how much users are concerned about their location and photo sharing practices and what sort of privacy enhancing approaches they employ to align it with their preferences. We then compared this with their reflections after using the tool to see what sort of revelations the tool has made and whether they would be able to make sense of what have been drawn from the tool. Genres of disclosure theory has been used as an analytical lens to capture the nuances of privacy concerns and actual behavior to evaluate to what extent the DataBait tool is playing role in raising awareness. Usability testing was another focus of the workshop study. Rubin & Chisnell (2008) "Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests" was used as the guide to plan and create (micro-) tasks. They note that "Exploratory tests usually dictate extensive interaction between the participant and test moderator to establish the efficacy of preliminary design concepts" (page 30). Therefore the designed questions for the workshops were targeted towards initiating this interaction between the testers and moderators to evaluate understanding, values, navigation, user experience compared to similar products and learnability of DataBait tool. A full list of questions is available in the Appendix.

Table 3 shows information about the sessions:

Workshop	**Date**	**Participants**	**Duration**	**Location**
1	02/09/2015	6 (2f, 4m)	1 H 41 M	Luleå
2	03/09/2015	3 (1f, 2m)	1 H 36 M	Luleå
3	04/09/2015	6 (3f, 3m)	1 H 42 M	Luleå

Table 3

The sessions were audio recorded and transcribed by a personnel in LTU. All participant names are anonymized and fake names will be used in this report for the sake of non-disclosure agreements. LTU gave a movie ticket voucher of worth of 150SEK as gratitude of participation to the testers. The analysis of the text was done exactly the same as described in section 5.2.3.

The Interview sessions took place in Flanders, Belgium over the first two weeks of September 2015. Since it is a part of the usability test to see if the structure and workings of the tool are clearly structured for the general public, we opted for a population of mixed age, varying between 20 and 53 years old. A total of 10 participants, which consisted of 5 male and 5 female respondents, were recruited throughout the region of Flanders. The interviews lasted on average 1 hour and 30 minutes. All sessions were tape-recorded and subsequently transcribed. To ensure anonymity, all respondents received pseudonyms in the transcriptions, which will also be used in this deliverable. For their participation in our research, they got rewarded with a voucher for a multimedia store (FNAC) with a value of €25.

The interviews touched the same topics as the workshops (see previous paragraph). If some time was left, we ended the interviews with discussing another feature of the DataBait-tool, namely the tracking function, on which we will elaborate further on this deliverable.

Table 4 summarizes the participant's demographics and timing of their interviews.

Interview	Date	Name (gender)	Age	Professional Situation	Frequency Facebook use
1	04/09/2015	Neil (m)	37	Employed	Several times a day
2	07/09/2015	Kathy (f)	45	Unspecified	Several times a day
3	07/09/2015	Nina (m)	27	Employed	Weekly
4	07/09/2015	Courtney (f)	38	Employed	Several times a day
5	08/09/2015	Paul (m)	32	Employed	Daily
6	09/09/2015	Bob (m)	20	Student	Several times a day
7	09/09/2015	Joni (f)	23	Employed	Several times a day
8	13/09/2015	Sylvie (f)	24	Employed	Weekly
8	17/09/2015	Rick (m)	53	Employed	Weekly
10	18/09/2015	Tim (m)	28	Employed	Several times a day

Table 4

We opted for 'only' 10 respondents, based on the results of (Nielsen & Landauer, 1993), who claim that after tests with 5 respondents up to 85% of the usability problems will be addressed (See Figure 12). Each additional user brings less new material to the metaphorical research table and the same problems will be brought up again and again. We added an extra 5 users since we did not only want to learn about usability issues but also about their attitudes

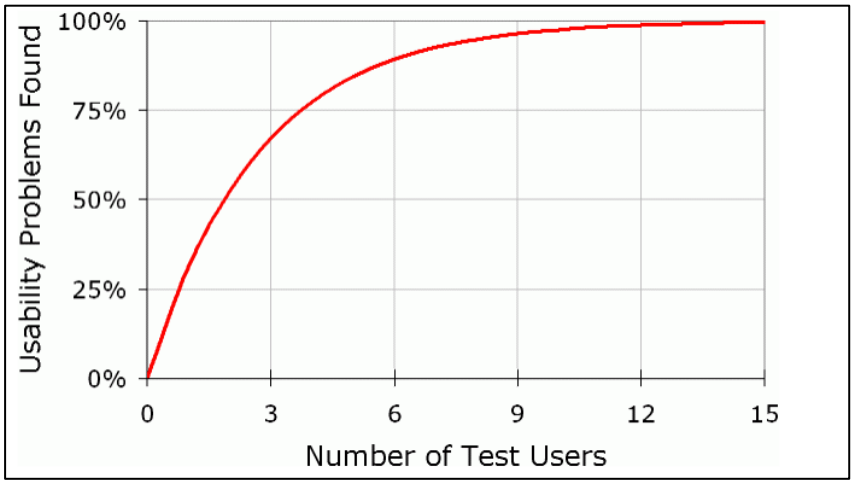


Figure 12 Nielsen et al. 5 participants rule

6. Ethical Aspects

USEMP strives to offer better control over OSN users' personal data and, consequently, personal data need to be processed during the project. The central role of experienced legal experts and of living labs that already work with panels of users will ensure that personal data are processed with strict observance of legal and ethical regulations. USEMP will give specific attention to any ethical issues that will arise and will address them in a professional way following established and upcoming EU regulations and corresponding national laws about privacy, digital and property rights issues and protection of minors very closely.

Any data collected for privacy and data protection experiments will be strictly anonymous. To enforce anonymity, established practises of living lab partners will be implemented in USEMP and the other partners will not have direct contact with participants to the experiments. In all cases the personal identity of the data will be strictly protected from third parties and will only be used for testing purposes within the project. USEMP will comply with data protection acts, directives, and opinions, both at European and at National level. These include:

- Directive 95/46/EC of the European Parliament and the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.
- The upcoming General Data Protection Regulation that will supersede the Directive 95/46/EC and National laws of EU member states and constitutes one of the main inspirations for USEMP works.
- The Charter of Fundamental Rights of the EU, specifically the article concerning the protection of personal data.
- The opinions of the European Group on Ethics in Science and New Technologies in their report "Citizens Rights and New Technologies: A European Challenge" on the Charter on Fundamental Rights related to technological innovation.

USEMP will perform user studies and tests and will operate with potentially sensitive data that are volunteered, obtained through behaviour analysis or inferred from the first two types. Following the best practice for ethics in Human-Computer Interaction (Ethics in HCI and Usability, 2010) and living labs partners existing practices, the personal data collected during the user evaluations will be automatically anonymised and used for research purposes only.

The data may include, but not limited to, personal information about the user such as: name, date of birth, interests, location, images, texts, opinions, or relations to other users, behavioural data such as clicks but also information derived from volunteered and behavioural data. It will not be transmitted to third parties, and will be handled with the experiment participants' explicit consents after clearly explaining what type of data will be collected and how they will be used, how it will be stored and destroyed after the experiments.

Consent will be obtained by formulating acceptance terms of usage, and depending on how far-reaching data collection is, informed consent will be requested at several levels of agreement (e.g. people may agree that USEMP analyses the data they upload, but not their user interactions, because this may intrude deeper into their privacy). Part of the terms of usage will be the information of users about the legal aspects of obtaining information for evidential purposes.

USEMP will observe European legal regulations concerning privacy. This is at a policy level, and will be monitored and reinforced by USEMP Coordinators, Boards and the CEA legal department. ICIS will have a particularly important contribution here since they have extensive expertise in working with personal data. At the technical level strong technical measures concerning data security of personal data will be applied. For instance transmission of personal data over open communication channels will be done in encrypted form only. Several partners of the consortium, most notably HWC, have considerable experience with such privacy protection measures. Another aspect of privacy is the protection from spamming for which appropriate tools will be devised. As a further measure to ensure compliance with legal and ethical conduct with private data, USEMP will provide a mandatory training day on data privacy for all project researchers at the project kick-off and two further ones before the start of the pilot studies.

The IMINDS panel management for Living Lab research – as foreseen in USEMP – fully complies with Belgian and European privacy regulation. This includes that a notification for their data processing operations has been submitted to the Privacy Commission (www.privacycommission.be). A clear user agreement form is foreseen with every respondent that participates, which includes the rights and obligations of the researchers and the test users regarding privacy, data protection and related issues. In this agreement the anonymity of personal data is guaranteed. In the few cases were personal data are transferred not on an anonymous basis, a ‘verwerkingsovereenkomst’ (data processing agreement) is foreseen.

Taking all necessary actions described above, to our knowledge no burdens exist, since participation is voluntary, anonymous and informed consent is requested in all cases.

Ethics of the Pre-Pilot

When participating in the pre-pilot of the USEMP project, be it in the workshop or interview sessions, our participants have a number of rights and obligations. First of all, any participation is voluntary and happens out of free will. Each and every participant can choose to stop their participation at any time, moreover they can ask the USEMP consortium to correct or delete their personal data. Since we are dealing in this project with information that is personal, and might be seen as sensitive in some cases (be it from a legal or user perspective), every participant must have reached the age of 18. The goal of the pre-pilot sessions is to gather comments and/or feedback on the product (DataBait) we are developing. We, the USEMP partners, hold the right to use, copy or make derived forms of the feedback we gather.

Before the user registers for the DataBait tool, he/she has to sign a contract called the Data Licensing Agreement, where his rights and obligations are clearly specified, as well as which data is collected, for what purpose this data is collected, which project partners are involved and reason on their data, on which dimensions the user will be profiled and which measures will be taken to secure the personal data of the respondents. This is all in line with the upcoming general data protection regulation.

Inside the DataBait tool, a specific part of the website is dedicated towards giving the user more insights in how profiling takes place and which data is currently being used.

