TOWARDS USABLE END-USER PRIVACY CONTROL FOR SOCIAL SOFTWARE SYSTEMS

by

MARYAM NAJAFIAN RAZAVI

B.A.Sc, Sharif University of Technology, 1995
M.A.Sc. University of British Columbia, 2000

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

September 2009

© Maryam Najafian Razavi, 2009

Abstract

The recent growth and wide adoption of social software systems have transformed the Web from an information pool to a platform for communication and social interaction. While often times social software systems are used with the goal of sharing information, studies have shown that many users struggle to properly manage selective sharing of the vast and diverse information artifacts they dispose in such tools. Most existing social software systems define privacy either as a private/public dichotomy or in terms of a "network of friends" relationship, in which all "friends" are created equal and all relationships are reciprocal. These models fail to support the privacy expectations that non-technical users bring from their real-life experiences, such as segregating one's disparate groups, enabling different degrees of intimacy within one's network, and providing flexible, natural means of managing the volatile social relationships that social software systems confront. Furthermore, both models suffer from lack of empirical grounding and systematic evaluation.

The research described in this thesis employs a qualitative research methodology to deepen understanding of the information sharing process in the context of social software systems, in order to propose guidelines for building privacy management mechanisms in this domain that provide users with more control over privacy, and yet, are intuitive and easy to use for the average, non-technical user population of social software. The research is based on a grounded theory study of users' information sharing behavior in a social software tool, and offers several contributions, including clarifying users' privacy needs, concerns, and strategies, and identifying factors that affect users' decisions regarding sharing various information artifacts with different audiences. The findings lead to the development of several design heuristics and a general framework for usable privacy in social software domain, which inform the design of OpnTag, a novel prototype that facilitates creation, organization, and sharing of information for an individual operating in various social contexts. Results of an empirical evaluation of OpnTag's privacy management mechanism show that our proposed privacy framework is flexible enough to meet users' varying information sharing needs in different contexts while maintaining adequate support for usability.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACL	Access Control List
RBAC	Role Based Access Control
PIM	Personal Information Management
GIM	Group Information Management
SPIMS	Social-Personal Information Management System
HCI-SEC	Human Computer Interaction-Security
BSP	Basic Social Process
OLT	Office of Learning Technologies
PLE	Personal Learning Environment
Trans	Transition Program
BREB	Behavioral Research Ethical Review
VSB	Vancouver School Board
Ubicomp	Ubiquitous Computing
eCommerce	Electronic Commerce

Acknowledgment

Completing a doctorate has been a long-standing goal of mine, which could only have been achieved with the support of some generous people. As I come to the end of this journey, I would like to acknowledge those who provided me with technical and moral assistance throughout this process.

Dr. Lee Iverson was my supervisor for the most part of this work. Over the years, he mentored and challenged me, and provided a sounding board for my initial ideas. I am grateful to him for helping me embark on a fascinating research question despite my initial reservation coming from an engineering background, and for his work on the development of the initial version of the OpnTag software.

I owe a special debt of gratitude to Dr. Philippe Kruchten, who took responsibility for my thesis upon Lee's sudden departure. Although his involvement with this project was only during the last few months, his intellectual, financial, and administrative support had a tremendous effect on making my transition towards graduation a smooth one. He welcomed me into his lab, made sure I met my deadlines even when they were self-imposed, and provided me with a much-appreciated sense of piece and serenity that I needed to carry out my work to the end; and for that, I will be eternally grateful. I would also like to thank my committee members, Dr. Sid Fels and Dr. Rick Kopak, whose stimulating questions and comments greatly improved this dissertation, and my external examiner, Dr. Lorrie Faith Cranor, for her encouraging comments and feedback on my work.

Many people had a positive impact on my experience over the course of this Ph.D. I would like to especially thank my dear friend and labmate Vanesa Mirzaee, for being a great partner to bounce ideas off during many brainstorming discussions, and for always lending a listening ear to my rumblings when things got though. Kelle Fleming, the UBC Office of Technology coordinator, and Betty Gilgoff, the Transition Program coordinator, assisted with recruiting students in the Transition program as participants for the initial grounded theory study. Also, some of the ideas in this thesis have benefited from lively discussion with the CrowdTrust group. In particular, David Botta and David Vogt provided valuable insight

and feedback on the concept of people-tagging. And finally, I thank the anonymous reviewers of the various related publications, the participants of the workshops where I discussed this work in progress, and all the informants who kindly agreed to participate in my two studies.

I could not have done this without the loving support of my family. In particular, I would like to thank my husband Amir, for encouraging me to start this journey which in retrospect, turned out to be one of the most fulfilling experiences of my life, and for being a solid rock to lean on during all the ups and downs of the past few years (Ph.D. related and else). Also, my son Hirad, whose constant achievements always fill my heart with maternal pride, for being the most responsible kid a parent could hope for, and my daughter Heeva, our little sunshine, for being the infinite source of joy and fun that she is.

