Facebook Users Attitudes towards Secondary Use of Personal Information

Completed Research Paper

Ali Padyab
Luleå University of Technology
Luleå, Sweden
ali.padyab@ltu.se

Tero Päivärinta
Luleå University of Technology
Luleå, Sweden
tero.paivarinta@ltu.se

Anna Ståhlbröst
Luleå University of Technology
Luleå, Sweden
anna.stahlbrost@ltu.se

Birgitta Bergvall-Kåreborn
Luleå University of Technology
Luleå, Sweden
birgitta.bergvall-kareborn@ltu.se

Abstract

This paper reports on a study of how user attitudes to institutional privacy change after exposing users to potential inferences that can be made from information disclosed on Facebook. Two sets of focus group sessions with Facebook users were conducted. Three sessions were conducted by demonstrating to the users, on a general level, what can be inferred from posts using prototypical software called DataBait. Another set of three sessions let the users experience the potential inferences from their own actual Facebook profiles by using the DataBait tool. Findings suggest that the participants' attitudes to secondary use of information changed from affective to cognitive when they were exposed to potential third-party inferences using their own actual personal information. This observation calls for more research into online tools that allow users to manage and educate themselves dynamically about their own disclosure practices.

Keywords: Secondary use of personal information, On-line Social Network, Facebook, Privacy, Attitude, Genre of Disclosure, Institutional Privacy, Privacy awareness

Introduction

Facebook is a global Online Social Network (OSN) site that enables users to present themselves via an online profile shared with their friends and encourages them to get involved in varying social activities online: to upload images, post comments, check-in to places, tip about hobbies and update status etc. In other words, users are encouraged to share their personal information, information about their networks and relations, and information about their on-line behavior and personal preferences through the platform. Some privacy risks associated with the use of social media (e.g. possible access of future employers to personal information) are commonly known, nevertheless, users post large amounts of information that could be traced back to them (Gross and Acquisti 2005; Kaspersky Lab 2016). Continuance of this sharing behavior depends on such factors as trust in the OSN site in question, perceived greater benefit gained in the exchange of losing privacy, and unawareness of the possible privacy implications of personal information disclosure (Acquisti 2004; Donath and Boyd 2004).

Facebook's possession and use of personal information can be defined as an "iceberg", also known as The Facebook Iceberg Model, which separates the visible part from the invisible (Debatin et al. 2009). The visible part (the small top), consists of interactions among end users with Facebook and each other, whereas

the invisible part (the big bottom) consists of Facebook's marketing profiling, data aggregation, third party sharing, and behavior surveillance. Some previous studies have focused on the visible part (e.g. Boyd and Hargittai 2010; Lipford et al. 2008; Tuunainen et al. 2009), however, fewer have examined the end user relation to the invisible part, i.e., user awareness and attitude toward varying secondary uses of personal information.

Secondary use of personal information has been studied in varying information system contexts, including OSNs. The concept involves "the use of personal information for other purposes subsequent to the original transaction between an individual and an organization when the information was collected." (Culnan 1993, p. 342). Secondary use, when recognized by the person whose information is concerned, correlates with privacy concerns about OSNs (Krasnova, Günther, et al. 2009; Soczka et al. 2015). The authors adhere to the notion presented in previous research that user activity for protecting privacy increases after negative experiences of information disclosure. (Yang 2012; Debatin et al. 2009). This situation is compounded by the fact that social media users are reported to lack understanding of business models of OSNs and of how their personal information is processed (Orito et al. 2014). While behavioral advertising continues to grow and to become more privacy invasive, user awareness lags behind. It is vital to emphasize that a large amount of data gathered about an individual, who is often unaware of its leakage and is not able to monitor, protect, and control it, is by nature a sensitive issue and debatable (Butler 2007; Kosinski et al. 2013; Narayanan and Shmatikov 2009). Consequently, this research adopts the value assumption that users should be made aware of secondary uses of personal information on OSNs and educated about online privacy (Krishnamurthy et al. 2011).

This paper reports on a study within the context of an EU-FP7 project called USEMP[1]. The USEMP framework encompasses disciplines and technologies that are relevant in understanding aspects of personal data and develops assistance tools for improved personal data sharing, management, and user awareness about sharing such data (Popescu et al. 2016).

As part of the USEMP project, the focus of this study was on whether and how improved understanding of secondary use of personal information influences user attitude towards privacy and disclosure. This study complements previous literature by enlightening end users about the sensitivity of their *own* personal data and its commercial value through a novel tool called DataBait[2]. Findings suggest that the use of tools that illustrate opportunities for secondary uses of personal information have an impact on user awareness of the risks related to sharing personal data and, thus, change their attitudes towards disclosure of personal information on OSNs.

The remainder of the paper is designed as follows: An overview of the discourse within the field of privacy research and secondary use of information is presented first. This is followed by the empirical set up with the DataBait tool and description of the research process through two sets of focus group sessions. After the findings of the study are presented, the contributions to and implications for end user education of the adverse effects of personal information disclosure are discussed. The paper is concluded with observations of the limitations of the study and proposed ideas for further research.

Background

Personal information shared on OSNs is a key area of privacy research. Legally, personal information has been broadly defined as any information that can be used to identify a person both directly or indirectly. For instance, Article 2 of the European Union Directive 95/46/EC emphasizes that personal data is any information that can directly or indirectly identify a person through factors such as physical, physiological, mental, economic, cultural, or social identity (EU 2011, p. 3). In the US, the essence of personally identifiable information has also been stressed in law involving any piece of information that can be used to directly or indirectly distinguish an individual. For example, the E-Government Act of 2002, Section 208 defines personal information as: "[a]ny representation of information that permits the identity of an

[1] www.usemp-project.eu, full title: User Empowerment for Enhanced Online Presence Management
[2] www.databait.eu, for an overview and positioning of the USEMP project and the tool, see Popescu et al. (2016).

individual to whom the information applies to be reasonably inferred by either direct or indirect means." (E-Government Act of 2002 n.d.).

Related research has largely focused on the directly recognizable entities of personal information, such as the public display of a person's name with a profile picture or different demographics (e.g. age, gender, location, and marital status) (Acquisti and Gross 2006; Hum et al. 2011; Liu et al. 2011). Such direct information is the first thing a user fills in when creating a profile, and, therefore, it is easily accessible. Although Facebook provides some features to control the accessibility of such information, a study shows that, despite participants' perceived high level of knowledge and skills in online privacy management on Facebook, the majority of the posted information was, nevertheless, shared beyond initial user intention (Suh and Hargittai 2015). This was because users are sometimes unaware of the potential audiences of their published posts (Johnson et al. 2012). On the other hand, indirect information, such as posts and pictures that could be used as a basis to extract information about, for example, personal habits, sexual orientation, or demographics, has been less in focus since studies in this area have mostly examined profiling algorithms (e.g. Liu and Terzi 2010; Theodoridis et al. 2015).

Direct and indirect information sharing can be viewed within the contexts of *social* and *institutional privacy*. Raynes-Goldie (2010) have defined the former as "the control of information flow about how and when personal information is shared with other people". Whereas institutional privacy is defined as "how institutions such as governments, banks, and other businesses, use or misuse the personal information." (Raynes-Goldie 2010). User attitude towards institutional privacy has gained less attention within the context of indirect information. The majority of studies concentrating on secondary usage assume that users are fully aware of these practices or should make hypothetical assumptions about their inferred attitudes by an OSN. It has been highlighted that human behavior is predictable through the digital records held by OSNs that make individuals more uniquely distinguished from each other (Kosinski et al. 2013). Individuals are susceptible to the harms caused by institutions and third parties and research indicates that users do not pay much attention to this (Brandtzæg et al. 2010; Raynes-Goldie 2010; Young and Quan-Haase 2013). Young and Quan-Haase (2013) argue that users seem to be helpless when faced with institutional privacy: "Little concern was raised [by users] about institutional privacy and no strategies were in place to protect against threats from the use of personal data by institutions. This is relevant for policy discussions, because it suggests that the collection, aggregation, and utilization of personal data for targeted advertisement have become an accepted social norm." (Young and Quan-Haase 2013, p. 479).