Each user also has the right to gain access to the personal data that is collected, and ask for full deletion. Since it is also a central feature of the DataBait tool to show how users are being profiled, they of course also get insight in how we try to **analyse or predict** in

particular that natural person's performance at work, economic situation, location, health, personal preferences, reliability or behaviour.

6.1. Privacy Agreements

In the context of USEMP the processing of personal data is a crucial and critical issue. It is the core purpose of USEMP to develop tools that empower users of OSNs to make more informed decision on which of their data to share.

Paradoxically this entails that USEMP

1. engages with a subset of USEMP users to develop and train the algorithms that aim to show users what can be inferred from their data and
2. engages with the same subset to inquire how they experience the use of the DataBait tools.

This necessitates the processing of a very sensitive subset of personal data, for which a special legal regime is in force. The USEMP consortium partners are very well aware of the duty of care they need to exercise and the liability they face if they fail to act as a trustworthy steward of this data. To make sure that the entire life cycle of collecting, processing, pseudonymisation, anonymisation and deletion of this data is done with the utmost prudence and care, we have developed a Data Licensing Agreement (which also contains explicit consent for the processing of sensitive data) between DataBait users and the Consortium partners, and a Personal Data Processing Agreement that clarifies the responsibilities of each USEMP partner in terms of the processing of personal data.

On top of this, two buttons will be placed on the USEMP platform, one to provide users with the relevant information on which of their data are processed how and by which partner, the other one to provide users with the means to object to further processing of their data. Finally, the collection, storage and transfer of the data will be governed by appropriate security measures that will be tested against a risk assessment template, making sure that the data cannot be accessed by unauthorized persons. A more elaborate explanation is to be found in Deliverable 3.1.

7.Results

In this section we report on the result of pre-pilot tests with some available functions of the DataBait tool. The result accumulates the outcome of the questionnaire, workshops and interview sessions, which all address the different parts of the tool.

7.1. Registration

Registration to the DataBait tool was the part of the test which required the users to enter to the DataBait[5] tool website, click on registration and follow the process until they are successfully delivered to the main screen of the website. We gave the testers guidelines with screens or web pages that the participants need on where to start and what they should see once the process is complete (see Appendix 9.2.2.). In order to gather as much as possible how smooth the logical workflow works, we rejected to give full description of how to proceed to the next level, unless were asked by the participants to avoid any bias. Once the registration was successful we asked the users to answer some questions in an online questionnaire (see appendix 9.2.3.). We were also noting the users' reactions during the process in terms of frustration, unclear messages, errors, problems and broken navigation system. From the registration process we also captured how users reacted to the DLA (whether they read the whole contract or not) and linking Facebook account to the DataBait tool.

Registration Process

In terms of registration flow, users found the process easy to follow and most of the participants completed the task without any problem. A remark that some participants of the interviews mentioned concerned the language. Not every part of the tool is already fully translated into Dutch and so it changes between languages, also some grammar mistakes were found (e.g. Are you older than **13th**). This needs to be updated in order to establish a better flow.

Another observation of some participants is that the first page of the DataBait website, does not explain what the purpose of this website is. This explanation should be given, or at least accessible, **before** logging in or registering to the tool. An option might be to change the place of the explanatory video (which is now only viewable after you log into the tool).

One respondent of the workshop and all respondents of the interview sessions had already created an account on the DataBait website for a related research. A problem occurred when they could not remember their passwords, since it is not possible to create a new account and link it to the same Facebook account. **A password retrieval function** should be implemented in the next version of the system.

Data Licensing Agreement

In terms of reading the DLA, half of the (Swedish) users read (fully or some parts) and the other half skipped it. Table 5 summarizes the clarity of the aims of objectives of the DLA to inform the (Swedish) users about all data processing taking place within the USEMP project.

[5] https://databait.hwcomms.com/

Which information did you find interesting?			
Answer Options	**Yes**	**No**	**Response Count**
Information about which partners form the consortium	4	5	9
Information about the obligations to which you commit in terms of downloading DataBait tools?	3	5	8
Information about the obligation to which the consortium partners commit in terms of providing profile transparency?	6	2	8
Information about the processing of your personal data?	7	2	9
Information about the purpose of processing your information (scientific research)?	7	1	8

Table 5

There have been some issues regarding the clarity of the message towards the information about the consortium, obligations of users and lastly the duration of availability of information handed over by the user to be used in the USEMP project. Other elements of the DLA (e.g. consortium obligations, aims and objectives) seem to have been communicated without issues. While analysing, why some testers have not read the DLA and why they ignored this part, the answers were mainly laid on the workshop setting where users felt that they were rushed to complete the process. Although it was not the intention of the workshop moderators to imply this and users were given sufficient amount of time to go through all the details of the DLA, testing how much the workshops had an impact on the way users skipped the part will be put into focus in pilot studies where users go through the process at their neutral settings.

In the interview sessions, similar results as the ones above were found. We let the users go through the DLA and gave them all the time they needed. Although the presence of a researcher might have an effect on the feeling of being 'pressured', most of the respondents skipped through the different parts of the DLA very fast, mostly stating that they never read these kinds of contracts. Only when asked later if they noticed anything different between the DLA and prevailing terms of service agreements, they stated that this one was clearer because it was cut in smaller pieces, which made it possible to glance over every part. They however also noted that to make things more clear, every little part should be preceded by one sentence or some key words that give them a hint of what is in the DLA.

Another thing that 2 respondents noted was that the progress bar should have a different colour than the background of the text box above it. Now they initially thought they were unable to scroll down to see the rest of the text.

Connect with Facebook

The DataBait tool requires the user to connect their Facebook profile in order to complete the registration process. We asked the users how comfortable they were when agreeing on linking their Facebook data to be processed by the DataBait tool. The answers were oriented towards two directions of scepticism and carefreeness. This was again supported by our test requirements gathered from the concept evaluation phase. Some users find this function as a very beneficial step (e.g. registration through Facebook) to save time, energy, avoid errors, etc, while other users were not totally comfortable to allow access to their Facebook data. The result shows that this can be viewed both as a challenge and as an opportunity to future

versions of the tool. While users can avoid some processes of registration (e.g. specifying username/password) through Facebook login, there is a danger that users drop off of using the tool at this stage. The users who were sceptic stated that they were continuing the process because they trusted the application as part of a research study. For this reason the DataBait tool need to attract user trust by different mechanisms, therefore it would be beneficial to measure the number of drop out during this phase in the pilot study when targeting wider audience.

"As always I felt a bit scared about adding a third party application to my profile." Tom

Another participant responded:

"It did feel a little "sketchy", as everything that wants to connect to my Facebook does." Micky

7.2. DataBait Information Page

Once logged in, the users are confronted with the Central Page of the DataBait tool. Most users of the interview sessions liked the aesthetic qualities of this webpage and felt it had a very clear overview. Subsequently, we asked them to look for the webpage with more information about the tool and the organization behind it[6]. Without much effort, most of the participants, correctly identified the 'DataBait: what, why, how' page. An issue that was mentioned is that neither the movie, nor the info graphics fit the screen of the page. Especially for the info graphics, this was considered frustrating since the user always had to scroll down and up to go through it.

The content of the movie is considered a success from the testings. All participants find it to be a useful addition that helps them understand the purpose of the tool. They liked the language and imagery: both playful and informative. As mentioned before, some participants felt however that it would be better to put the video on the homepage before logging in to the system.

Hereafter, participants clicked on the 'Practical Info' tab, the naming should be changed since it does not fully encompass the scope of the information on it. As it only gives contact information, a more logical name would be 'Contact'. They all agreed that this type of information should be included in this webpage and that an e-mail address is sufficient. Two users however also mentioned that if the tool was to be provided by a commercial organization, they also expect a twitter account. Since they feel that when they have a complaint it can be propagated more rapidly through Twitter.

The 'Which of your personal data do we process?' page was well received. Its intentions are clear and the structure is suitable.

[6] See D3.1 and D3.6 for more information

7.3. Location Leaks

In order to capture user's attitudes towards sharing their location, we asked them about the situations and occasions when they reveal their location in their Facebook profile. However we kept the question on a high level to avoid any bias. We were especially interested in user's perception of location sharing and to understand how they interpret the location sharing. By this we mean the ways in which they could think of their location revealed. We asked them questions like give examples of when you share location, where, with whom etc. without specifically mentioning about the ways. We kept the concept abstract by stimulating the user's thought about what locations could be revealed through Facebook. The results show that users' understanding of revealing location is bound to the functions offered by Facebook but not what could be inferred from the profile itself. Even the awareness among the users about the location sharing functions of Facebook varies between the users. Most of the users thought of revealing location through check-in function while a few could state that also location could be acknowledged through hometown, school, current town, tagged by others and etc. check-in function has been used at least once for every participant but the level of sharing between them varied a lot. Most of the users use this function rarely while a few use it regularly. The reason behind sharing is mainly for informing with the purposes varying between enjoyment, to meet up, remembrance and show off. Main reasons for not sharing location is mainly due to personal and security reasons like being afraid of theft while abroad. We asked and captured how users feel about revealing certain location information and how far they would go before reaching their red lines. Consequently this has allowed us to compare their expectation about certain practice (either location or photo sharing) into their reaction against the information provided to them through analysis of their Facebook profile by the DataBait tool.

In the next step we asked them to go back to the DataBait tool and check the 'Location Leaks' page (as depicted in Figure 13).