Last but not least, I would like to thank my father who to this date, has been my primary source of encouragement and inspiration. From an early age, he instilled in me self-esteem and an appreciation for hard work and dedication. Looking back, I can see now how the lessons he taught me have shaped my character and the paths I chose throughout my life. Today, I can only hope that this work, although far from perfect, will make him proud.

Chapter 1
Introduction

Social software systems are a family of Web 2.0 applications, characterized by their primarily user-driven content and the ability to mediate personal and social information across collectivities such as teams, communities, and organizations. The advent of the "Social Web" has made users producers as well as consumers of information, resulting in publishing and distributing huge amounts of user-created data. Examples of social software systems include social authoring tools (e.g., wikis), social bookmarking tools (e.g., del.icio.us), and social networking tools (e.g., LinkedIn, Facebook).

Social-personal information management systems (SPIMS) are social software systems for supporting personal information management in a social context. Personal information management refers to the act of creating, collecting, organizing, maintaining, or retrieving one's collection of personal information (Bergman et al. 2004). Such collection of information is called "personal" in the sense that the person who keeps the information has control over it. It does not mean that this information is necessarily about that person; nor that it is wholly created by that person, or that it is private (Lansdale 1988).

Conventionally, personal information management is considered a private activity. However, personal information is often created with sharing in mind or as a result of information sharing activities. This gives PIM a social dimension and therefore we consider sharing and exchange of information a part of PIM activity and not something separate (Erickson 2006). However, when people transfer their personal information from a private repository (e.g., one's desktop) into a social space (e.g., the Web) they are typically forced to give up control of some aspects of their information. One such aspect is that people are often limited as to how to define what information (and perhaps to whom) to reveal or conceal.

The recent adoption of social software applications for personal information management has created new opportunities for users. These applications not only allow their users to create personal information spaces that are easily accessible from anywhere on the Web, but also give them the tools to share their personal, social, and professional artifacts with others and take advantage of others' shared artifacts. Although use of social software systems for personal information management implies sharing with groups beyond users' control, people are often willing to do so for two main reasons: 1) social tools provide significant enhancements in utility and cost over similar desktop tools (e.g., Google Mail is free, intuitive, reliable and available anywhere); and 2) it becomes remarkably easy to share information and generate an audience when one chooses to put that information and knowledge online. These two advantages are in many cases so strong that users are either explicitly willing to give up control of that information or do so without any real awareness of the degree to which they are doing so. Nevertheless, providing users with a persistent sense of information ownership and control is important, especially in the case of SPIMS where the information being stored and potentially exchanged is creative, analytic, and/or work-related. This is because now in addition to the information itself, the way it is organized and its patterns of use and sharing also have value as personal knowledge.

Existing theories of knowledge sharing have compared the exchange of information between people with the exchange of money in economic systems. Fuller, for example, argues that having knowledge is to be able to solve problems, predict outcomes, and influence others (Fuller 2002). All of these have great economic potential and in our "knowledge economy" it is the content, organization, and control of one's knowledge that creates economic advantage for both the individuals and communities.

The serious ramifications of improper use of one's online data are well documented in the news media. People have lost their jobs when their employer discovered the employee's personal blog (Simonetti 2004). Bloggers have been stalked based on the opinions and personal information placed on their blog (Rowse 2006). Universities have disciplined students using photographs published on social-networking sites (Barnes 2006). Companies have regularly used information on social networking sites to screen and drop potential job candidates (Kumar 2008). These examples all emphasize that as the level of

social interaction supported by social software systems increases, there is also a pressing need to support end-users to properly describe and execute access control on the large, complex conglomerations of personal data they dispose on these applications.

1.1 Problem Statement

From a user's point of view, the primary concern in managing information sharing is the ability to define the particular audiences for each piece of personal artifact or attribute, generally in terms of a group of others who one trusts with that particular piece of information. Traditionally, group definition for access control has been based on organizational roles (Sandhu et. al 1996) or the equivalent (i.e., task (Thomas and Sandhu 1997)). However, the dynamic, administrator-less nature of social software makes these models inapplicable to this domain. These schemes rely on the existence of a central administrator to define and manage groups, assign access rights to each, and maintain access control lists on objects (e.g., using ACL (Access Control List) or RBAC (Role-Based Access Control) models). However, unlike organizational information management approaches for which access policies are predefined and rather stable, personal information policies are dynamic, highly individualized, and dependent to a large degree on current personal preferences, needs, and intended activities. Furthermore, no central administration exists in social software systems, and even if the users were willing to self-administer their data, the act of constantly specifying "who" should have "what kind" of access to "what objects" under "what conditions" would quickly become a burden. This is one reason such systems thus allow sharing either everything or nothing at all.