Palen and Dourish (2003) define the concept *genres of disclosure* as "socially-constructed patterns of privacy management". Upon the moment of disclosure, e.g. on an OSN, an individual needs to be able to find a balance between three boundaries: of self and other, of privacy and publicity, and of past and future. These arrangements of coming to a conclusion whether to decide disclose something or to limit the depth and breadth of a single communicative action can, thus, be characterized through genres of disclosure recognized by the user. However, problems do rise from the fact that it is difficult to abide to a genre without being aware of personal information being potentially misappropriated after it is stored by an OSN. Institutional practices can potentially affect an individual's privacy by manipulating sensitive information in certain situations, such as citizens occasionally targeted for surveillance or information sold to data aggregators (Padyab 2014). In this paper, it is argued that individuals participating in digital media make decisions based on improper knowledge about the boundary of private and public as it is unclear to them what could potentially breach the boundary behind the system in question (e.g. Facebook). The awareness of possible invasions of an individuals' personal information (through readily recognized or plausible ways) can lead to more informed communication choices. For example, if a user is made aware that by posting comment X on Facebook, it could be inferred that they belong to a certain religion, or it may reveal something of more or less sensitive habits. Such awareness could have an impact on the communicative decisions made by the individual at that moment or in the future.

Derived from psychology studies, *attitude* has been the focus of privacy research for some time. An attitude is "an evaluative integration of cognitions and affects experienced in relation to an object" (Crano and Prislin 2006, p. 347). The *affective* and *cognitive* components are two determinants of attitudes (Abelson et al. 1982). The affective component contributes to the emotional, sensational, and feeling-related aspects of an attitude while the cognitive component involves the subject's rational reactions to the object of the attitude. The importance of studying privacy attitudes is the power of these attitudes over privacy behavior.

For example, Dienlin and Trepte (2015, p. 294) have concluded that "informational, social, and psychological privacy attitudes are significantly related to informational, social, and psychological privacy behaviors". Both cognitive and affective reactions to information privacy are equally important in understanding the psychological reactions of individuals to invasions of privacy (Choi et al. 2012). The Information Systems (IS) literature on privacy has seen a large number of studies conducted on user attitudes towards sharing information on OSNs especially through the cognitive component of attitude (i.e. cost-benefit analysis and risk perception) while the study of attitudes over secondary uses of information remains scant (Adjei and Olesen 2012; Iyilade et al. 2015; Soczka et al. 2015).

Privacy awareness is defined as the extent to which users are informed about privacy problems, violations, and procedures on OSNs (Nemec Zlatolas, Welzer, Heričko, & Hölbl 2015). Previous studies have shown that general-level awareness (e.g. media coverage) has a significant impact on self-disclosure and privacy concerns (ibid). Bateman, Pike, & Butler (2011) showed that an individual's intention to self-disclose items related to user likes and affiliations is impacted by the perceived publicity of OSNs. The present study is motivated by Hull, Lipford, & Latulipe's (2010) call for more research on user awareness of the invisible part of Facebook. They suggest that "users need to be aware of what, exactly, might happen with their and their friends' information, in order to make informed decisions about how to share that information" (ibid, p. 300).

Current literature lacks the means to capture the nuances of user attitudes toward secondary use. Little or no previous research has been found that measures attitudes contextualized to a user's own information disclosure setting on OSNs. It seems that the previous studies on secondary use have been largely based on researcher-centric pre-conceptions of privacy concerns, often perhaps a bit disconnected from the users' own sense-making of privacy. For example, data from users has been collected through simple questions (e.g. Dinev and Hart 2006; Son and Kim 2008), fictional case scenarios (Culnan 1993; Krasnova, Hildebrand, et al. 2009), or stimuli for conjoint analysis of an imaginary financial portal based on a participant's judgment of a set of alternatives of privacy invasive scenarios (Hann et al. 2007). Such approaches alone do not fully capture the phenomenon if privacy concerns do not take place in the context of the end-user's own OSN actions and experience, and this area still requires better means to explore the end user's demeanor. This explains the present involvement in a study in which end users are no longer examined through simple speculative questions or scenarios about secondary use, but are exposed to instantiations of their own disclosures.

The next section describes how the DataBait tool was used in a focus group setting in which Facebook users were introduced to the possible breaches of their online profile through image and text mining algorithms.

Methodology

Exploratory focus groups were conducted in this research with the goal to capture participant attitudes towards disclosure by means of an OSN after being exposed to illustrations of potential secondary uses of information from their Facebook profiles. The focus group method helps to identify and clarify emerging concepts through a group discussion (Edmunds 2000; Belanger 2012). The method also allows for the observation of change of attitude (if any) through what Morgan (1996) suggests as the "synergistic effects" of focused interactions, which can provide greater insights than the sum of individual interviews. In this research, the participants could follow up on each other and then explain their evolving view, which made the emerging concepts more traceable for the researcher. For example, participants were asked to give examples of what they had experienced in relation to their privacy to make the other participants more familiar with the theme and also to trigger them to tell their own interesting stories. This discussion then helped to capture the transformation of attitude that occurred influenced by awareness through the course of the focus groups.

Focus Group Procedure and Data Collection

This research is based on empirical data stemming from two separate rounds, or studies, of focus group interviews. In the first study, the focus was to present and jointly discuss an illustrative mock-up of the DataBait tool, and, in the second study, the participants could test and experiment with the tool based on their own Facebook profiles. The aim of this approach was to investigate participant attitudes related to privacy when exposed to a presentation of possible conclusions that could be drawn based on their Facebook

shares (Study 1) versus their attitudes when they experienced possible inferences made on their own Facebook photos and geographical locations (Study 2). The first round of focus groups was carried out in February and March 2015 and consisted of three group sessions with twelve participants in total. In the second round, conducted in August 2015, three sessions with fifteen participants were carried out. The sessions were in English and were situated on a university campus. The participants were recruited via a Living Lab[3] and by inviting students and employees at the university to participate. In the pursuit of dependability of the process of inquiry, participants overlapped from Study one to Study two. Bélanger and Crossler (2011) argue that privacy research is heavily reliant on student-based samples, and this need to be alleviated. For this reason, non-students were recruited as well. After receiving an initial expression of interest and analyzing the candidate profiles, a mixture of participants in terms of occupation (13 students, 14 non-students), educational background (5 high school, 11 BA and 11 MA level), gender (15 male, 12 female), cultural backgrounds (14 Swedish, 13 non-Swedish) and age (from 18 to 58 years old) were invited. The reason for selecting a panel of people with varying backgrounds was to minimize bias imposed by a specific demographic (Bouma et al. 1995). However, the intention of this paper is not to provide stratification of the findings based on demographic variation, albeit that would be interesting for future research. At this phase, it is not argued that the findings represent thoroughly tested theoretical knowledge, instead the findings imply a theoretically interesting proposition and motivation for further research to be complemented with other methodological means.

To stimulate the discussion in the group, the notion of *genres of disclosure* was used to capture what users disclose and account for situations of potential privacy violations. Palen and Dourish (2003, p. 6) define a genre of disclosure as "the relationship between *forms of disclosure* and *expectations of use*". For this reason, the situations where expectations of use are misaligned with the potential use (e.g. in our case inferences made from one's private photos) are called violations. The two notions of 'forms of disclosure' and 'expectation of use' were deemed appropriate to design the general-level, semi-structured focus group protocol (Appendix). The former notion was captured by asking general level questions. Examples of these are: What do you think that you disclose on the Internet in your daily activities or what do you think Facebook and Google know about you? The latter notion was captured through the introduction of (possibly) new forms of disclosure (e.g. location traces) using the DataBait tool and, thus, changing the expectation of use. Normally, disclosure activity is based on the expectation of appropriate use and any identified deviations can be regarded as abusive or otherwise disruptive changes to such genres (ibid). Observed changes in attitude can be viewed as raised awareness regarding institutional privacy, which adheres to the goal to investigate the extent to which this awareness can affect an individual's disclosures on OSNs.

The questions were first piloted internally within the research team and with the project manager, who was not part of this research, to determine if the questions were understandable and helped assisted the authors in recognizing biases (Shenton 2004). Based on this piloting, a semi-structured interview guide consisting of both open-ended questions as well as specific questions, as presented in the appendix, was created.