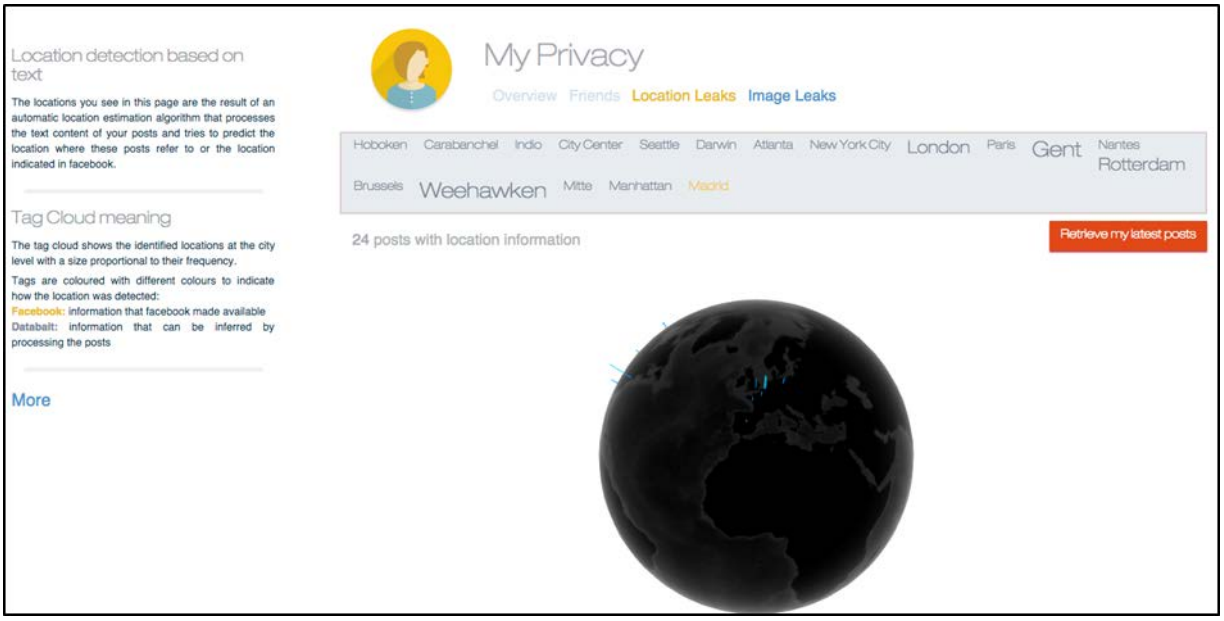


Figure 13 Location Leaks

After receiving their feedback it was shocking to most of the audience how much they are revealing about their location which, went beyond their initial expectations. The interesting part for them was to see all of their location exposures gathered in a single panel. This gave the idea of about the trajectory of their locations at one view. However the users found this function not very accurate. Among the concepts shown, users believed that most of them they did not make sense and were unrelated with the actual source. (This criticism was pointed only to the inferred locations). While a few paid attention to the confidence level of each concept which was low in all cases, most of the participants expressed that they would like to see the places where the source of information inferred has more certainty. Figure 14 shows that majority of the participants found the concepts were little or with non-accuracy.

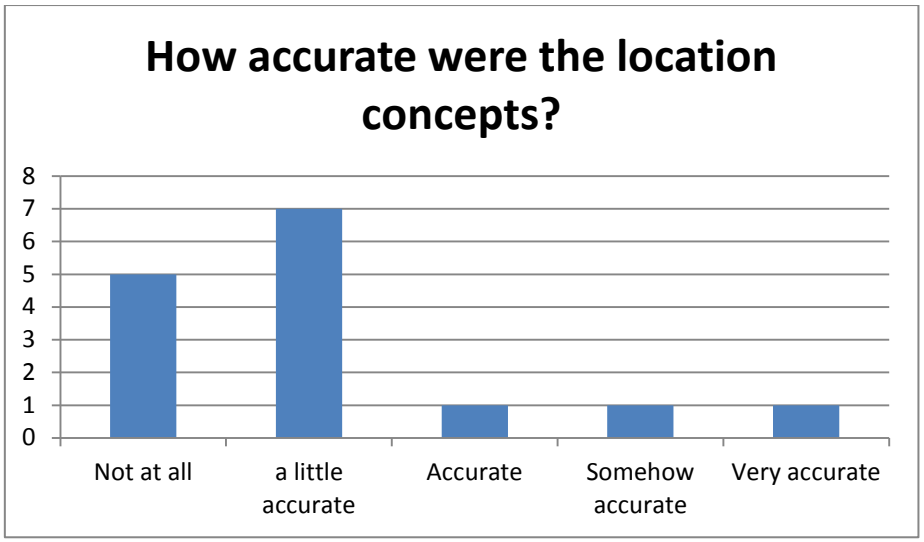


Figure 14

Although some concepts inferred did not make sense of the source, users could deduce the idea behind the function. When asking them about what this function could afford to, they replied with some interesting remarks all focusing on how information could be inferred to predict the whereabouts of one used for advertising or other malicious purposes.

> *"I am not that worried if people know where I am, and also looking at this (DataBait), it has a 1/50 chance of getting it right from what I see right now. If it would be better and exactly pinpoint where I have been and make correct assumptions then I would think it would be pretty good and probably useful to see OK this is what Facebook and all of Facebook's partners can see that I have done or where I have been…" Doris*

> *"I think it's good because it gives people more clean picture of what companies and other people get out of what you do on Facebook, like she visits this place a lot so they can keep the records of that." Maria*

We also faced some minor issues which affected the usability of DataBait tool. There were issues with some technical aspects like two occasions when the page was non-responsive and the user had to logout and login again. Some concepts had no source associated with them. In one case the user did not see any result. However, he strongly believed that he has used location service or at least some status updates that could have contained location. There were also occasions where the text source were inferred wrongly because of similarity between the work to the name of a city for example 'nice' which means 'cordial,

kind, good' were mixed up with the name of the city Nice located on the south east coast of France. Therefore the following versions of the DataBait application need to resolve the abovementioned issues.

7.4. Image Leaks

Similarly as with the testing of the location leaks, we started the discussion with capturing the users' attitudes towards sharing images through Facebook. Again we asked them if they shared photos on a regular basis and if so, on which occasions. Here it became clear that most of the respondents only shared pictures when attending special events, like concerts, when eating in restaurants or (and most often mentioned) when on holiday. When people did share images, their main reason was that they wanted to show their close friends that they were doing well (this was especially true when the picture was taken on holiday). One respondent claimed (and this was later underwritten by the image leak function) that she never posts pictures anymore on Facebook and deleted all the old ones because she didn't want to take part in to the online good news show that she calls Facebook use. Another reason for not posting images was, similarly as with the location sharing, due to security reasons. Some participants claimed they would only post pictures of an event after they are home again. Furthermore, we questioned them about which device they primarily used to upload pictures. Most users only use the smartphone to do this. Only two of them said they still have another camera that they take with them to special events and link to their computer afterwards to upload the pictures on Facebook. When asked if they thought they sometimes share information they wish to keep private through images, the majority of the participants answered negatively. They feel they have enough control over image sharing and don't think that they themselves reveal sensitive information. Some had however deleted pictures from the past, because they don't find them relevant anymore and they also use Facebook's review function to see which pictures others tag them in before allowing it to appear on their profile.

After this we explained to the users that we were developing a feature that attaches concepts to their pictures and asked them to locate the function. Most users wanted to click on 'Trackers' instead of my privacy. This is due to the fact that the word 'trackers' does probably not ring a bell to a layperson, but also that 'my privacy' does not fully encompass the functions behind it. When arriving at the image leaks page (as depicted in Figure 15), we asked them to read the explanation on the left hand side of the screen. This text was clearly understood by our participants and they were eager to find out which concepts were linked to their pictures. The visualizations with the word cloud and coloured circles also appealed to them. Only for one user these visualizations did not appear and the text 'your images are being processed' stayed. They thought it was an interesting feature that they could see all their pictures gathered in one screen as an alternative visualization to what Facebook provides. Some participants did not find it intuitive to click on the concepts and thought that the concept's panel was only for information.

When going over the different topics, most users were slightly impressed about the amount of information could be attached to visual images. They also found the concepts that were attached to the images to be accurate to very accurate. Though they stated it did not make them more privacy conscious since they were aware that they had released these pictures in the past. Also the concepts are not very privacy intrusive. A recommendation would be to

limit the amount of concepts in this tool to only those pertaining sensitive information, or to aggregate the concepts on a higher level in fewer overarching dimensions (such as e.g. health). They did however recognize how this kind of algorithms could be very useful to advertising agencies to target them in a better way. Another recommendation that was made by a user is that the tool also shows to how many people the pictures are visible (friends, friends of friends, public or custom).

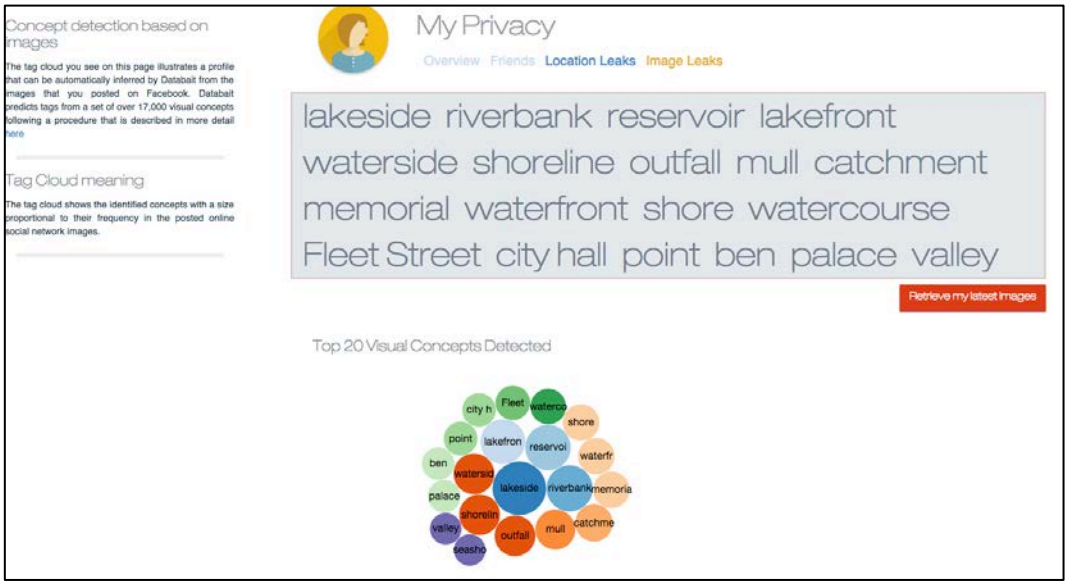


Figure 15 Image leaks

7.5. Tracking

If there was some time left in the interview sessions, we asked the participants to have a first look at the tracking function of the DataBait tool. This feature visualizes the different data brokers and companies that track the users on websites and collect their data. We first asked the users who they thought that were following their moves on the Internet. Most often the respondents answered with big companies such as Google and Facebook and other websites where they filled in a registration form. After this we navigated them to a Flemish commercial news site (hln.be), which relies heavily on income from advertising. When showing the different companies that were tracking them on this website, the general response was one of surprise. However, they immediately understood that the information they provided on websites was useful for marketing purposes and they were not immediately upset by this situation. They also said that the information was probably aggregated and not used to identify an individual.