Alternatively, in social software systems access control is often defined either as a private/public dichotomy (i.e., a document can be either completely private or completely public), or the slightly more flexible "network of friends" model, in which users create one or more list(s) of friends and restrict content to be visible only to this/these group(s). The friends model has been widely adopted in social systems because it provides a relative balance in the trade-off between ease-of-use and flexibility (Hart et. al 2007). However, although "friendship" is more suitable than access control models for representing the interpersonal relationships reflected in social systems, this approach still fails to satisfy the expectation that non-technical users of social software systems bring from real-life

information sharing situations. One problem is that while in social systems, various groups that are relevant or important to an individual might simultaneously be present in one context, the friends' model does not allow users to segregate the disparate social groups with whom they share information. Users often have different degrees of intimacy with various people in their network and the friends notion is simply too coarse to capture these distinctions. A recent survey (Hart et. al 2007) shows that MySpace users had a median of 115 friends, suggesting that the friends notion was being stretched to cover a wide range of social acquaintances.

We thus suggest that neither the rule-based access control schemes (primarily intended for security administrators who are privacy experts) nor the friends model are directly applicable to the social software domain, and that there is a clear need for better models of privacy in this domain that allow non-technical end-users to easily and precisely express, apply, and update their information sharing preferences.

1.2 Towards Usable End-User Privacy Control for Social Software Systems

While many definitions of privacy exist in various domains, from an HCI perspective privacy is defined as "the capabilities of individuals in a particular situation to control what they consider to be personal data" (Cranor et. al 2006). Based on this definition, the goal of this research is to propose a simple, yet usable privacy management framework for social software systems, to support the average, non-technical user population of these tools with better control over selective sharing of their various online information artifacts. Research suggests that security and privacy mechanisms are only effective when used correctly, and that this is often not the case due to usability issues (Whitten and Tygar 1999). We thus believe that usability must be a central part of a privacy management model for social software, because if privacy aspects are so complex that end-users cannot understand or use them properly, then it is impossible to protect the privacy of an individual's personal information on the system. In order to be usable, privacy mechanisms should be intuitive, non-intrusive, and map well to users' mental model.

While privacy has been extensively studied in many domains (e.g., ubicomp, eCommerce, and media spaces), the topic of end-user privacy in social systems - how people manage privacy of their own information with respect to other individuals, as opposed to organizational privacy (Iachello and Hong 2007) - and usable tools to support it have received little attention in comparison. This can be partly attributed to the fact that social systems are a fairly new phenomenon and research in many areas is just beginning. Another factor is the inherent difficulty of usability research in the privacy domain. The unique challenges of the problem of usability engineering for privacy and security are well documented in the HCI literature (Ackerman and Mainwaring 2005, Karat et al. 2005, Cranor et. al 2006, Brunk 2006). To rephrase, security and privacy mechanisms are often not users' primary task. While users value these mechanisms, they consider the functionality secondary to completing their primary task. Also, privacy concerns are highly nuanced and contextual and as such, designs must accommodate various kinds of attitudes. The problem is further complicated by the fact that risks of non-usable privacy and security mechanisms are often higher than any other application. Also, it is often the case that most users are not privacy experts, which makes it difficult to capture their privacy needs and preferences, and to communicate their privacy options, current privacy state of their data, or future implications of their privacy decisions. Finally, use of privacy mechanisms often poses a trade off with functionality, performance, or convenience.

Despite these challenges, the importance of research into privacy and usability (and most importantly, how to make the two go hand-in-hand) cannot be underestimated. The HCI community has long been aware of the importance of usability in the success of any design. Recently, the privacy and security research community has also become increasingly aware of the impact of the usability problems on the effectiveness of privacy and security mechanisms for end-users. As an example, the 2003 Computing Research Association Conference on Grand Research Challenges in Information Security and Assurance identified the ability to "*Give end-users security controls they can understand and privacy they can control for the dynamic, pervasive computing environments of the future*" as one of the four grand research challenges in information security (CRA Report 2003). More recently, Iachello and Hong (2007) have proposed that "*developing more effective and efficient ways for end-users to manage their privacy*" and "*gaining a deeper understanding of people's attitudes and behaviors toward privacy*" are two

of the five grand challenges of privacy and HCI research; and finally, Karat et al. (2007) call the need for usable and trusted privacy in the PIM domain an emerging and critical need.

1.3 Research Objectives

There is recognition in the literature that appropriate attention to usable privacy issues will require more than just a focus on technology, and that there is need for a deep understanding of the range of user concerns and preferences about privacy in order to build usable and effective privacy interface mechanisms (Karat et al. 2007). As such, the objectives of this dissertation research fall under three general areas:

(1) Developing a better understanding of users' perspective on personal privacy in the social software domain, in terms of the extent of the problem, specific privacy needs and concerns, and strategies that users employ in order to achieve their desired level of personal privacy

(2) Identifying factors that impact users' information sharing behavior in this domain, in order to build a conceptual model of personal privacy that matches a user's mental model

(3) Devising guidelines for building privacy management mechanisms in this domain that satisfy users' varying privacy needs, and yet, are usable for the average, non-technical user population of social software systems.