Since little, if any, in-depth previous research on the institutional privacy of OSNs was found, the exploratory nature of research demanded an openness to unexpected findings (Wilkinson 1998). Little control over focus group research could be seen as a benefit in order to give greater opportunity to the participants to 'develop the themes most important to them' (Cooper et al. 1993). The principle of minimum control led the authors to find interesting themes in the first focus group. A lot of affective responses were observed among the participants, such as being scared, curious, annoyed and shocked, or surprised by secondary use. One part of the affective reactions was due to the discussions driven by group interactions (e.g. hearing about privacy-related experiences) and other part was related to the demonstration of the tool. The affective responses then shaped the focus of subsequent group interviews by investigating in depth the attitudes of participants towards institutional privacy and its building components, i.e. affective and cognitive components.

The focus group sessions started with a short introduction of the facilitator, the practicalities of the session, and a chance for each participant to introduce themselves and to explain motivations for participation. They were informed that all information gathered during the discussions was to be analyzed based on themes

[3] www.testplats.com

and individual responses and would be anonymized if anything was quoted direct quote. The duration of a focus group interview was, on average, 100 minutes. The participants were then familiarized with the idea of secondary use to get everyone on the same page. The participants were free to raise and discuss issues and concepts of secondary use that they regarded as the most important, with minimum influence from the moderator to reach credibility in the study, as described by Lincoln and Guba (1985). This was followed by a general discussion about their use of social media, the type of information participants disclose, and privacy concerns related to social media. This phase counted as attitudes *before* being exposed to potential inferences, by looking at *cognitive* and *affective* aspects of the attitudes of the participants. Participant awareness about indirect information sharing was examined by gathering insights into what personal information they think they reveal through Facebook. The focus was on why participants share their locations and/or photos and the impact of *institutional privacy* over their shares. A brief introduction of the DataBait tool followed this (more details are given below). The participants saw screenshots of the tool with no active interaction with their own data. Finally, a group discussion took place that focused on attitudes towards secondary information use and its effects on the private information disclosures of the participants. This phase counted as *after* by gathering data about *affective* and *cognitive* aspects of participant attitudes. The impacts of *awareness* on possible secondary uses of participant information and the intended attitude of future disclosure were also examined.

In the second round of focus groups, the overall flow of tasks was designed in a way that participants would be able to compare their (intended) privacy preferences with their actual behaviors (i.e. photos and/or locations shared) to capture if any change of attitude had occurred. The design of the workshop was identical to Study 1 except that participants used the DataBait tool by themselves on their own profiles. First, participants discussed their attitudes towards photo and location sharing and then used the tool to analyze their Facebook profile. The tool visualized some possible inferences that could be made to predict their locations (based on posts) and extracting concepts from their photos (similar to what is presented in Figures 1 and 2). After seeing the result of potential leaks, participants reflected upon the result and to what extent it was close to their preconception of institutional privacy. The focus group was conclude with a discussion that focused on participant attitude towards secondary information use and its effects on their information disclosure. All six sessions were transcribed verbatim from the audio recordings captured during the sessions.

DataBait Tool

The DataBait tool is the result of the USEMP project, and its functionality has been peer-reviewed a number of times by external reviewers of the European Council as well as academic reviewers involved in publications from the project. The project adopted Facebook as a use case. Once a user registers with DataBait and links one of his/her external Facebook accounts, DataBait requests access to the content and messages that the user shares through the Facebook account. As a result, DataBait gains access to the media content (photos) and activity data (status updates, comments) that the user shares through the linked Facebook API. The platform, in this respect, allows personal data extraction (e.g. locations) and classification after the initial processing through advanced algorithms and interprets these results in order to provide feedback to the end-user (i.e. designate whether the automatically detected results are meaningful for friends, for Facebook, and for third parties, such as advertisers). DataBait data analysis, thus, generates inferred data. It is not claimed here that the inferences made using this tool are exactly what Facebook inducts, however, it is not far from the mark since Facebook privacy policy is long, complex, and allows the company to conduct similar analyses (cf. Fuchs 2013).

In this study, two services were used to illustrate user privacy inferences: Image Leaks and Location leaks.

Image Leaks/Visual Concept Mining Tool provides the user with the list of concepts that can be inferred from the images the user has uploaded or shared with others on Facebook. These visual concepts can act both as an indicator of privacy (what algorithms can infer about 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 a user has posted on Facebook. DataBait predicts tags from a set of over 17,000 visual concepts (Ginsca et al. 2015). 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 in which this concept has been detected are shown along with a measure of confidence for the detection from the corresponding algorithm. A screenshot of the image leaks function is illustrated in Figure 1.

Figure 1. Concepts retrieved from a Facebook user uploaded/shared images presented in concept cloud and a complete view over the top 20 most repeated visual concepts with different colors to easily distinguish the bubbles

Location Leaks/Prediction gives the user a list of locations that can be inferred from the posts the user has shared with others on 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 these posts refer to or the location indicated on Facebook by the user. The tag cloud shows the identified locations on the city level with a size proportional to their frequency. Tags are colored with different colors 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 2).

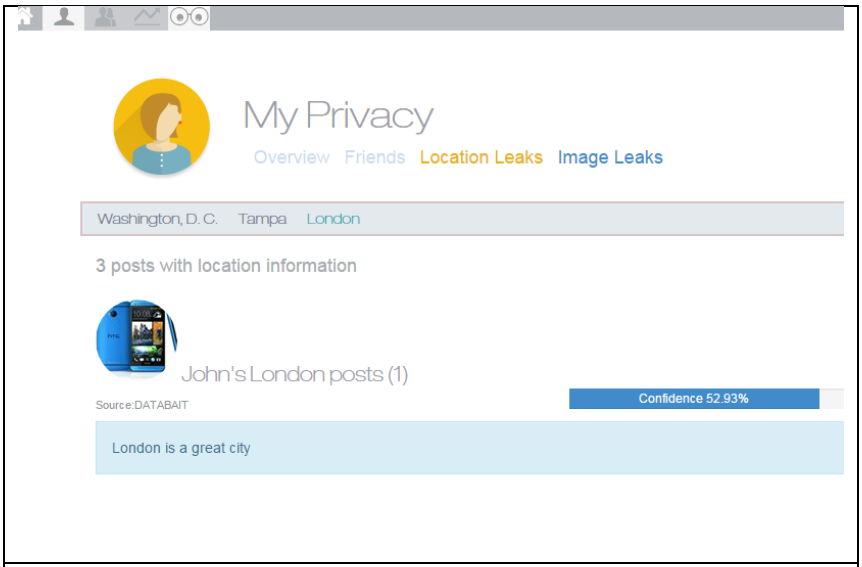

Figure 2. Users are able to obtain a holistic view of their leaked locations along with the related post

Data Analysis

This study has used qualitative content analysis with a deductive category application to support the process of analyzing data (Mayring 2000). With this approach, the researcher determines the initial coding scheme based on identified core categories (Potter and Levine-Donnerstein 1999) with the objective of investigating the relationships among these concepts. To support the deductive application of categories, the core concepts were chosen to be affective and cognitive attitudes both before and after the presentation and usage of the described tool. After that, sub-categories under the core concepts were identified according to the principles of inductive content analysis (cf. Elo and Kyngäs 2008). In this way, it was possible to identify changes in both constructs influenced by the presentation as well as by actual usage of the tool. The attitudes (affective and cognitive) related to information disclosure through Facebook and attitudes (affective and cognitive) towards institutional privacy, respectively, were analyzed. Table 2 presents a summary of the resulting categorization matrix (cf. Elo and Kyngäs 2008). An additional coding categorization of direct and indirect information emerged later during the data analysis as a concept describing the level of awareness of self-disclosure. The analysis of the transcripts was conducted in NVivo, which made the analysis process more manageable by organizing meanings assigned to text selections in the form of nodes or concepts. Associations between nodes created higher level categories, which facilitated detection of trends across the collection.