We then returned to the DataBait website where the visualization of the different trackers for the different websites was shown. The users thought it was a nice overview and liked the feature that they could turn of the trackers if they wanted to. They also wondered if this kind of tool would also influence their browsing behavior. This seems to be a big predictor of the eventual use of this kind of tools. Two users said they used Ghostery[7] in the past (a similar

[7] https://www.ghostery.com/

plugin) but that they stopped using it because some video's on websites did not load properly anymore and they simply did not look at it anymore after awhile. A third person who also used a similar web plugin, noticed how she could not see videos on a certain website anymore but that she just stopped going to that website (vtm.be).

One time, it was not possible to show the feature because once installed, instead of receiving the DataBait icon in the browser, the icon of Disconnect[8] appeared, which is a similar app. But it did not work, as shown in Figure 16.

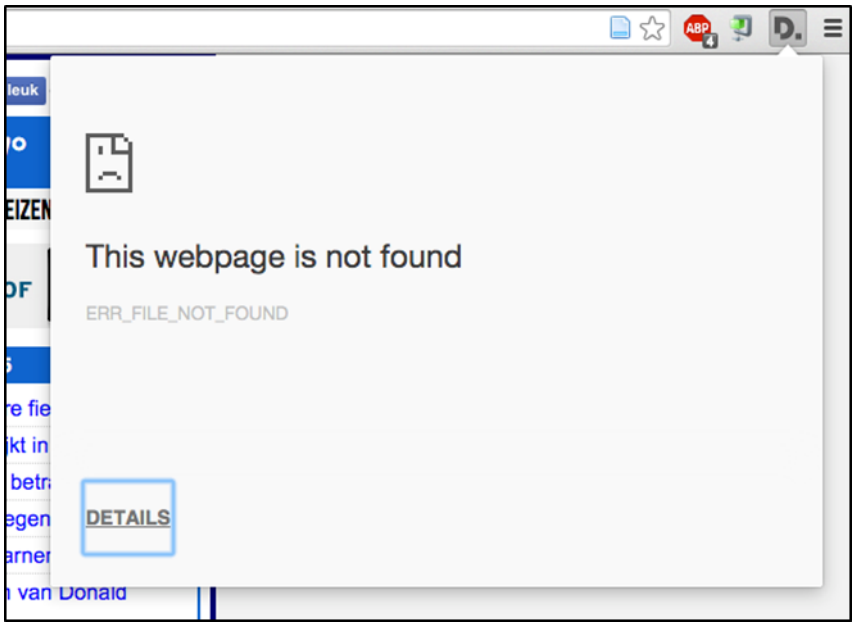


Figure 16 Disconnect Problem

7.6. System Usability Scale

At the end of each interview session, we let the participants (n=10) fill in the System Usability Scale in relation to their experience with the DataBait tool. Eventually we got a mean SUS score of 78,25 (s=5,90). In order to understand what this score actually represents; Bangor et al. (2009, p.19) developed an interpretation as shown in Figure 17. Here you see that according to SUS, DataBait is perceived as a system with good usability.

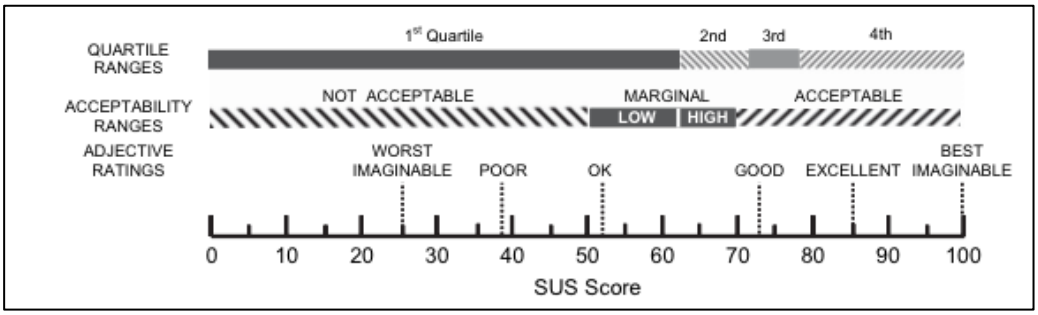


Figure 17 A comparison of mean System Usability Scale (SUS) scores by quartile, adjective ratings, and the acceptability of the overall SUS score. (Bangor et al., 2009, p.19)

[8] https://disconnect.me/

We have to note though, that some other reasons may have influenced this fairly high score. First of all, the sample size (only ten participants) was on the low side of the acceptable spectrum. Secondly, while performing the task, the researcher was sitting close to the participant. This could result in some unintended pressure to give socially accepted responses. We also want to note that, in line with Bangor et al. (2009), a single metric should not be used in isolation to make absolute judgements. The nature of the failures that the participants have encountered should also be taken into account when estimating a system's usability, hence the importance of the other results as described earlier in this report. Finally, it is clear from past research that systems with a SUS score below 50 will have usability difficulties in the field, it is not certain that high SUS scores predict adoption (Bangor et al., 2009).

7.7. Summary

All in all, the DataBait tool fulfilled the conceptual ideas behind the USEMP project and some practical steps to make user more aware of their online privacy. Based on test requirements gathered from concept evaluation, we performed an evaluation of the actual system (with some functions). Users could easily relate the basic ideas behind the tool to what sort of affordances could be derived. For example one participant saw some "Muslim" related concepts in her profile. Since she was not a Muslim herself, she wondered, in the beginning, the reason behind this information. After investigating through the concepts and seeing the related pictures, she found that in those pictures she was accompanied by her friend with hijab. She raised a strong concern over the issue that OSN or other aggregators might make wrong inferences based on pictures that a person has shared. She told us she will immediately delete those pictures because it might affect her future career and that she was concerned that her profile can expose religion-related indications. This example showed us that we were able to fulfil the requirements mentioned in 5.3. The user was able to see some interesting concepts that was against her preferences to be presented online (i.e. religion) by understanding more about her historical data. She could immediately see the affordances by matching her expectations (religious related contents) of privacy and the risks towards it (being profiled as a something she doesn't likes to project). Hence general level objectives of the project seem to be going to the right direction.

There are minor issues, namely the usability aspect, which need to be resolved for the next round. Here is the list of issues that are recommended to the developers:

No.	Related area	Issue
1	Image/location leaks	Low frequent tags found to be unnecessary
2	Location leaks	World map progression is not intuitive
3	Image/location leaks	It is not intuitive to click on a concept
4	Location leaks	Some concepts have no post associated with them
5	Registration	Password reset/forget function is required
6	DLA	Colouring in the progress bar
7	Info page	Movie and info graphics does not fit in the page

8	Info page	'Practical info' does not fully encompass the scope of the information on it
9	Location leaks	Users did not pay attention to the confidence level
10	Location leaks	Resemblance of words to concepts like 'nice' to City 'Nice'
11	Navigation	Trying to find 'Image leaks' within trackers section
12	Image leaks	The concepts are not very privacy intrusive
13	Tracker plugin	Only available for Chrome browser

Table 6 List of minor issues

8. Feedback

8.1. Technology acceptance model

Based on table 6 of list of minor issues, we have rendered a list of suggestions focusing on the usability aspects of the DataBait tool.

No.	Area	Suggestions
1	Registration	Facebook login as single sign on to eliminate the problem with forgotten passwords
2	Registration	Password retrieval function as an alternative to 1
3	DLA	Every little part should be preceded by one sentence or some key words that give them a hint of what is in the DLA
4	DLA panel	Changing the color of progress bar text
5	Information video	Moved to the home page (before login) to be more explicit
6	Information page	Change 'Practical Info' tab to 'Contacts'
7	Location leaks	Highlight the confidence level
8	Image leaks	Limit the amount of concepts in this tool to only those pertaining sensitive information, or to aggregate the concepts on a higher level in fewer overarching dimensions
9	Image leaks	Show to how many people the pictures are visible

Table 7 List of suggestions

8.2. Discussing test requirements

In section 5.2.3 we enumerated the overarching test requirements that need to be fulfilled after the test. We highlighted three main categories of Awareness, Benefit realization and conveying the right message. Below we will look at the results through the requirements to strengthen the advantages and to overcome the deficiencies in the pilot studies.

We argued that awareness needs to be an essential part of the tool to give the users knowledge about the dangers towards the personal information. Specifically we argued that awareness should incorporate in a way that users are educated about their online profile and get to know about the possible secondary usages of their information.