Our approach towards meeting these objectives involved four steps, which are summarized below.

1.3.1 Phase One: Understanding Personal Privacy Needs and Concerns in Social Software Systems

Our first step was to gain a deeper understanding of users' perspective on personal privacy in the social software domain. Research into privacy preferences has generally concentrated on information disclosure to online retailers, sharing and privacy in the workplaces, and data collection and handling by businesses, government, and other formal organizations. However, preferences regarding selective sharing of personal artifacts in open online environments have been minimally studied. For this reason, we needed to conduct

foundational research to determine which aspects of information privacy manifest in our particular domain. While various social software systems for serving varied information sharing purposes exist, we specifically focused our research in the SPIM domain; since compared to other social software systems, privacy issues in SPIM are more important due to the fact that they mostly contain personal information in which users have vested interest. We used a qualitative research methodology to identify privacy needs, concerns, and challenges in SPIM domain from users' perspective. We conducted a grounded theory study (Glaser and Strauss 1967) with 12 participants who were using a social system with integrated blog, wiki, social bookmarking, and social networking functionality for over a year. The outcome of this phase was a grounded theory of information sharing behavior of users in SPIMS, which provided a deeper understanding of the problem space by portraying a clearer picture of users' perspective on the privacy of their information in a social context.

1.3.2 Phase Two: Developing A Privacy Framework

The second phase of the research involved translating the social requirements of a privacy system for SPIM (as identified by the grounded theory) into technical requirements of a system that can satisfy those requirements. For that, we proposed eight heuristics for designing privacy management mechanisms in SPIM based on the findings of the grounded theory, and consolidated those heuristics into a framework for privacy in this domain. The central structure upon which we based our design framework was the description of both the kinds of control of privacy that would be required in various dimensions, and how those controls must be presented to users to ensure usability. This included artifact control, audience control, relationship control, change control, and clarity.

1.3.3 Phase Three: Creating a Test Bed for Evaluation

The third phase of the research involved creating an environment for testing our proposed privacy framework in action, in order to illustrate its suitability for design. For this purpose, we designed and implemented OpnTag, an experimental web-based utility for social-personal information management, and built a privacy management mechanism based on the proposed framework for it.

1.3.4 Phase Four: Empirical Evaluation of The Framework

The last phase of the research involved evaluating our proposed privacy framework through controlled studies to see whether it fulfills the privacy requirements as expected. A laboratory study was conducted to validate the appropriateness and the overall usability and utility of OpnTag's privacy management mechanism in satisfying users' privacy needs across a variety of usage scenarios. The results of this study showed that the framework is flexible enough to meet users' varying privacy needs in terms of giving them more control over privacy, and that it provides reasonable usability, which suggested that it is possible to build usable privacy management systems that gives users more control over privacy based on the proposed framework. Findings of the study also yielded several potential improvements to the design of OpnTag's privacy management system.

1.4 Organizational Overview

Chapter 2 provides an overview of the literature relevant to the research presented in this thesis. The chapter begins with a summary of privacy definitions and theories. This is followed by an overview of the literature on group information management (GIM), and users' information sharing preferences both in the social software domain and other areas. We then discuss various frameworks that have previously been proposed for supporting privacy in different domains. The chapter ends with a discussion on usability engineering for privacy, including usability principles, design guidelines, and usability evaluation criteria.

Chapter 3 begins the presentation of the first phase of the research, by introducing the research methodology that was employed for our foundational study of information sharing behavior of users in an SPIM tool. Research questions and goals that motivated the study are discussed, followed by a comprehensive description of the grounded theory methodology and the data collection and analysis techniques that were used. We also discuss specific challenges of locating the study, along with strategies employed to address them.

The presentation of the phase one of the research continues in chapter 4, where the findings of the grounded theory study are discussed. The findings provides a set of propositions for better understanding of information sharing behavior of the users in a

social-personal information management system: the factors that shape users' perception of information privacy in such an environment, some of the challenges they face in ensuring privacy of information, and strategies they employ to achieve the desired level of privacy. Specific factors that affect users' information sharing behavior are also discussed.

Chapter 5 presents the second phase of the research. This chapter starts by an analysis of the validity of the grounded theory study, by providing evidence that it satisfies the four proposed criteria of fit, relevance, workability, and modifiability. We then proceed to the description of the design heuristics that we developed based on the results of the grounded theory and how they were consolidated into a framework for privacy in the SPIM domain. Some of the limitations of the grounded theory study are also discussed in this chapter.

Chapter 6 presents the third phase of the research. In order to show that it has practical benefit for design, our proposed framework needed to be systematically tested for validity under experimental conditions. For that, we implemented our proposed framework in an experimental SPIM tool and performed empirical evaluation through user studies. Chapter 6 explains the design and implementation of our test bed, OpnTag, an experimental SPIM system whose privacy management mechanism instantiates the proposed framework. The technical structure of OpnTag is discussed, along with a discussion of how our framework as embodied in this system supports each of the five user-oriented privacy controls.