Three researchers were involved in the data analysis. First, the authors read 15% of the whole transcription independently to acquire an overview of participant attitudes towards disclosure and institutional privacy both before and after the introduction of the DataBait tool separately. Sample codes extracted from the transcription are presented in Table 1.

Table 1. Example codes from transcription

Excerpt	**Codes**
nowadays I just try to have like things like not so, very special photos, these photos of me, but maybe something you could use for anything or maybe I mean use the picture of me standing in front of something is not bad, you could use that much maybe and but yeah, so lately I have deleted a lot of things on Facebook, but I am afraid they still have it, at least you can´t see it as a private person maybe Facebook has it but at least people who visit my Facebook don't see it…	Keeping personal information private Being afraid
I think it is good to sort of visualize to see what is happening behind the scene, because this is something that the end-users never get to see usually, you know behind Facebook´s wall, so in visualizing and seeing what is exactly happening, it might make you think differently about sharing photos.	What others know about me Unintended disclosure Being interested
Then I think it would be interesting. Because, for example, if you see a lot of kids in the picture, then you know you would be probably at risk or targeted by companies wanting to sell children´s stuff, that would be private to me though, you know, would be more private than knowing places I am, for example.	Awareness Avoiding unintended disclosure Unintended disclosure Clearer picture of commercialization Being interested

Second, to reach a common view of the main themes and acquire a shared understanding, all categories were iteratively discussed within the research group. This also contributed to a transferability of the study

both by reaching a consensus about the interpretations generated by the data and by gathering demographic information about the participants with careful documentation and transcriptions from the group sessions to allow such comparisons in the future, if desired (Lincoln and Guba 1985; Shenton 2004). Third, after reaching agreement consensus among the researchers, the rest of the transcription was analyzed by the first author, and, when uncertainties arose, all authors read the transcript and discussed it to ensure a high level of credibility of the analysis. Although all six focus groups were analyzed, after analyzing the fifth transcription, no new concept emerged, and, therefore, it was agreed that at least an initial saturation of concepts was reached and the amount of gathered data was found to be reasonable (Tong et al. 2007).

Result

Participant attitudes towards information disclosure in relation to institutional privacy revealed three important findings. First, it was found that participants were more affective than cognitive towards institutional privacy, and this changed from affective to cognitive when they practiced the inferences on their own data. Second, it was found that by comparing Study 1 and 2, the attitude of participants towards disclosure changed when confronted with inferences made from their own Facebook contents compared to the participants who were only told that it was possible to make inferences from their Facebook profile. Third, indirect information sharing was found to be propagated through different channels of communication in Facebook, which is beyond user awareness. These three findings will be explained in detail below.

Attitudes towards disclosure and institutional privacy

The two constructs of affective and cognitive attitudes towards disclosure of photos and locations over Facebook and also attitudes towards institutional privacy were examined.

Attitude towards disclosure

All participants stated from the beginning that they are generally aware of what they disclose on Facebook and try to keep their personal information private. The main reasons for sharing photos and their locations were affective, such as fun, interesting photo and/or location to share. They were cautious based on feelings of whether they liked certain content to be shared and to whom they liked it to be available. One of the observed trends was that participants felt indifferent to sharing their home address. For example, in Sweden this information is public and anyone can look up online to see where someone lives, his/her marital status, who he/she lives with and his/her phone number.

"*Nowadays you can locate a person on the internet, their phone number and address and things like that, Facebook even that especially I would, I do not write my address or anything like that, but people find it anyway*."

The cognitive aspect of disclosure relates to personal preferences for sharing photos and locations. Users are selective based on cognition (and not just feelings) with regard to impacts on other people's privacy if they are mentioned in the photo and/or location, sensitivity of the content, and limiting access to friends. They form a cognitive policy approach in their mind, which rules out their disclosure practices. For example, some participants mentioned that they have developed rules for sharing only photos that include no children.

"I have two small children and we have a family policy not to share our children on Facebook, but there are other people, who have connections with us they could take pictures of our children and post them, they are on Facebook, we don't. So, I try to think about what to put out there because I don't think it depends a lot about what settings I set on Facebook, if it is shown or not shown for others."

Affective attitudes towards sharing, when confronted with the tool, converged to 'awareness' as the result of acquiring insights into direct and indirect disclosure, i.e. knowing that the personal information could be inferred by different means other than revealing it directly. The combination of all inferences made from locations and photos and a summary of them was of interest to the participants, as the tool could give predictions of their hobbies and hints of their personality traits. Therefore, participants felt more frightened

and predicted to become more cautious and to give their decisions a second thought before posting anything in the future. Some participants even showed interest in a tool that would make them more pre-aware of inferences in parallel before posting anything to Facebook. One participant in the first study mentioned:

"Oh, that would be so cool and like a pop-up thing on your computer like,"Oh, this would be the consequences of your post, your picture, or like, put up a picture, and then, like, it comes up and this message, like, "you can see a beer in the background, have you seen that." Blah blah I can see that it is a Heineken and, like, "Do you really want to post this photo… Like, you know before what you post."

Avoiding unwanted disclosure and inferences were the main cognitive aspects of user attitude towards disclosure after use of the DataBait tool. It alerted them to see that previous disclosure could result in unwanted inferences, against their initial intention. It was bothering for the participants to see that unexpected inferences could be made from their content. Although the inferences were legitimate, they could cast an undesired image of the user. Even if no unwanted disclosure happened, the participants, by and large, wanted to have some assent through an automatic mechanism that would somehow assure them that nothing undesirable could be inferred from the information ready to share. Some participants in the workshops, in turn, mentioned that the inferences analysis for their part assured them that their photos and/or location shares were not interpreted otherwise (at least by the DataBait tool). For example, one participant in Study 2 noted:

"*The best case scenario for this tool [in the future], is that it infers something that I would never have thought of, like posting a media or status update and it says to me that it really make me really stop from sharing*."

Attitude towards institutional privacy

The affective part of attitudes towards institutional privacy was salient in both studies. Participants expressed feelings of how their privacy might be handled by Facebook or third parties as mainly negative using words such as afraid, bothered, not trusting, paranoiac, scary, skeptical, and weird. It made sense to have such feelings because the main channels participants have of knowing how their information is handled by institutions is through common news with adverse headlines, e.g. the Snowden revelations or personal awareness through experience. For example, one participant has worked for a Swedish telecom company with a policy that allows them to keep track of the proximity of their customers all the time, which was shocking to the other participants because they were not aware of this company's practice. As a consequence, most of the cognitive aspects of attitude in this context evolved around ambiguity concerning how user-shared information might be used despite initial awareness that it will be somehow used by the companies. It was interesting to see that participants referred to Facebook and its potential customers for secondary use as "they" all the time, without having a clear idea of who *they* actually are, which forms a nimbus of ambiguity around Facebook's institutional privacy.

"*You know like before you put out your name, your address and whatnot, you don't know the consequences … now it is kind of like it will come back to you in some way, somehow, magically or not. Yeah, I think lately, I've been more aware of it*."

Interestingly, after showing the tool, participants started to contemplate about the possible uses of their shared content to Facebook, thus changing their attitude as a result. Affectively, it was interesting to them to see that their photos and/or locations had the potential advantage to Facebook for mining relative concepts. Some even expressed feelings of being shocked and frightened as if this was something completely new to them. For example, it was interesting to see that even though some pictures were shared privately, they were still prone to secondary analysis. They also found this kind of analysis scary and quite invasive especially considering that this analysis could be done by others than Facebook, e.g. a potential employer gathering intelligence about people. The cognitive aspect then triggered more possible ways that such inferences could be utilized, such as being profiled, acquiring a clearer picture of the commercialization of content, and becoming familiar with what others could potentially find out about me. It was more understandable for the participants to see how the profiling mechanism works as the result of having a picture of possible value posted on their Facebook and how one person could possibly be categorized based on habits and lifestyle and some participants believed that the inferences could be unfair due to the fact that

one person might end up in a wrong profile. Targeted commercial advertisement is one example that is based on a person's activities in OSN, and false inferences could be annoying to the individual. For example, one participant saw a lot of snow-related inferences in connection to her photos because she lived in the north, therefore, she was being potentially profiled as a person who likes winter. However, she stated that she likes summer most, and if she was to receive ads, she would prefer them to be related to summer.