Translating through genres of disclosure theory (section 3.1) (Padyab, 2014; Palen & Dourish, 2003), defined as "socially constructed patterns of privacy management," it draws attention to the communicative practices involved in a system to insinuate about the expectations of use according to the users. Therefore it suggests that privacy could be violated if the expectation of use is different than what the system is/could present(ing). Testing of this theory should reveal that by comparison of what users were thinking of how their information handled in a system (e.g. Facebook) to how information could be handled in the system will indicate to what extent the users feel that their privacy is at danger. If we argue that the users came to conclusion that their privacy is at risk, then we can start analyzing the factors that will lead this transformation of awareness. In order to achieve this goal, we compared the user's pre-DataBait-use to their thought on the post-DataBait-use. The dominant benefit realization of the tool lies within its ability to inform users about the

institutional privacy and the logic behind the OSN hungrily demanding more information sharing. Benefit realization was achieved through giving a tool to the users which allows them to uncover some of the conventions happening behind the scenes of OSN and other connected parties. Here we could argue that awareness is a result of benefit realized through conveying the message that illuminates their online social behaviors and traps that users are usually dragged into. Users ventured into an experience of self-revelation about some of their communicative practices that led them to double think about their sharing behaviors. Most of the users stated that they will change the current setting of their current Facebook profiles accordingly to restrict the access to their contents. Even some expressed deep concerns about some of their contents and immediately deleted those.

Translating the results based on the values table presented in the section 3.4 shows achievements of major objectives of the DataBait tool and gives directions for the next steps.

Values	**Defining goal**	**Interpretations of values**	**Translated through DataBait**
Self-Direction	Independent thought and action. Need of control and mastery	When considering this value in design, it is important to let the user take control over their own data and over what they share through the system	Users are more aware of their digital footprint however more control is demanded
Stimulation	Excitement, novelty and challenge in life	This value highlights the fact that people needs to be stimulated in the use and they need to be challenged and have the ability to learn new things	Visualization of data privacy related concepts from photos and pics stimulated users to be more cautious in the future
Hedonism	Pleasure or sensuous gratification for oneself	Related to this value it becomes clear that the system need to support having fun while using the system.	The systems proved to be easy to work with in the general level however gamification is still lacking
Achievement	Personal success through demonstrating competence according to social standards	This value put emphasis on the importance of seeing individual success and to feel competent.	Users are able to see the leaks within their Facebook profile and trackers over internet, and are more sieged over their information
Power	Social status and prestige, control or dominance over people and resources	In our interpretation of this value, control and social status came in focus. Here the feeling of being important is in focus.	Control is still lacking in this version of the application
Security	Safety, harmony, and stability of society, or relationship and of self	This value set focus on avoiding anxiety and to make citizens feel safe while using the system.	Opposite to this value the app gives users a sense of danger which ultimately leads to more harmony of expected privacy
Conformity	Restraint of actions, inclinations, and impulses likely to upset or harm others and violate social expectations and norms	Related to this value, the importance of encouraging a particular behaviour and restraining another is highlighted to keep a good and sound	The tool allows comparison of what is expected as an socially constructed pattern of privacy management and information disclosure

		community.	into actual/potential secondary use of information
Tradition	Respect, commitment, and acceptance of customs and ideas that one's culture or religion provides	This value becomes a bit difficult to translate in this context since the focus of the system is to stimulate innovation and change. However, in relation to systems design, we interpret this as making the use of the system as familiar as possible.	Does not apply to DataBait application
Benevolence	Preserving and enhancing welfare of those with whom one is in frequent personal contact	Our interpretation of this value encourages us to consider a way to socialise in the system both with new and old friends.	Current version does not support this in a fully extent however users are able to see the impact of the photos/posts which includes friends in them on their inferred profile
Universalism	Understanding, appreciation, tolerance, and protection of welfare of all people and for nature	The system focus on contributing to make users more aware of their personal information value and enhancing Internet privacy, hence this value inspire us to think of how to simulate a focus on the common good and how to improve society through peoples compassion.	The tool allows the users to see behind the scenes and look at their own data through the eyes of data controllers

Table 8 Values translated by DataBait

8.3. What is next? Getting ready for pilots

In Formative evaluation the aim is for an artifact which is still under development to be evaluated to determine areas for improvement and refinement. Formative evaluations focus on consequences and support the kinds of decisions that intend to improve the evaluand (Wiliam & Black, 1996). Looking at the future of USEMP tool (DataBait), there are several considerations that need to be addressed. From a usability point of view, the DataBait tool should enhance some minor issues regarding the navigation and on screen guides. Users appreciate being aware about their historical digital footprint, but their on-the-fly actions (e.g. the time they upload a photo) should be stressed as well. In the USEMP project, we need to consider the functionality of online scanning of material and align it better with the historical footprint. It that sense both historical and on-the-moment privacy is intertwined.

However raising of awareness alone is not enough nor is not the only aim of the USEMP project. We need to give users more control once they are informed about the inferences. Since in the current version of the tool these controls are not the focus, future plans need to take this into account. Control could be made through the DataBait's interface or by linking the contents to its source and use the hosting's (e.g. OSN) control features. In D8.3 we need to tackle the main challenge of engaging people to be more concerned about their privacy

and online footprint. Therefore it is more demanding to absorb more testers and measure their usage of the Tool during the time.

From the user experience point of view, in pre-pilot studies users showed to be more deviated toward being feared while checking their online content and the more they see the dangers the more they would be interested in these topics. This aspect also needs to be studied in the future pilots and see how much visualization changes needed, how that could help to convey a sense of fright and what is impact on the changes of behavior.

During the pre-pilot study we have encountered many requests from the users to integrate Instagram into the DataBait tool since they found the tool's functions very helpful. In the USEMP project we will take this into account and conduct feasibility studies on other OSNs as well.

There have been some issues with the reliability of the service provided during the pre-pilot studies. Although the issue was resolved right after the problem from the technical point of view, it had an impact on how well the tool could be adopted. In the pilot studies we need to make sure that the back-end, front end and hardware equipment work with the minimum disruption. The pre-pilot studies also suggested that, in the future, a support team should be built to ensure that any issues are resolved immediately while the pilots are running.

9. Appendix

9.1. Concept evaluation focus group process and questions

1. **Agenda for session**
   - Introduction of people in the group
   - Introduction of project
   - Presentation of DataBait tool
   - Concept validation and discuss DataBait
2. **People in this focus group**
   - Who are you?
   - How and why are you part of this workshop?
   - What is your background?
   - What do you work with?
3. **Introduction of the project**
4. **Social media use**
   - Describe your daily usage of social media
   - Why do you feel it is important to use social media?
   - What functions you use most often and which ones seldom
   - What sort of information/content do you share?
   - What sort of information do you mostly put into social network sites?
   - What are your concerns when doing this?
   - What sort of expectations do you have from social media providers who hold your data?
   - Whether social media providers fulfill your expectations?
   - What are your thoughts about privacy issues in your everyday life today?
   - How interested are you in privacy issues? Has your interest changed over time? If so, how and why has it changed?
5. **Introduction of DataBait tool**
   - Connecting to Facebook
   - What are your responses usually when apps ask you to connect your Facebook profile?

- What are your thoughts on giving your Facebook personal data to DataBait?
- What could motivate/discourage you to do so?

6. **Personality traits: location**
   - How often do you share your location? Why?
   - What sort of other information do you reveal along with that?
   - How important is this function?
   - What benefits do you see when having this function?
   - Could you think of a situation(s) in which this function becomes necessary?
7. **Multimedia Function**
   - How often do you share your photos? Why (not)?
   - What sort of precautions do you consider while uploading?
   - What sort of other information do you reveal along with that?
   - How important is this function?
   - What benefits do you see when having this function?
   - Could you think of a situation(s) in which this function becomes necessary?
   - What could improve your motivation of using this?
8. **Trackers function**
   - What do you know about the trackers?
   - Are you interested to see who is tracking you? Why? How could this be 'more' interesting for you to follow who is tracking you? What do you want to know more about the trackers?
   - What does 'White List" and "Black List" mean to you? Does it make sense to have two lists? Why (not)?
   - Do you generally use browser plugins and what about them do you like/dislike or how do you like them to function within the browser?
9. **General Discussion**
   - What are the benefits of such a tool with the mentioned functionalities?
   - Will you use this application? Why? Why not?
   - In what situations would it make most sense to use?
   - Lets consider about your expectations when it comes to protecting your personal information. Whether DataBait or any other ideal tool could succeed to do so? how?
10. **Dream freely**

- If you could dream freely in the area of security, how would the solution work. Never mind technical or economical hindrances which might exist today. Just tell me your ideas. Some science fiction ideas. ;-)

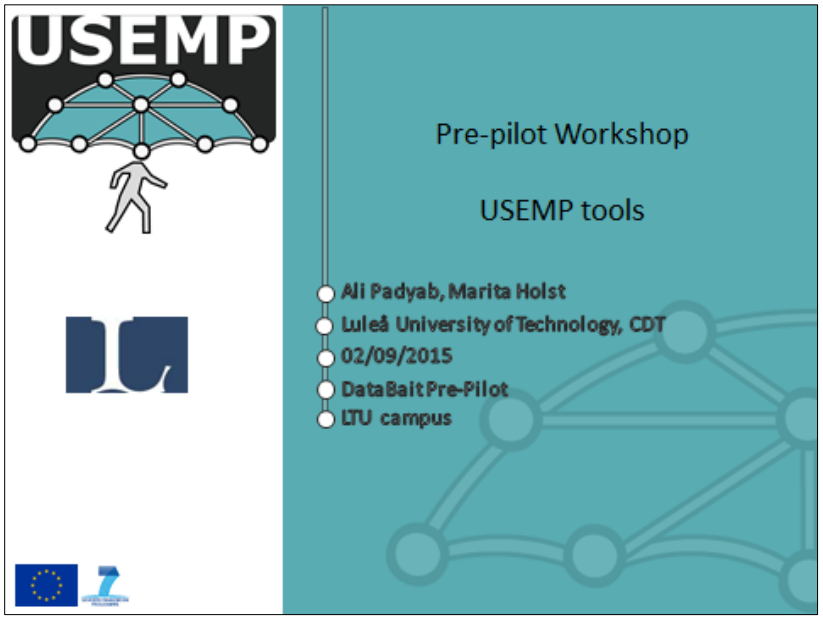
11. Missing points

- OK, we have asked our questions now. Is there an important aspect we have missed to ask about?
- What is that?
- Why is this important?