Chapter 7 describes the last stage of our research; an evaluation of the framework. This chapter starts by clarifying the goals of the evaluation process and proceeds to a discussion of how we used OpnTag to evaluate our privacy framework in terms of both utility and usability. Through a preliminary empirical evaluation of OpnTag's privacy management model in a laboratory study with 10 participants, we provide initial validation of the viability of the proposed framework for building privacy management systems that provide users with more control over privacy without sacrificing usability. This chapter also discusses the challenges of the evaluation process, and some limitations associated with our laboratory evaluation methodology.

Chapter 8 offers a summary of the research presented in this thesis, followed by a description of the major research contributions. The chapter concludes with an overview of the planned future work.

Chapter 2

Related Work

This chapter summarizes the literature relevant to the research presented in this thesis. It covers several areas including privacy definitions, theories, and frameworks; information sharing behavior in various domains; personal and group information management; and usability design and evaluation guidelines for privacy mechanisms.

2.1 Privacy Definitions and Theories

Privacy is a multi-faceted phenomenon, in the sense that it carries different meanings in various domains. As such, we start this section by trying to disambiguate privacy, by discussing various definitions of privacy in the literature and existing theories that deal with the concept.

Iachello and Hong (2007) clarify the distinction between various philosophical perspectives on privacy and suggest that privacy researchers clearly specify their stance. One distinction that is related to this research is the one between *data protection* view (which refers to the management of personally identifiable information, typically by governments or commercial entities) and that of *personal privacy* (how people manage their privacy with respect to other individuals, as opposed to large organizations). In the data protection view, privacy policies and regulations are used to ensure end-users' control over personal information that is gathered about them by organizations. In contrast, in the personal privacy view the focus is on interpersonal relationships and social circles and how they affect information sharing.

Iachello and Hong suggest that recognizing this distinction is important, since it provides a framework for designers to select the most appropriate approach for solving a

specific privacy problem. They also argue that the personal privacy view seems to be a better model for situations where the data that is the subject of protection is not well defined as these kinds of situations tend to be difficult to model using rigid privacy policies that are typical of data protection guidelines. Clearly, the problem of end-user privacy in social software is best aligned with the personal privacy view, because of the highly contextual and situational nature of users' information sharing behavior.

From a sociological perspective, Phillips (2004) distinguishes between four kinds of privacy: freedom from intrusion (autonomy), construction of the public/private divide (social negotiation), separation of identities (individual's right to control, edit, manage, and delete information about themselves), and protection from surveillance (creation of social knowledge about individuals or groups). Although this segmentation is neither exhaustive nor mutually exclusive, it is helpful in clarifying the point of focus in addressing the problem of privacy in a certain context. End-user privacy in social systems, for example, is directly related to the issues of construction of the public/private divide and separation of identities.

Boyle and Greenberg (2005) provide a definition of privacy in the domain of media spaces. They define privacy as "a property of a piece of information that can be defined as a perception of how important it is to maintain control over access to it." They propose that privacy in this domain can be considered to have three dimensions: solitude (i.e., control over interpersonal interactions), confidentiality (i.e., control over access to one's personal information), and autonomy (i.e., control over one's own actions and expression of identity). Considering Iachello's classification of privacy perspectives, solitude and autonomy reflect the personal privacy view, while confidentiality is more related to the data protection view.

Alan Westin has been investigating consumer privacy rights since mid sixties (Westin 1967, 1991, 2003). His research primarily focuses on collection and use of consumers' information by commercial entities. Westin (2003) defines privacy as "the claim of an individual to determine what information about himself or herself should be known to others." Westin suggests that individuals are continually engaged in a balancing act to adjust their need for privacy with their desire for communication and disclosure (Westin 1967). He also suggests that many factors affect how the individual maintains this balance, including

social class, education, and psychological composition, and that individuals' need for privacy changes continuously based on contextual factors, such as different situations or events.

The Westin-Harris privacy segmentation model (Westin 1991) has been the dominative model for categorizing US consumers in terms of their privacy preferences and attitudes towards how their personal information is collected and used by companies. The model classifies consumers into three privacy categories: *privacy fundamentalists*, who feel strongly that current information practices are a threat to their privacy, and believe that current legislation is not sufficient to keep companies from using personal information irresponsibly; *privacy unconcerned*, who are not strongly concerned about the handling of their personal data and believe that sufficient safeguards are in place; and the *privacy pragmatists*, which constitute the majority, who are somehow concerned about privacy, but would weigh the benefits of compromising their privacy against potential risks. Although Westin classifications present a consistent categorization of US consumers based on their opinion on the use of personal information by commercial entities (the data protection view), there is no research to support that the same categorization could be applied to personal preferences in the personal privacy view. In fact, some recent research suggest that there is no correlation: Consolvo (2005), for example, found no correspondence between users' attitude towards disclosing their current location to people they know and their privacy classification, and Acquisti and Gross (2006) study showed that individuals' privacy concerns were only a weak predictor of their membership to Facebook, since many privacy concerned individuals join Facebook and reveal great amounts of personal information. As such, it has been suggested that Westin classifications should not be used to categorize participants of experimental studies that address the personal privacy issues, or to interpret the results of such studies (Iachello 2007).