"When you share your photos you share maybe, one, two, or three photos at this point, and you kind of forget the ones you previously posted, and then when you see them all together, it gives you a kind of summary of the pictures that you are introducing of yourself, the profile that you are actually producing."

Table 2 summarizes general attitudes towards disclosures on Facebook and institutional privacy.

Table 2. Categorization matrix: Attitudes on disclosure and institutional privacy, before and after

Attitude towards disclosure	Before demonstration of DataBait	**Affective:** afraid, being tracked, cautious, convenient, unlike, embarrassing, fun, interesting, passive, scary, upset, hesitant, skeptical, dangerous
		Cognitive: communicate, keeping other's information private, keeping my personal information private, be selective in sharing, felt like being watched
	After demonstration of DataBait	**Affective:** awareness, cautious, scared
		Cognitive: alerting, avoid unwanted disclosure, avoid unwanted inferences, restrictive, note of other's privacy, educated
Attitude towards institutional privacy	Before demonstration DataBait	**Affective:** afraid, bother, deceiving, unlike, ignorant, uninformed, untrusted, paranoiac, scary, skeptical, trusted, weird
		Cognitive: commercial purposes, obscure use, they get to know me, makes me restricted
	After demonstration of DataBait	**Affective:** bother, curious to know more, interesting, scary, shocked, suspicious, need to be more cautious
		Cognitive: more aware, being profiled, commercialization, what others know about me, wrongly profiled, undesirable disclosure

Attitude changes when secondary use is practiced

By comparing the observed participants' attitudes between the two sets of group sessions (Table 2), it is apparent that the end users lean more towards affective rather than cognitive attitudes if they cannot directly see the possibilities for secondary uses of their information. Participants from study 1, after seeing the functionality of the DataBait tool, made affective responses to the possible secondary uses, like being scared and shocked. They became interested with the institutional privacy while they had doubt if secondary use on their own data was something worth to heed. They were rather certain about their disclosure habits and saw no deviation from first-hand disclosure because of their confidence in that nothing out of the ordinary could be inferred from what they were sharing. They preferred to govern their own self-awareness without a tool because they believed they could analyze better in their own mind. It could be concluded from Study 1 that participants were even resistant towards the idea of needing any automatic inference detection tool. However, Study 2 provided some contradictory insight. After using the DataBait tool on their own profiles, the participants were more cognitive with clearer ideas of how their information could be utilized for different purposes (e.g. commercial or government surveillance). Study 2 showed that users saw

more dangers in secondary use, and this insight appeared more menacing because they could see how their photos and location shares were actually prone to inferences beyond their initial sharing purposes. For example, one participant in Study 2 said:

"[Facebook] has all kinds of crazy programs and algorithms and whatever, that analyze all your habits and likes and what links you click on, it is how they tailor a lot of the advertising you see on the sites and everything, I think. Facebook probably does worse than the developers of this [DataBait] on a regular basis. Plus they are a massive multi-national company with the service base on the people but this [DataBait] is just a couple of developers in a couple of companies and universities in Europe."

It was evident that users have an incomplete view of the potential use of their shared information over OSNs. While being unconfident about institutional privacy, their disclosure practices had allowed them to share photos and locations that they thought could not be used for other purposes. When participants were confronted with the potential inferences (related to, e.g. religion or employment), they thought more about how an individual's disclosure can build up an image of a person when analyzed by a third party, which could be unwanted. Compared to Study 1, participants were unsure about the use of the DataBait and thought that their disclosures were within their control. In Study 2, however, the participants mentioned that the tool could empower them to see what an OSN (e.g. Facebook) could possibly find in their data. This led them to express interest in using the tool in the future. Several participants said that they will delete some of their pictures based on the analysis provided by the DataBait tool. Seeing the inferences generated by their own data triggered thoughts about how technological advances are drastic, and the participants then felt potential dangers more clearly.

"I mean, if I had this tool that I could go and see, I would definitely do it. First of all because I was recently looking for a job and I want to be very careful about what others tag me in and what I post myself, and I went through and looked, does it look like a profile that is ok or clean?, if someone else goes in and checks it then I would probably, if I knew that, I would delete photos and then this would all go away, and that they (Facebook) wouldn't save it and still keep a record of this, for example, then I would go and make changes so that I get things that only I think are ok to see."

Information disclosure is not only direct but also indirect

In terms of personal information sharing on Facebook, participants were asked what kind of personal information they think they share. Most of the participants were aware of the direct sharing of information, such as name, home address, place of residence, marital status, although the notion of personal information was debatable among the participants. Some participants thought of personal information as the information that is too private or secret to share in the first place, such as a bank account number or a home security code, while some involved also basic personal information in their own concept. Therefore, when asked about sharing personal information, they reflected in relation to their *direct* information sharing and things that, according to their own pre-thoughts, could be directly linked to them. For example, one participant in Study 1, in answer to the question of whether or not he was concerned about Internet privacy, he replied:

"No, I think things that you think are private, you should not put up on the internet, like I don´t want anyone to know my bank account number or my code to my alarm system. I don't put those up on the Internet really, because those are secrets for me."

In terms of sharing indirect information, a little awareness was detected among participants that had higher privacy literacy. However, after the tool presentation all participants could reflect upon their indirect information revelations and things that could potentially be traced back to them. For example, some could see the possibility of an aggregation of their photos to be linked to certain place, therefore, their locations would be possible to track in time. We saw that users may need to stimulate their sharing practices in contrast to what could potentially happen in order to make sense of how information they shared could be indirectly related to them. For example, by tracking the attitude of one participant it could be seen that she stated first that she does not use the Facebook app on her cell phone because she was afraid that her location would be revealed thorough GPS, and was, thus, aware directly sharing information. After seeing the result

of her inferences through the location leaks function, she now saw how she was, nevertheless, giving hints about her locations through her posts.

Discussion

The aim of this research was to investigate to what extent demonstrations of how to infer conclusions from uploaded personal Facebook content impacts user attitudes towards disclosure and institutional privacy. This research contributes to the privacy literature by illustrating the importance of personal awareness of institutional privacy and effects on attitude towards disclosure after exposure to automated inferences made from an individual's own personal information on Facebook.

As a first contribution, this research complements previous approaches, in which informants were exposed to imaginary or speculative scenarios, to study secondary use of personal information. For example, there can be many different interpretations of personal contents that end users might not be aware of at the time of answering the questions in surveys or case scenarios. The findings show that attitude changes become stronger when a user experiences inferences from their own personal data. Future research should not only rely on a user's preexisting knowledge of secondary use because users feel that secondary use of information is within their realm of expectations and, thus, in their control. The findings also show that individuals consciously select and disclose personal information in alignment with their self-regulatory preferences to avoid secret revelation (Boyd 2007) but lack the awareness that the data holder is capable of processing information to dig up a great many additional inferences. These findings are a means to describe why future employers, governments or corporations, do not have an impact on the visibility of user profiles (Tufekci 2008), because users have a more fragmented view of their own profiles than what third parties can potentially have. Users care about the information retrieved from their profile by third parties, however, their lack of awareness of actual information processing routines influences their disclosures. These findings are in line with previous research, which found that the higher the perceived Internet privacy risk, the lower the willingness to provide personal information over the Internet (Dinev and Hart 2006).

A second contribution of this research is that it demonstrates the significance of awareness of secondary use as a factor that influences disclosure intention. Users tend to disclose due to the fact that they are unable to see how advances in technology can upset the balance of public and private boundaries in genres of disclosure (Palen and Dourish 2003) related to OSNs. Users manage their disclosures depending on the level of privacy of the information and the publicness of the communication channel (Masur and Scharkow 2016). In the present case, when confronted with possible views of their own disclosed information, the participants saw that the inferences could become too private. On the other hand, participants distinguished between who could have made the inferences and how public an inference can go e.g. if it remains with Facebook for commercial purposes or if it extends to governments, secret services, or other third parties. Being aware that information could potentially be traced indirectly impacts the disclosure decision of the individual. Participants disclosed differently based on expectations of use of information, and since this secondary use is very vague, the disclosure decision is weakly tied to preferences and actual foreseen private and public boundaries. Users may disclose something assuming that there is nothing private in a photo or that it does not look private (Krasnova, et al. 2009) but, before being concretely exposed to potential analyses (such as with the DataBait tool) they can only imagine or guess, how private or public a photo will be. User cognitive ability lags behind technological advances. Consequently, it is argued that users need assistance in terms of determining if their disclosure is really aligned with their socially constructed and self-regulatory privacy practices. In the present study, individuals first thought that their disclosures were aligned with their preconceived policies, but after seeing that their content could be open to other interpretations through the DataBait tool, they started to defy their initial disclosure.