12. End of focus group

9.2. Live system evaluation workshop process and questions (LTU)

9.2.1. Workshop



Agenda for session

- Introduction of people in the group
- Introduction of project
- Presentation of DataBait tool
- Answering an online questionnaire
- Discussion on photo and location sharing
- Introduction and testing of functions
- Discussion of different features

Set-up of workshop

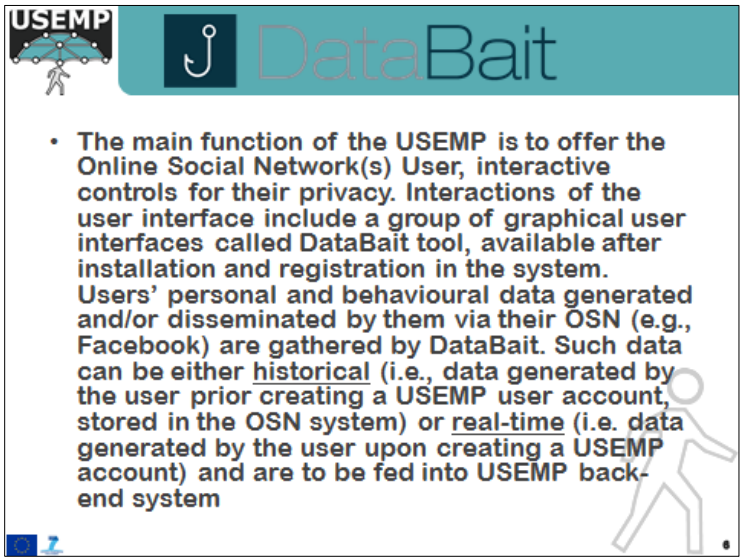
- Workshop, what is important to think about
- ~ 90 minutes
- Discuss with each others and not with the moderator. Moreover, it is not the purpose to find true or false answers; it is different perspectives of the phenomenon we are interested in.
- Even if it is a discussion it is important to try not to interrupt each other or that two or more talk at the same time, i.e. try to talk one at the time.
- Speak English so that all participants understand you and makes it easier to transcribe later

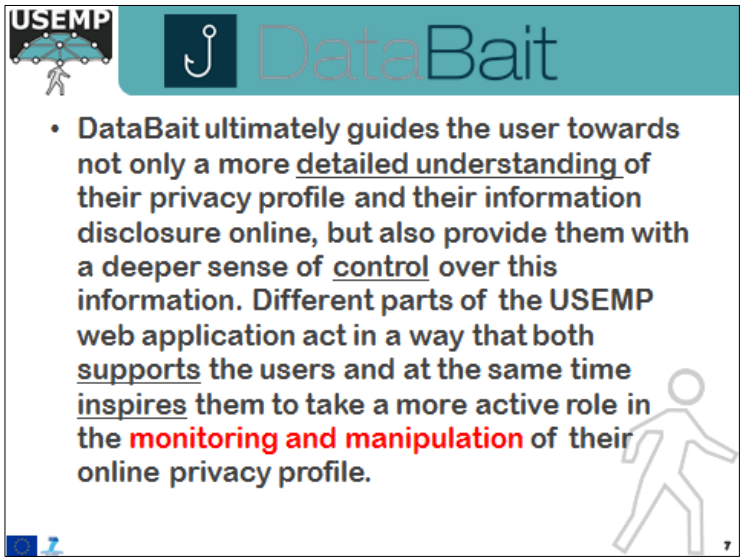
People in this focus group

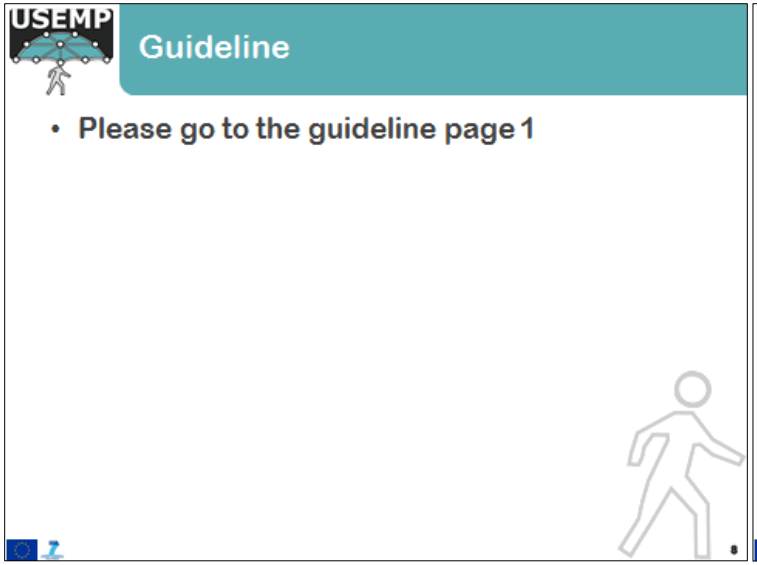
- Who are you?
- How and why are you part of this workshop?
- What is your background?
- What is your occupation?

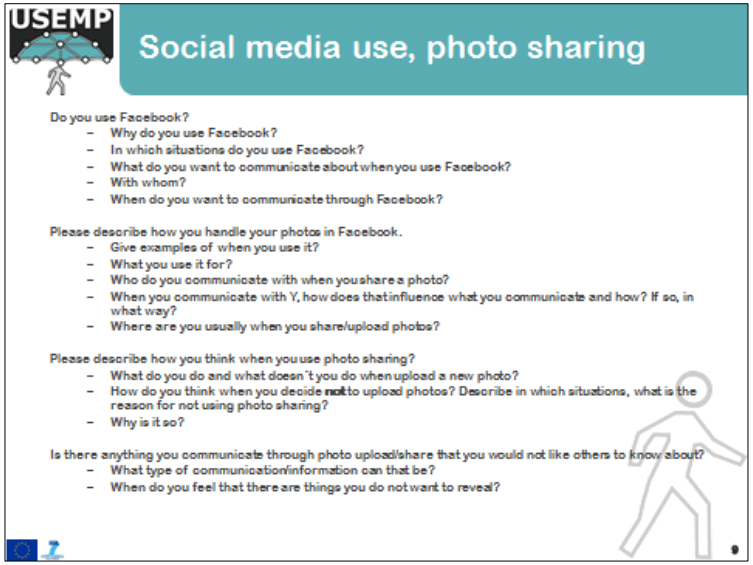
USEMP

- USEMP aims at developing a framework that will empower users by enhancing their control over the data they distribute or interact with
- Objectives:
  - Advance the understanding of privacy issues
  - Empower users to assist them in personal data management tasks
  - Raising users' awareness concerning the advantages and risks related to sharing personal data.
- USEMP will demonstrate the added value of its outcomes through (a) OSN presence control and (b) economic value of personal information
- LTU contributes to:
  - User Studies of Personal Information Sharing
  - Pre-Pilot
  - Pilot Studies









Social media use: locations

- Please describe how you handle your locations in Facebook.
  - Give examples of when you use it?
  - What you use it for?
  - Who do you communicate with when you share a location?
  - When you communicate with Y, how does that influence what you communicate and how? If so, in what way?
  - Where are you usually when you share location?
- Please describe how you think when you use location sharing?
  - What do you do and what doesn´t you do when share a new location?
  - How do you think when you decide **not** to update location? Describe in which situations, what is the reason for not using location sharing?
  - Why is it so?
- Is there anything you communicate through photo upload/share that you would not like others to know about?
  - What type of communication/information can that be?
  - When do you feel that there are things you do not want to reveal?

10

What will we do now

- Short intro to the two functions
- You will test the two features with your own Facebook account
- Questions to each function
- Closing discussions

11

Photo and location leaks

- The two functions are designed to extract the potential secondary usage of related photos and location you have shared before.
- Photo leaks function analyzes your previously uploaded photos and shows you the concepts that can be inferred by these images.
- The location leaks analyzes your posts and shows you the location related concepts

12

- Now please go to the guideline page 5

13

Location leaks Questions

- what disclosures have you made? Were you aware of them?
- Could you understand what this function does through the in-app instructions?
- How does the information represented relates or differs compared to your own expectations?
- Please describe what sort of information was useful for you when reading about your location traits. What do you expect this feature to do?
- Is this going to change the way you share your location in the future? Have you already done action to limit this?
- Please list the things you find most positive for this feature.
- Please list the things you would like to change for this feature; were there any incorrect information?
- How do you see the style of each concept (e.g. font sizes)?

14

- Now please go to the guideline page 7

15

Photos Questions

- What are your first impression when seeing the concepts? what disclosures have you made? Were you aware of them?
- Could you understand what this function does from the instructions?
- Please describe what sort of information was useful for you when reading about concepts close to your photos. What do you expect this feature to do?
- Is this going to change the way you share photos in the future? Have you already done action to limit this?
- How do you see the style of each concept (e.g. font sizes)?
- Please list the things you find most positive for this feature.
- Please list the things you would like to change for this feature; were there any incorrect information?

16

- Now please go to the questionnaire and answer the rest of the questions and then wait for the moderator to start the general discussion.

17

General discussion

- What sort of benefits do you see here to help controlling your privacy?
- What are the benefits of such a tool with the mentioned functionalities? What have you learned?
- Will you use this application? Why? Why not? Will you recommend it to others?
- In what situations would it make most sense to use?
- Lets consider about your expectations when it comes to protecting your personal information. Whether DataBait or any other ideal tool could succeed to do so? how?
- How do you like the future of the tool could look like?

18

Dream freely

- If you could dream freely in the area of security, how would the solution work. Never mind technical or economical hindrances which might exist today. Just tell me your ideas. Some science fiction ideas. ;-)

19

Missing points?

- OK, we have asked our questions now. Is there an important aspect we have missed to ask about?
- What is that?
- Why is this important?