Many researchers have discussed the situational nature of privacy. Goffman's works on social and interpersonal relations in small groups (Goffman's 1959) were one of the first to suggest that users project different personas in various social interactions. Goffman noted that release of personal information is a highly nuanced everyday activity, and that it is critically important to people to be able to present themselves differently to different audiences. The choice of what *persona* or *face* to present in a given situation is determined by

the combination of a number of factors, including the audience (i.e., a person might present him/herself as a loyal employee to a supervisor and a job seeker to another company), and social appropriateness (i.e., a teacher might present a formal face to his/her students and a more casual face to friends). While in real life it is easy for users to identify these factors, they are hard to capture or mathematically model in the online world.

Palen and Dourish (2003) build upon privacy regulation theory developed by social psychologist Irwin Altman (in turn influenced by Goffman's work), and extend it to the domain of information technology. Altman's theory (Altman 1975, 1977) suggests that privacy is a dialectic and dynamic social process. Palen and Dourish build upon the argument and suggest that end-user privacy is a continuous process of negotiating boundaries of disclosure, identity, and time, rather than a definitive entitlement. They observe that people may act *simultaneously* in different spaces: as individuals, as members of a family, members of some occupational group, etc. In each of these affiliations, they may choose to disclose different information to different audiences. Palen and Dourish then expose the unsuitability of existing access control models for end-user privacy management, since the conventional separation of one's network into "roles" (as done by existing access control models) fails to capture the fluid nature of these various *genres of disclosure* in which one acts. Palen and Dourish also point out that human and social factors may affect privacy as much as technology, and suggest that applications must be carefully designed to enable such dynamics. We follow on Palen and Dourish's lead and adopt the term *privacy management* to mean the user-centered expression of personal (and organizational) constraints on information sharing. This is distinguished from *access control*, which is the means by which systems enable and enforce these choices. The problem of usable end-user privacy in social systems is thereby defined as to how to provide users with intuitive and flexible means to specify and control what they want to share with whom in what context.

Ackerman (2000) has also discussed the nuanced and fluid nature of privacy, particularly in the domain of Computer Supported Collaborative Work (CSCW). He argues that CSCW applications present a huge information space for social interaction where it is critically important for people to be able to control the details of their interactions. However, because of the situational and complex nature of the privacy concerns for users, it is unlikely

that we can prescribe a "one-size-fits-all" approach that works adequately for all users. As a result, we are likely to be facing a situation he calls the *socio-technical gap*, where we know what we need to do socially, but we don't have the technical ability to do it. Ackerman sees the socio-technical gap as one of the major impediments in the design of effective and usable end-user privacy controls.

2.2 Information Sharing Preferences

This section covers previous studies of information sharing preferences in various domains. We start by studies of privacy in social software domain and then proceed to similar studies in other domains.

2.2.1 Privacy Issues In Social Software

In recent years, use of social software has moved from niche phenomenon to mass adoption. This increase in use has been accompanied by diversity of purposes and access patterns. As a result, researchers have studied several issues that pertain to these tools, including people's attitudes towards disclosing personal data. Most studies of privacy in social software focus on information sharing behavior of users in online social networking sites, and particularly on Facebook. This is partly due to Facebook's success in terms of user adoption and use, and most importantly, because of the amount and the quality of personally identifiable information that users make available on it, which make it a unique window of observation on the privacy attitudes of individuals.

One of the first academic studies of privacy in social systems is Gross and Acquisti's study on patterns of information revelation in online social networks and their privacy implications (Gross and Acquisti 2005). Their results are based on the analysis of actual field data from more than 4,000 Carnegie Mellon University Facebook profiles. They report that users' patterns of information revelation depend on a number of factors, including pretense of identifiability, type of information revealed or elicited, and the degree of information visibility. They also found that users rarely change the -often insufficient- default privacy settings of the tool, and that users' privacy expectations are often not met in reality.