With respect to the notion of boundary of the self and others, inferences can be made from a person present in a photo or tagged in a post. Therefore, one person can have an impact on the privacy of others and vice versa. For example, being in a photo album with someone else, which contains a lot of alcohol inferences, puts the privacy of self and others into jeopardy. It is commonly known that Facebook is able to automatically recognize people in photos (Ingram 2015), and it could be seen here that the affordances of such technologies combined with those of data mining can lead to undesirable third-party discoveries about people. It should be noted that awareness of secondary use has an impact on the boundaries of 'self and other' as well as on 'public and private'. The findings herein, consequently, suggest that when users are

actively made more aware of potential secondary genres of disclosure (cf. Padyab 2014) that can be inferred from their direct disclosure genres (cf. Palen and Dourish 2003), their attitudes to disclosing will change. Moreover, it is asserted that better means and tools to empower users with such awareness will be needed, as common pre-conceptions do not seem to equip most users with enough critical knowledge in this regard.

Altogether, this research reinforces the established value orientation towards educating users about the risks to their privacy and finding ways to change user behavior (Debatin et al. 2009; Krishnamurthy et al. 2011). IS researchers need to focus more on the invisible part of the OSN iceberg and make users more aware of possible secondary uses of their information in a way in which they can test by themselves. The complexity of behavioral profiling and potential secondary use of personal information is hard to grasp by an average user, and a shift needs to be made towards more experimental computing (Yoo 2010) in this field of research.

Limitation and future work

This study has limitations. Although it was found that the perceived concepts found in the data analysis were saturated, future study should consider whether the results are transferable to other settings. For example, similar research could be conducted on other OSNs, while this research let users practice with their own Facebook profiles. Web trackers are also an emerging issue directed towards secondary use of data gathered from end users (Srivastava et al. 2000).

Focus groups are generally prone to group effects (Bryman 2012). Therefore, follow-up interviews would give a better understanding of participants' change of attitude due to the sensitivity of the issue, which might not have been possible to discuss in a group setting. For this reason, the aim of our future research is to capture actual disclosure behavior (Smith et al. 2011), and actual changes in behavior, in relation to greater awareness of institutional privacy.

Conclusion

This paper reported on a study of user attitudes towards institutional privacy after exposing users to potential inferences from their own personal information on Facebook. The empirical approach to enlighten users with their own information complemented previous studies that have been mostly based on capturing the pre-conceptions of users or exposing users to imaginary scenarios and cases. Those approaches do not take into account the full sensitivity of the information inferred by OSN companies and do not consider the fact that secondary uses of information might not be familiar to users. It was found that user attitudes towards institutional privacy and disclosure changed more when participants experienced the intrusiveness of the possible secondary information use on their own data. This was compared to a setting in which the potential of the DataBait tool was only discussed on a general level. Before introducing the tool to the participants, most attitudes were affective while, after introduction of the tool, attitudes shifted towards being cognitive. It was also observed that pre-awareness of indirect personal information disclosure was relatively low. Participants mainly felt that they disclosed only the same information that they shared directly and actively. Through observing some of the inferences that could be made indirectly from a user profile, the participants were able see a connection between the value of their information disclosed and its implications for their privacy. In the uneven battle between corporations to obtain knowledge of behavioral advertising algorithms for secondary purposes and end users, who are unaware of these practices, more education is required to empower users with more awareness and a balanced view of their institutional privacy. Practitioners can develop tools to assist users at the same time in evaluating their own information to be shared on OSNs and the Internet.

Acknowledgements

This work was funded by the European Commission in the context of the FP7 project USEMP (under Grant Agreement No. 611596), FP7 project IoT Lab (Grant Agreement No. 610477) and Horizon 2020 project PrivacyFlag (Grant Agreement No. 653426), which are gratefully acknowledged.

References

Abelson, R. P., Kinder, D. R., Peters, M. D., and Fiske, S. T. 1982. "Affective and semantic components in political person perception," *Journal of Personality and Social Psychology* (42:4), pp. 619–630 (doi: 10.1037/0022-3514.42.4.619).

Acquisti, A. 2004. "Privacy in Electronic Commerce and the Economics of Immediate Gratification," in *Proceedings of the 5th ACM Conference on Electronic Commerce*, New York, NY: ACM, pp. 21–29 (doi: 10.1145/988772.988777).

Acquisti, A., and Gross, R. 2006. "Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook," in *Privacy Enhancing Technologies*Lecture Notes in Computer Science, G. Danezis and P. Golle (eds.), Springer Berlin Heidelberg, pp. 36–58.

Adjei, J. K., and Olesen, H. 2012. "Secondary Uses of Personal Identity Information: Policies, Technologies and Regulatory Framework," *Communications & Strategies* (1:88), pp. 79–98.

Bateman, P. J., Pike, J. C., and Butler, B. S. 2011. "To disclose or not: publicness in social networking sites," *Information Technology & People* (24:1), pp. 78–100 (doi: 10.1108/09593841111109431).

Belanger, F. 2012. "Theorizing in information systems research using focus groups," *Australasian Journal of Information Systems* (17:2).

Bélanger, F., and Crossler, R. E. 2011. "Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems," *MIS Q.* (35:4), pp. 1017–1042.

Bouma, G. D., Atkinson, G. B. J., and Dixon, B. R. 1995. *A handbook of social science research*, Oxford: Oxford University Press.

Boyd, D. 2007. "Why Youth (Heart) Social Network Sites: The Role of Networked Publics in Teenage Social Life,"MacArthur Foundation Series on Digital Learning: Youth, Identity, and Digital Media Volume, D. Buckingham (ed.), Cambridge, MA: MIT Press, pp. 119–142 (available at http://www.danah.org/papers/WhyYouthHeart.pdf).

Boyd, D., and Hargittai, E. 2010. "Facebook privacy settings: Who cares?," *First Monday* (15:8) (doi: 10.5210/fm.v15i8.3086).

Brandtzæg, P. B., Lüders, M., and Skjetne, J. H. 2010. "Too many facebook 'Friends'? Content sharing and sociability versus the need for privacy in social network sites," *International Journal of Human-Computer Interaction* (26:11–12), pp. 1006–1030 (doi: 10.1080/10447318.2010.516719).

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

Butler, D. 2007. "Data sharing threatens privacy," *Nature News* (449:7163), pp. 644–645 (doi: 10.1038/449644a).

Choi, B. C. F., Jiang, Z., and Yap, E. 2012. "Information Sharing in Online Dyadic Exchange: A Relational Dialectic Perspective," in *2012 45th Hawaii International Conference on System Science (HICSS)*R. H. J. Sprague (ed.), Maui, Hawaii, pp. 743–752 (doi: 10.1109/HICSS.2012.324).

Cooper, P., Diamond, I., and High, S. 1993. "Choosing and Using Contraceptives - Integrating Qualitative and Quantitative Methods in Family-Planning," *Journal of the Market Research Society* (35:4), pp. 325–339.

Crano, W. D., and Prislin, R. 2006. "Attitudes and Persuasion," *Annual Review of Psychology* (57:1), pp. 345–374 (doi: 10.1146/annurev.psych.57.102904.190034).

Culnan, M. J. 1993. "'How Did They Get My Name?': An Exploratory Investigation of Consumer Attitudes toward Secondary Information Use," *MIS Quarterly* (17:3), pp. 341–363 (doi: 10.2307/249775).

Debatin, B., Lovejoy, J. P., Horn, A.-K., and Hughes, B. N. 2009. "Facebook and Online Privacy: Attitudes, Behaviors, and Unintended Consequences," *Journal of Computer-Mediated Communication* (15:1), pp. 83–108 (doi: 10.1111/j.1083-6101.2009.01494.x).