20

Thank You

... & don't forget to claim your reward!

21

9.2.2. Workshop user guidelines

Page 1

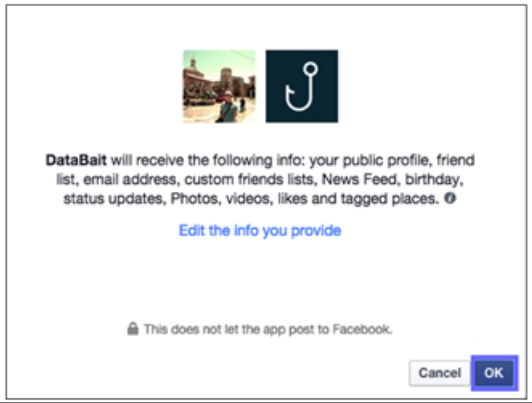
- Please go to http://bit.ly/1FQ9t2S and click on Register Now and follow the process until you see the screen in Page 2

Page 2

- Click on Login with Facebook. This will grant us temporary access to all the data stored in your profile. After clicking go to page 3.

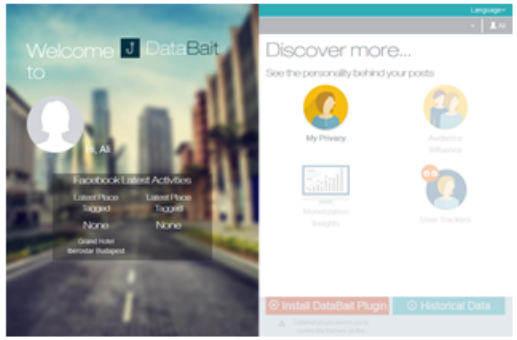
Page 3

- By clicking OK, your Facebook data is ready to be analyzed by the DataBait tool. After clicking go to page 4.



Page 4

- Once you have landed in this page, open this questionnaire link (http://svy.mk/1KC3ilY) and answer questions **1** to **9**. once completed wait for the workshop moderator to start the privacy and social media discussion. [If your answer to Question 3 was NO then only continue to Q4.]



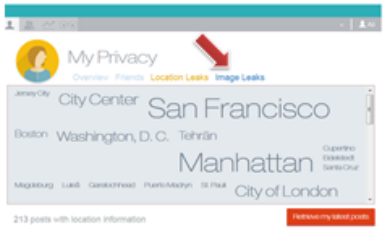
Page 5

- Click on My Privacy and then go to page 6

Page 6

- Click on Location Leaks
- Look at different 'concepts' presented there
- Click on a concept and see the text underneath
- Play a bit around on this page
- go back to the questionnaire and answer questions 10 to 12.(or 5-7)
- After answering the questions wait for the workshop moderator to start the open discussion.

Page 7



- Click on Image Leaks
- Look at different 'concepts' presented there
- Click on a concept and see the photos underneath
- Play a bit around on this page
- go back to the questionnaire and answer questions 13 to 15. (or 8-10)
- After answering the questions wait for the workshop moderator to start the open discussion.

9.2.3. Survey Questions

Welcome to DataBait Questionnaire.

This test contains 19 questions regarding the DataBait application.

*** 1. Please specify your name**

*** 2. Gender**

- Male
- Female
- Other

*** 3. After the registration process is complete answer the following.**

Most service providers combine very extensive and usually incomprehensible Terms of Service with a request for consent. USEMP offers users a contract that determines mutual obligations, the so-called Data Licensing Agreement (DLA).

Did you read (some of) the articles of the DLA?

- Yes
- No

4. Did you understand most of the information that you read in the DLA?

- Yes
- No

5. Which information did you find interesting?

	Yes	No
Information about which partners form the consortium		
Information about the obligations to which you commit in terms of downloading DataBait tools?		
information about the obligation to which the consortium partners commit in terms of providing profile transparency?		
information about the processing of your personal data?		
information about the purpose of processing your information (scientific research)?		

Which other information did you find interesting?

6. Which information did you miss in the DLA?

7. Please specify how did you perform the registration?

- Did not succeed to perform the task
- Task was hard to complete
- Completed the task with a little problem
- Performed task without any problems

Please specify any problems encountered

*** 8. The process of connecting with Facebook was...**

- not successful

- Task was hard to complete
- Completed the task with a little problem
- Performed task without any problems

Please specify any problems encountered

9. What did you feel when DataBait asked you to link your Facebook profile?

*** 10. Location leak function**

Please go to the main interface and click on 'My Privacy', then on the 'location leaks' link.

- Did not succeed to perform the task
- Task was hard to complete
- Completed the task with a little problem
- Performed task without any problems

11. In general, How do you feel about the accuracy of the concepts with respect to the presented text?

	Not at all	a little accurate	Accurate	Somehow accurate	Very accurate
How accurate?	○	○	○	○	◉

*** 12. What benefits can you see in this function from privacy point of view?**

- No benefit

- Little benefit
- Somehow beneficial
- Very beneficial

13. Multimedia function

Please go to the main interface and click on 'My Privacy', then on the 'Image Leaks' tab.

- Did not succeed to perform the task
- Task was hard to complete
- Completed the task with a little problem
- Performed task without any problems

*** 14. In general, How do you feel about the accuracy of the concepts with respect to the presented images?**

	Not at all	a little accurate	Accurate	Somehow accurate	Very accurate
How accurate?					

*** 15. What benefits can you see in this function from privacy point of view?**

- No benefit
- Little benefit
- Somehow beneficial
- Very beneficial

16. Besides the DLA, the USEMP Platform offers an information button and a button to withdraw your participation. Did you check the information behind the infobutton?

- Yes
- No

17. Please write us if you have encountered any specific problem during the test or any general comment that could help us improve the product.

Thank you very much for participating in this questionnaire.

9.3. iMinds Interview Topic List + Survey Questions

Topic List

- Introduction
  - User Profile Questions
- General Questions about Online Privacy
- DataBait Testing
  - Registration process + Survey Questions
  - Information Page + Survey Questions
  - Image Leaks Test + Survey Questions
  - Location Leak Test + Survey Questions
  - Tracker function

DataBait Pre-Pilot Survey Questions (Dutch)

User Profile Questions

Naam
Geslacht
Wat is jouw geboortejaar?
Hoeveel uur per dag breng je gemiddeld door op het internet (voor privé doeleinden)?
Via welke apparaten raadpleeg jij Facebook/surf je op het internet?
Wanneer ben je Facebook voor het eerst beginnen gebruiken?

Wat is jouw huidige beroepstoestand

- ❍ Leerling/ (doctoraats-) student
- ❍ Betaald werk (bediende, ambtenaar, zelfstandige, ...)
- ❍ Gepensioneerd (brugpensioen, pre-pensioen, enz.)
- ❍ Huisvrouw, huisman
- ❍ Op tijdelijk ziekte- of bevallingsverlof
- ❍ Met verlof zonder wedde/loopbaanonderbreking
- ❍ Arbeidsongeschikt, langdurig ziek, invalide
- ❍ Op zoek naar eerste werk
- ❍ Werkzoekend/werkloos
- ❍ Andere _____

Hoe vaak raadpleeg jij Facebook?

- ❍ Nooit
- ❍ Zelden
- ❍ Maandelijks
- ❍ Wekelijks
- ❍ Dagelijks
- ❍ Meermaals per dag

Duid hieronder aan in welke mate je akkoord gaat met de stellingen:

	Helemaal niet akkoord	Niet akkoord	Eerder niet akkoord	Neutraal	Eerder akkoord	Akkoord	Helemaal Akkoord
Het internet zorgt voor ernstige privacy problemen.	❍	❍	❍	❍	❍	❍	❍
In vergelijking met anderen ben ik gevoeliger voor de manier waarop online bedrijven omgaan met mijn persoonlijke informatie.	❍	❍	❍	❍	❍	❍	❍
Mijn privacy beschermen	❍	❍	❍	❍	❍	❍	❍

tegen online bedrijven is voor mij het meest belangrijke.							
Ik vind dat andere mensen veel te bezorgd zijn over online privacy.	❍	❍	❍	❍	❍	❍	❍
In vergelijking met andere onderwerpen, vind ik mijn privacy heel belangrijk.	❍	❍	❍	❍	❍	❍	❍
Vandaag de dag ben ik bezorgd over bedreigingen voor mijn privacy.	❍	❍	❍	❍	❍	❍	❍

Na dat je het registratieproces hebt doorlopen, gelieve volgende vragen te beantwoorden.

Hoe moeilijk vond je het om deze taak uit te voeren? (1= zeer moeilijk, 7= heel makkelijk)

- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7

Ik heb er vertrouwen in dat ik deze taak succesvol heb uitgevoerd

- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7

Welke informatie vond je interessant?

	Helemaal niet nuttig	Niet nuttig	Eerder niet nuttig	Neutraal	Eerder nuttig	Nuttig	Zeer nuttig
Informatie over de onderzoekspartners	❍	❍	❍	❍	❍	❍	❍
Informatie over de verbintenis die jij aangaat? (Het downloaden van de DataBait tools, installeren van de plugin)	❍	❍	❍	❍	❍	❍	❍
Informatie over de verplichtingen van de onderzoekspartners (aanbieden van profiel transparantie)	❍	❍	❍	❍	❍	❍	❍
Informatie over hoe wij jouw persoonlijke informatie zullen verwerken	❍	❍	❍	❍	❍	❍	❍
Informatie over het doel van het verwerken van jouw informatie (wetenschappelijk onderzoek)	❍	❍	❍	❍	❍	❍	❍

Nadat je op zoek bent gegaan naar extra informatie over het project, gelieve volgende vragen te beantwoorden

Hoe moeilijk vond je het om deze taak uit te voeren? (1= zeer moeilijk, 7= heel makkelijk)

- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7

Ik heb er vertrouwen in dat ik deze taak succesvol heb uitgevoerd

- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7

Welke informatie vond je interessant?