In related work, Acquisti and Gross (2006) compared survey data from about 300 Facebook users to empirical data they mined from about 7000 users, looking for potential gaps between Facebook members' stated attitudes with actual behavior, and for the impact of privacy concerns on actual information sharing behavior of users. They found that there is often a disconnect between students' desire to protect privacy and their actual behaviors, evidenced by the fact that even privacy-concerned individuals joined the network and revealed great amounts of personal information. Acquisti and Gross argue that this is partly due to users' (sometimes wrong) confidence in controlling access to their online information, and partly due to a wrong conception regarding the size and the composition of the audience or the degree of accessibility of their information. Stutzman (2006) explores a similar idea, that users of social systems are just not sufficiently aware of potential privacy risks and threats to attempt proper protection. In a quantitative analysis of survey data from 38 Facebook users, Stutzman compares users' identity information disclosure on Facebook with their subjective opinions regarding identity protection and information disclosure. Stutzman's findings also suggest that there is a disconnect between the new types of identity information being disclosed through Facebook (photo, political views, sexual orientation) and users' stated opinions regarding identity information protection. Other researchers have made similar observations. (Barnes 2006), for example, suggests that we live in a "privacy paradox" where most people are not aware of the extent that social systems make their personal data available to take proper steps to stop further disclosure without authorization. These studies suggest that bringing awareness to users might change their behavior. However, the (Govani and Pashley 2007) study suggested that this may not be the case. Although participants in this study were mostly aware of possible consequences of revealing personally identifiable information to strangers (e.g., identity theft or stalking), and also knew how to easily limit over-exposure of their profile data, they did not take the initiative to protect their information.

Privacy is not just about availability of personal data. A recent study by Boyd (2006) asserted that Facebook's introduction of the "News Feed" feature disrupted users' sense of control, even though data exposed through the feed was previously accessible. This emphasizes the importance of users' ability to control impressions and manage social contexts in their sense of privacy. Preibusch, et al. (2007) notes that in social system, users'

privacy could also be affected by their social networks' privacy management, and that current privacy mechanisms in social software do not provide functionality to handle conflict with members of one's network who have different conceptions of privacy. Preibusch et al. suggest refining the public/private dichotomy with a tiered confidentiality level that considers group and community data in addition to public and private. They also propose a framework for analyzing privacy requirements and for analyzing privacy-related data in social networking systems.

2.2.2 Information Sharing In Other Domains

Researchers have also studied users' attitude towards revealing information in several other domains, including e-commerce, location-aware mobile services, CSCW, incidental information, and ubiquitous computing. Works in this category often focus mainly on identifying clusters of users with similar patterns of information sharing behavior, as these can then be translated into "templates" for building privacy management tools. While privacy issues in these domains are not exactly the same as privacy concerns in social software, this body of research informed our overall understanding of the appropriate methodologies and approaches for gathering and analyzing user data, and for translating users' needs into technical system requirements.

(Ackerman et al. 1999) surveyed privacy concerns of 381 Internet users regarding e-commerce and other Web-based transactions, mainly looking for people's reactions in situations where their personal data is collected by online entities and the reason behind their sensitivity to particular privacy practices. Their study identified different levels of sensitivity about collection and use of personal data by commercial sites, ranging from little concern about providing such information as favorite television show to great concern over credit card and medical information. The authors found critical differences in privacy perception and practices among privacy fundamentalists, pragmatists, and marginally concerned clusters, and noted that it will probably be easier to design privacy interfaces for the fundamentalists and marginally concerned groups, whereas the pragmatic group (the majority) will require more sophisticated and varied interface mechanisms to accommodate their highly nuanced and varied privacy strategies. They thus suggest that an individualized approach to privacy management would be necessary given the large variance in users' reactions.

Olson et al. (2005) conducted a two-phase survey to probe people's comfort level in sharing a range of everyday information with various others. Although the main finding of the study was the mere confirmation of the fact that individuals would share more sensitive information with closer acquaintances, Olsen et. al were able to also identify clusters of information types (i.e., availability to communication, contact information, web sites visited, or health status) and recipients (i.e., spouse, family, friends, close colleagues, remote colleagues, and public) that remain fairly consistent across different privacy attitudes and concerns. They also point out that whether data is anonymized or can be tied directly to people identity plays a major role in users' willingness to disclose. Personal judgment regarding "appropriateness" (i.e., relevance) of sharing certain information with certain groups was also found to be a relevant factor.

Patil and Lai (2005) conducted a study on privacy-awareness tradeoff to identify the kinds of information that users of an awareness application are willing to share with various others (team mates, family, friends, managers, etc.) for various purposes in the context of workplace. Their study also identified clusters of awareness information that are more likely to be shared with certain clusters of recipients in various contexts (i.e. "team members" received comparable level of awareness sharing with "family" during work hours). In another work, Patil and Kobsa (2005) studied privacy issues related to the use of Instant Messenger. They found their participants to have a wide range of privacy concerns with respect to the use of Instant Messenger, which suggested a more fine-grained approach might be required for managing privacy levels for contact lists. Whalen and Gates (2005) report on a small-scale study on the type of personal information that users would be willing to disclose in open online environments, primarily focusing on uncontrolled spaces such as search engines. Their results, although limited in scope, point to the existence of consistencies in the way people treat certain classes of information, which suggests it might be possible to group related information into clusters that are treated similarly.