Dienlin, T., and Trepte, S. 2015. "Is the privacy paradox a relic of the past? An in-depth analysis of privacy attitudes and privacy behaviors," *European Journal of Social Psychology* (45:3), pp. 285–297 (doi: 10.1002/ejsp.2049).

Dinev, T., and Hart, P. 2006. "An Extended Privacy Calculus Model for E-Commerce Transactions," *Information Systems Research* (17:1), pp. 61–80 (doi: 10.1287/isre.1060.0080).

Donath, J., and Boyd, D. 2004. "Public Displays of Connection," *BT Technology Journal* (22:4), pp. 71–82 (doi: 10.1023/B:BTTJ.0000047585.06264.cc).

Edmunds, H. 2000. *Focus Group Research Handbook* (1 edition.), Lincolnwood, Ill.: McGraw-Hill.

E-Government Act of 2002. (n.d.). "Bill Text - 107th Congress (2001-2002) - THOMAS (Library of Congress)," (available at http://thomas.loc.gov/cgi-bin/query/F?c107:1:./temp/~c1079aXJzp:e72517:; retrieved February 6, 2016).

Elo, S., and Kyngäs, H. 2008. "The qualitative content analysis process," *Journal of Advanced Nursing* (62:1), pp. 107–115 (doi: 10.1111/j.1365-2648.2007.04569.x).

EU. 2011. "Directive 95/46/EC," *Official Journal L 281 , 23/11/1995 P. 0031 - 0050* (available at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML; retrieved February 6, 2016).

Fuchs, C. 2013. *Social Media: A Critical Introduction*, London: SAGE.

Ginsca, A. L., Popescu, A., Borgne, H. L., Ballas, N., Vo, P., and Kanellos, I. 2015. "Large-Scale Image Mining with Flickr Groups," in *MultiMedia Modeling*Lecture Notes in Computer Science, X. He, S. Luo, D. Tao, C. Xu, J. Yang, and M. A. Hasan (eds.), Switzerland: Springer International Publishing, pp. 318–334.

Gross, R., and Acquisti, A. 2005. "Information Revelation and Privacy in Online Social Networks," in *Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*, New York, NY: ACM, pp. 71–80 (doi: 10.1145/1102199.1102214).

Hann, I.-H., Hui, K.-L., Lee, S.-Y. T., and Png, I. P. L. 2007. "Overcoming Online Information Privacy Concerns: An Information-Processing Theory Approach," *Journal of Management Information Systems* (24:2), pp. 13–42 (doi: 10.2753/MIS0742-1222240202).

Hull, G., Lipford, H. R., and Latulipe, C. 2010. "Contextual gaps: privacy issues on Facebook," *Ethics and Information Technology* (13:4), pp. 289–302 (doi: 10.1007/s10676-010-9224-8).

Hum, N. J., Chamberlin, P. E., Hambright, B. L., Portwood, A. C., Schat, A. C., and Bevan, J. L. 2011. "A picture is worth a thousand words: A content analysis of Facebook profile photographs," *Computers in Human Behavior* (27:5), pp. 1828–1833 (doi: 10.1016/j.chb.2011.04.003).

Ingram, M. 2015. "Facebook's new algorithm can identify you even if your face is hidden - Fortune," June (available at http://fortune.com/2015/06/23/facebook-facial-recognition/; retrieved May 6, 2016).

Iyilade, J., Orji, R., and Vassileva, J. 2015. "Factors Influencing User's Attitude to Secondary Information Sharing and Usage," *CIT. Journal of Computing and Information Technology* (23:3), pp. 231–244.

Johnson, M., Egelman, S., and Bellovin, S. M. 2012. "Facebook and Privacy: It's Complicated," in *Proceedings of the Eighth Symposium on Usable Privacy and Security*SOUPS '12, Washington, DC: ACM, p. 9:1–9:15 (doi: 10.1145/2335356.2335369).

Kaspersky Lab. 2016. "Online Data Sharing Wrecks Marriages and Careers | Kaspersky Lab," April 12 (available at http://www.kaspersky.com/about/news/product/2016/Online-Data-Sharing-Wrecks-Marriages-and-Careers; retrieved May 4, 2016).

Kosinski, M., Stillwell, D., and Graepel, T. 2013. "Private traits and attributes are predictable from digital records of human behavior," *Proceedings of the National Academy of Sciences of the United States of America* (110:15), pp. 5802–5805 (doi: 10.1073/pnas.1218772110).

Krasnova, H., Günther, O., Spiekermann, S., and Koroleva, K. 2009. "Privacy concerns and identity in online social networks," *Identity in the Information Society* (2:1), pp. 39–63 (doi: 10.1007/s12394-009-0019-1).

Krasnova, H., Hildebrand, T., and Guenther, O. 2009. "Investigating the value of privacy in online social networks: conjoint analysis," in *Proceedings of the 30th International Conference on Information Systems (ICIS 2009)*, Phoenix, AZ, p. 173.

Krishnamurthy, B., Naryshkin, K., and Wills, C. 2011. "Privacy leakage vs. protection measures: the growing disconnect," in *Proceedings of the Web 2.0 Security and Privacy Workshop* (Vol. 2), Oakland, CA, pp. 1–10.

Lincoln, Y. S., and Guba, E. G. 1985. *Naturalistic Inquiry*, SAGE Publications.

Lipford, H. R., Besmer, A., and Watson, J. 2008. "Understanding Privacy Settings in Facebook with an Audience View," in *UPSEC'08 Proceedings of the 1st Conference on Usability, Psychology, and Security*E. Churchill and R. Dhamija (eds.) (Vol. 8), Berkeley, CA: USENIX Association, pp. 1–8.

Liu, K., and Terzi, E. 2010. "A Framework for Computing the Privacy Scores of Users in Online Social Networks," *ACM Trans. Knowl. Discov. Data* (5:1), p. 6:1–6:30 (doi: 10.1145/1870096.1870102).

Liu, Y., Gummadi, K. P., Krishnamurthy, B., and Mislove, A. 2011. "Analyzing Facebook Privacy Settings: User Expectations vs. Reality," in *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, New York, NY: ACM, pp. 61–70 (doi: 10.1145/2068816.2068823).

Masur, P. K., and Scharkow, M. 2016. "Disclosure Management on Social Network Sites: Individual Privacy Perceptions and User-Directed Privacy Strategies," *Social Media + Society* (2:1), p. 2056305116634368 (doi: 10.1177/2056305116634368).

Mayring, P. 2000. "Qualitative Content Analysis," *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* (1:2) (available at http://www.qualitative-research.net/index.php/fqs/article/view/1089).

Morgan, D. L. 1996. "Focus Groups," *Annual Review of Sociology* (22), pp. 129–152.

Narayanan, A., and Shmatikov, V. 2009. "De-anonymizing Social Networks," in *2009 30th IEEE Symposium on Security and Privacy*, Presented at the 2009 30th IEEE Symposium on Security and Privacy, Berkeley, CA: IEEE, May, pp. 173–187 (doi: 10.1109/SP.2009.22).

Nemec Zlatolas, L., Welzer, T., Heričko, M., and Hölbl, M. 2015. "Privacy antecedents for SNS self-disclosure: The case of Facebook," *Computers in Human Behavior* (45), pp. 158–167 (doi: 10.1016/j.chb.2014.12.012).

Orito, Y., Fukuta, Y., and Murata, K. 2014. "I Will Continue to Use This Nonetheless: Social Media Survive Users' Privacy Concerns," *International Journal of Virtual Worlds and Human Computer Interaction* (doi: 10.11159/vwhci.2014.010).

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).

Palen, L., and Dourish, P. 2003. "Unpacking privacy for a networked world," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, NY: ACM, pp. 129–136 (doi: 10.1145/642611.642635).