	Helemaal niet nuttig	Niet nuttig	Eerder niet nuttig	Neutraal	Eerder nuttig	Nuttig	Zeer nuttig
Het filmpje met meer informatie over hoe DataBait werkt	❍	❍	❍	❍	❍	❍	❍
De infographics met meer informatie over hoe DataBait werkt	❍	❍	❍	❍	❍	❍	❍
De informatie over de verschillende onderzoekspartners in het project	❍	❍	❍	❍	❍	❍	❍
De contactinformatie van de makers van DataBait	❍	❍	❍	❍	❍	❍	❍
Informatie over welke data van mij gebruikt wordt	❍	❍	❍	❍	❍	❍	❍
Informatie over de contracten	❍	❍	❍	❍	❍	❍	❍

Nadat je de Image Leaks functie hebt getest, gelieve volgende vragen te beantwoorden

Hoe moeilijk vond je het om deze taak uit te voeren? (1= zeer moeilijk, 7= heel makkelijk)

- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7

Ik heb er vertrouwen in dat ik deze taak succesvol heb uitgevoerd

- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7

Hoe makkelijk was het voor jou om deze functie terug te vinden?

- ❍ Het lukte me niet om dit te bekijken
- ❍ De taak was moeilijk, maar is wel gelukt
- ❍ Het is me gelukt, maar ben wel een klein probleem tegengekomen
- ❍ Ik heb deze functie terug gevonden zonder een probleem

Hoe accuraat vond je de concepten die aan jouw foto's werden verbonden?

- ❍ Helemaal niet accuraat
- ❍ Niet accuraat
- ❍ Eerder niet accuraat
- ❍ Neutraal
- ❍ Eerder accuraat
- ❍ Accuraat
- ❍ Zeer accuraat

Duid hieronder aan in hoeverre je akkoord gaat met onderstaande stelling

	Helemaal niet akkoord	Niet akkoord	Eerder niet akkoord	Neutraal	Eerder akkoord	Akkoord	Helemaal akkoord
Deze functie helpt mij mijn privacy beter te beschermen	❍	❍	❍	❍	❍	❍	❍
Ik word meer bewust van wat ik allemaal online deel door deze functie	❍	❍	❍	❍	❍	❍	❍
Door deze functie zal ik mijn gedrag in de toekomst aanpassen	❍	❍	❍	❍	❍	❍	❍

Nadat je de lokatie leaks functie hebt getest, gelieve volgende vragen te beantwoorden:

Hoe moeilijk vond je het om deze taak uit te voeren? (1= zeer moeilijk, 7= heel makkelijk)

- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7

Ik heb er vertrouwen in dat ik deze taak succesvol heb uitgevoerd

- ❍ 1
- ❍ 2
- ❍ 3
- ❍ 4
- ❍ 5
- ❍ 6
- ❍ 7

Hoe makkelijk was het voor jou om deze functie terug te vinden?

- ❍ Het lukte me niet om dit te bekijken
- ❍ De taak was moeilijk, maar is wel gelukt
- ❍ Het is me gelukt, maar ben wel een klein probleem tegengekomen
- ❍ Ik heb deze functie terug gevonden zonder een probleem

Hoe accuraat vond je de concepten die aan jouw foto's werden verbonden?

- ❍ Helemaal niet accuraat
- ❍ Niet accuraat
- ❍ Eerder niet accuraat
- ❍ Neutraal
- ❍ Eerder accuraat
- ❍ Accuraat
- ❍ Zeer accuraat

Duid hieronder aan in hoeverre je akkoord gaat met onderstaande stelling

	Helemaal niet akkoord	Niet akkoord	Eerder niet akkoord	Neutraal	Eerder akkoord	Akkoord	Helemaal akkoord
Deze functie helpt mij mijn privacy beter te beschermen	❍	❍	❍	❍	❍	❍	❍
Ik word meer bewust van wat ik allemaal online deel door deze functie	❍	❍	❍	❍	❍	❍	❍
Door deze functie zal ik mijn gedrag in de toekomst aanpassen	❍	❍	❍	❍	❍	❍	❍

Duid hieronder aan in hoeverre je akkoord gaat met onderstaande stelling

	Niet akkoord	Eerder niet akkoord	Neutraal	Eerder akkoord	Akkoord
Ik denk dat ik dit systeem graag regelmatig wil gebruiken	❍	❍	❍	❍	❍
Ik vond het systeem onnodig complex	❍	❍	❍	❍	❍
Ik vond het systeem makkelijk te gebruiken	❍	❍	❍	❍	❍
Ik denk dat ik ondersteuning nodig heb van een technisch persoon om dit systeem te kunnen gebruiken	❍	❍	❍	❍	❍
Ik vond dat de verschillende functies in dit systeem erg goed geïntegreerd zijn	❍	❍	❍	❍	❍
Ik vond dat er teveel tegenstrijdigheden in het systeem zaten	❍	❍	❍	❍	❍
Ik kan me voorstellen dat de meeste mensen zeer snel leren om	❍	❍	❍	❍	❍

dit systeem te gebruiken.					
Ik vond het systeem er omslachtig in gebruik	❍	❍	❍	❍	❍
Ik voelde me erg vertrouwd met het systeem	❍	❍	❍	❍	❍
Ik moest erg veel leren voordat ik aan de gang kon gaan met dit systeem	❍	❍	❍	❍	❍

Duid hieronder aan in hoeverre je akkoord gaat met onderstaande stelling

	Niet akkoord	Eerder niet akkoord	Neutraal	Eerder akkoord	akkoord
Het gebruik van DataBait helpt me om sneller inzicht te verwerven in mijn online data	❍	❍	❍	❍	❍
DataBait helpt me om meer controle te krijgen over mijn online data	❍	❍	❍	❍	❍
Ik vind DataBait een handige tool	❍	❍	❍	❍	❍
Deze tool maakt het gemakkelijk om inzicht te verwerven in mijn privacy	❍	❍	❍	❍	❍
Wanneer DataBait beschikbaar komt, zou ik deze tool op regelmatige basis gebruiken	❍	❍	❍	❍	❍
Ik vind dat DataBait me een meerwaarde biedt om inzicht	❍	❍	❍	❍	❍

te krijgen in mijn online privacy De functies die DataBait aanbiedt geven me kwaliteitsvolle output	❍	❍	❍	❍	❍

10. References

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, *24*(6), 574–594. http://doi.org/10.1080/10447310802205776

Bangor, A., Kortum, P. T., & Miller, J. T. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, *4*(3), 114–123.

Belanger, F., Hiller, J. S., & Smith, W. J. (2002). Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. *The Journal of Strategic Information Systems*, *11*(3–4), 245–270. http://doi.org/10.1016/S0963-8687(02)00018-5

Bergvall-Kareborn, B., Holst, M., & Stahlbrost, A. (2009). Concept Design with a Living Lab Approach. In *42nd Hawaii International Conference on System Sciences, 2009. HICSS '09* (pp. 1–10). http://doi.org/10.1109/HICSS.2009.123

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability Evaluation in Industry*, *189*(194), 4–7.

Bryman, A. (2012). *Social Research Methods, 4th Edition* (4th edition). Oxford ; New York: Oxford University Press.

Caballero, R. (2008). Theorizing about genre and cybergenre. *CORELL: Computer Resources for Language Learning*, *2*, 14–27.

Goodwin, K., & Cooper, A. (2009). *Designing for the Digital Age: How to Create Human-Centered Products and Services* (1 edition). Indianapolis, IN: Wiley.

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, *24*(2), 105–112. http://doi.org/10.1016/j.nedt.2003.10.001

Ioinson, A. N., & Paine, C. B. (2007). Self-disclosure, privacy and the Internet. *The Oxford Handbook of Internet Psychology*, 2374252.

Jourard, S. M., & Lasakow, P. (1958). Some factors in self-disclosure. *The Journal of Abnormal and Social Psychology*, *56*(1), 91–98. http://doi.org/http://dx.doi.org/10.1037/h0043357

Nielsen, J., & Landauer, T. K. (1993). A Mathematical Model of the Finding of Usability Problems. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems* (pp. 206–213). New York, NY, USA: ACM. http://doi.org/10.1145/169059.169166

Orfanou, K., Tselios, N., & Katsanos, C. (2015). Perceived Usability Evaluation of Learning Management Systems: Empirical Evaluation of the System Usability Scale. *International Review of Research in Open and Distributed Learning*, *16*(2), 227–246.

Padyab, A. M. (2014). Getting More Explicit On Genres Of Disclosure: Towards Better Understanding Of Privacy In Digital Age (Research In Progress). *Norsk Konferanse for Organisasjoners Bruk Av IT*, *22*(1). Retrieved from http://ojs.bibsys.no/index.php/Nokobit/article/view/48

Palen, L., & Dourish, P. (2003). Unpacking privacy for a networked world. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 129–136). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=642635

Preece, J., Sharp, H., & Rogers, Y. (2015). *Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons.

Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: how to plan, design and conduct effective tests*. John Wiley & Sons.

Teague, R., De Jesus, K., & Ueno, M. N. (2001). Concurrent vs. post-task usability test ratings. In *CHI'01 extended abstracts on Human factors in computing systems* (pp. 289–290). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=634238

Tomasello, M. (2010). *Origins of Human Communication*. Mit Press.

Tullis, T. S., & Stetson, J. N. (2004). A comparison of questionnaires for assessing website usability. In *Usability Professional Association Conference* (pp. 1–12). Retrieved from http://home.comcast.net/~tomtullis/publications/UPA2004TullisStetson.pdf

Wiliam, D., & Black, P. (1996). Meanings and Consequences: a basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, *22*(5), 537–548. http://doi.org/10.1080/0141192960220502

Yates, J., & Orlikowski, W. J. (1992). Genres of Organizational Communication: A Structurational Approach to Studying Communication and Media. *The Academy of Management Review*, *17*(2), 299. http://doi.org/10.2307/258774

11. Table of Figures