Popescu, A., Hildebrandt, M., Breuer, J., Claeys, L., Papadopoulos, S., Petkos, G., Michalareas, T., Lund, D., Heyman, R., van der Graaf, S., Gadeski, E., Le Borgne, H., deVries, K., Kastrinogiannis, T., Kousaridas, A., and Padyab, A. 2016. "Increasing Transparency and Privacy for Online Social Network Users – USEMP Value Model, Scoring Framework and Legal," in *Privacy Technologies and Policy*B. Berendt, T. Engel, D. Ikonomou, D. Le Métayer, and S. Schiffner (eds.) (Vol. 9484), Cham: Springer International Publishing, pp. 38–59.

Potter, W. J., and Levine-Donnerstein, D. 1999. "Rethinking validity and reliability in content analysis," *Journal of Applied Communication Research* (27:3), pp. 258–284 (doi: 10.1080/00909889909365539).

Raynes-Goldie, K. 2010. "Aliases, creeping, and wall cleaning: Understanding privacy in the age of Facebook," *First Monday* (15:1) (doi: 10.5210/fm.v15i1.2775).

Shenton, A. K. 2004. "Strategies for ensuring trustworthiness in qualitative research projects," *Education for Information* (22:2), pp. 63–75.

Smith, H. J., Dinev, T., and Xu, H. 2011. "Information Privacy Research: An Interdisciplinary Review," *MIS Q.* (35:4), pp. 989–1016.

Soczka, L., Brites, R., and Matos, P. 2015. "Personal Information Disclosure and Perceptions About Data Usage by Facebook," in *Proceedings of the Second European Conference on e-Learning*A. Mesquita and P. Peres (eds.), Reading, UK: Academic Conferences Limited, January, pp. 413–420.

Son, J.-Y., and Kim, S. S. 2008. "Internet Users' Information Privacy-Protective Responses: A Taxonomy and a Nomological Model," *MIS Quarterly* (32:3), pp. 503–529.

Srivastava, J., Cooley, R., Deshpande, M., and Tan, P.-N. 2000. "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data," *SIGKDD Explorations* (1:2), pp. 12–23 (doi: 10.1145/846183.846188).

Suh, J. J., and Hargittai, E. 2015. "Privacy Management on Facebook: Do Device Type and Location of Posting Matter?," *Social Media + Society* (1:2), p. 2056305115612783 (doi: 10.1177/2056305115612783).

Theodoridis, T., Papadopoulos, S., and Kompatsiaris, Y. 2015. "Assessing the Reliability of Facebook User Profiling," in *Proceedings of the 24th International Conference on World Wide Web*WWW '15 Companion, New York, NY: ACM, pp. 129–130 (doi: 10.1145/2740908.2742728).

Tong, A., Sainsbury, P., and Craig, J. 2007. "Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups," *International Journal for Quality in Health Care* (19:6), pp. 349–357 (doi: 10.1093/intqhc/mzm042).

Tufekci, Z. 2008. "Can You See Me Now? Audience and Disclosure Regulation in Online Social Network Sites," *Bulletin of Science, Technology & Society* (28:1), pp. 20–36 (doi: 10.1177/0270467607311484).

Tuunainen, V., Pitkänen, O., and Hovi, M. (n.d.). "Users' Awareness of Privacy on Online Social Networking Sites – Case Facebook," in *BLED 2009 Proceedings*, Presented at the 22nd Bled eConference eEnablement: Facilitating an Open, Effective and Representative eSociety, Bled, Slovenia.

Wilkinson, S. 1998. "Focus group methodology: a review," *International Journal of Social Research Methodology* (1:3), pp. 181–203 (doi: 10.1080/13645579.1998.10846874).

Yang, H. 2012. "Young American Consumers' Prior Negative Experience Of Online Disclosure, Online Privacy Concerns, and Privacy Protection Behavioral Intent," *Journal of Consumer Satisfaction, Dissatisfaction & Complaining Behavior* (25), pp. 179–202.

Yoo, Y. 2010. "Computing in Everyday Life: A Call for Research on Experiential Computing," *MIS Quarterly* (34:2), pp. 213–231.

Young, A. L., and Quan-Haase, A. 2013. "Privacy Protection Strategies on Facebook," *Information, Communication & Society* (16:4), pp. 479–500 (doi: 10.1080/1369118X.2013.777757).

Appendix

In structuring the Focus Group discussion, a semi-structured interview guide was developed to help stimulate the discussions in the group. Each study followed a similar protocol with minor changes since one study was based on the illustrations of the DataBait tool and another study was based on actual use, which required instructions to facilitate running the web application.

Study 1

- Practical information to the participants such as: express your thoughts irrespective of what others might think about them, there are no right or wrong answers, the session takes at least 90 minutes, discuss with each other and not with the moderator, try to talk one at a time and please speak English
- Please introduce yourself, tell us how and why are you part of this workshop, and your background
- A short introduction of the USEMP project, consortium, aims, and objectives
- Social Media Use
    - Describe your daily usage of social media
    - Why do you feel it is important to use social media?
    - What functions do you use most often and which ones seldom?
    - What sort of information or content do you share?
    - What sort of information do you mostly post on 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?
    - Do 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?
- An introduction to the DataBait tool through screen shots describing how it functions
- Introducing the location leaks module and starting the discussion by asking the following questions:
    - 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 with this function?
    - Could you think of a situation(s) in which this function becomes necessary?
- Illustrating the photo leaks module and starting the discussion by asking the following questions:
    - How often do you share your photos? Why? 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 module?
    - What benefits do you see with this function?
    - Could you think of a situation(s) in which this function becomes necessary?
    - What could improve your motivation for using this?
- 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 the most sense to use?
    - Let's consider your expectations when it comes to protecting your personal information. Could DataBait or any other ideal tool succeed in doing this? How?
    - Please tell us if there is an important aspect we have missed?

Study 2

- Practical information to the participants such as: express your thoughts irrespective of what others might think about them, there are no right or wrong answers, the session takes at least 90 minutes, discuss with each other and not with the moderator, try to talk one at a time and please speak English
- Please introduce yourself, tell us how and why are you part of this workshop and your background
- A short introduction of the USEMP project, consortium, aims, and objectives as well as one slide description of the DataBait tool

We asked participants to go to www.databait.eu and register themselves. The registration requires agreeing to the consortium's data license agreement, which specifies how Facebook data is handled by the consortium to make sure that the personal data is handled in a fair and transparent way, in accordance with EU law. The next

step in the registration was to choose a valid email, specify a password, and then to link their Facebook profile to their DataBait account. At this point, the registration was complete.

- Social media use and
    - Why do you use Facebook?
    - In which situations do you use Facebook?
    - What do you want to communicate when you use Facebook?
    - With whom?
    - When do you want to communicate through Facebook?
    - 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?
- Sharing photos on Facebook
    - Please describe how you handle your photos on Facebook. Give examples of when you use it.
    - Who do you communicate with when you share a photo?
    - What do you think when you decide not to upload photos? Describe in which situations, what is the reason for not using photo sharing? Why is this so?
    - Is there anything you communicate through photo upload or share that you would not like others to know about? What type of communication or information could that be?
- Sharing locations in Facebook
    - Please describe how you handle your locations on Facebook. Give examples of when you use it?
    - What do you think when you decide not to share a location? Describe in which situations, what is the reason for not using location sharing? Why is this so?
    - Is there anything you communicate through location sharing that you would not like others to know? What type of communication or information could that be?
- An introduction to the DataBait tool and a description of the photo and location inference modules
- Now click on "location leaks" and look at different 'cities' shown there. Click on a city and see the text (i.e. Facebook post) underneath. Play around on this page.

After 7-10 minutes of using the 'location leaks' feature, we continued with the discussion

- 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 relate or differ compared to your own expectations?
    - Please describe what sort of information that 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 taken action to limit this?
- Click on "image leaks" and look at different 'concepts' presented there. Click on a concept and see the photos underneath. Play around on this page.

After 7-10 minutes of using the 'image leaks' feature, we continued with the discussion

- Image leaks questions
    - What are your first impressions 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 taken action to limit this?
- General discussion
    - What sort of benefits do you see here to help control 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 the most sense to use?
    - Let's consider your expectations when it comes to protecting your personal information. Could DataBait or any other ideal tool succeed in doing this? How?
    - How do you imagine the future of this tool?
    - Please tell us if there is an important aspect we have missed?