Consolvo et al. (2005) conducted a three-phased study with 16 non-technical users (including both in-lab and in-situ) to explore users' behavior regarding disclosing their location information to social relations in different situations. They found the most

important factors to be the inquirer (*who* was requesting), the reason behind requesting the information (*why* the requester wanted the participant's location), and the level of detail that was requested (*what* information would be most useful to the requester). They suggested that these three factors were sufficient for the users to make an informed decision regarding how much they wanted to reveal about their location to the person who was requesting it. One interesting outcome of this study was that participants' privacy classification (as determined by the Westin-Harris Privacy Segmentation Model) proved not to be a good predictor of users' attitude towards location disclosure.

Other studies have focused on the role of specific factor(s) on information sharing preferences of users. (Lederer et al. 2003), for example, examine the impact of inquirer (i.e., spouse, employer, stranger, merchant), and the situation on the preferred accuracy of personal information disclosed in a ubiquitous computing system. They find the inquirer to be a more determining factor than the situation; i.e., users were more likely to disclose the same information to the same inquirer in different situations than to disclose the same information to different inquirers in the same situation.

(Hawkey and Inkpen 2007) examine incidental information privacy in Web browsers. Through the use of scenario-based surveys, they identify privacy concerns and privacy management approaches of users while browsing or searching for information on the Web. Their studies show that a majority of users are in fact concerned about privacy of their incidental information that may be revealed through browser history and search functionality. Like other studies, Hawkey and Inkpen found that the comfort level of users is impacted by the sensitivity of the information, the viewer, and his/her relationship with the user. However, Hawkey and Inkpen also found an additional factor that is specific to the domain of incidental information privacy, which is the amount of control that the viewer has on the input devices (such as mouse or keyboard).

2.3 Privacy Frameworks

In chapter 5, we present a framework for designing for privacy in social systems. Scientists have also developed privacy frameworks in various other domains. From this perspective, our work is mostly comparable with Belloti and Sellen (1993) privacy framework

for ubiquitous computing, and Adams' privacy framework in multimedia environments (Adams 1999a, Adams 1999b, Adams and Sasse 2001). Here we discuss each framework and how ours differs.

In a series of related works, Adams and Sasse summarize the details and results of four qualitative empirical studies to model users' perception of privacy in multimedia environments (Adams 1999a, Adams 1999b, Adams and Sasse 2001). From their results, they derive a model of privacy factors for multimedia environments from users' perspective, plus a theory of the process behind privacy invasion that details how ignoring these factors could lead to privacy invasion. In their model, Adams and Sasse break the privacy factors into three main categories: users' judgment of information sensitivity, their trust in the information receiver, and the costs and benefits of the usage of information; meaning, the model tries to explain how the designers must provide users with control over sensitive information about them being captured and delivered to the wrong audience and used for non-appropriate purposes. The model also explains the relationship between the factors and how they are affected by the context within which the information is captured, such as the technology in use, social groupings, and national/international settings.

A very similar decomposition of privacy is suggested by (Lederer et al. 2003) for the ubiquitous computing environments. Lederer et al. propose three dimensions for privacy that are of high impact in the domain of ubicomp and discuss how they relate to each other. The dimensions include *system properties* (the HOW factor), which include the important aspects of the mechanisms of disclosure and how participatory the disclosure is (i.e., surveillance vs. transaction); *actor relations* (the WHO factor), which include the history and nature of the relationship between information owner and information recipient (i.e., interpersonal vs. institutional); and *information types* (the WHAT factor), which represents critically distinct types of disclosable information (i.e., intentional vs. incidental). Lederer et. al suggest that this deconstruction of the privacy space can help focus the scope of discourse for the analysis or design of privacy-related systems, since privacy issues related to each of the dimensions of the model will be different and typical design solutions will vary. They don't, however, discuss appropriate design solutions for privacy problems that are mapped

to various dimensions of the model, and as such, their model is better viewed as a descriptive one rather than a design framework.

Bellotti and Sellen (1993) propose a privacy framework for ubiquitous computing environments. They argue that the main problem with ensuring privacy in ubiquitous computing is that the technology weakens the natural mechanisms of feedback and control over what information is captured and released about users. As such, their framework for privacy in ubiquitous computing is primarily based on the appropriate types of *feedback* and *control* that must be conveyed to users regarding information capture, construction, accessibility, and usage. In other words, Bellotti and Sellen's framework suggests that to provide users with appropriate level of privacy, ubiquitous computing systems must support users to understand and control when and what information about them gets into the system, what happens to those information once it gets inside the system, who (people and/or software) will have access to that information, and why the information is needed.

While both frameworks emphasize the importance of privacy of personal data, there are fundamental differences between the environments they are targeting (multimedia, ubiquitous computing) and the social software environment that limit the use of their frameworks in the context of our work: both frameworks are based on the same assumption that in these particular environments, users can't easily know what information is being captured and released about them. The main issue with ensuring privacy in social software however, is not the problem of hidden data. Rather, it is how to provide users with a flexible, non-overwhelming way to manage the vast amount of information they dispose in the tool in various situations, accounting for issues such reconsideration and moving from one context to another (e.g. from personal to professional). In other words, the main problem of privacy in social software is providing a user-centered mechanism for controlling user-created content.

2.4 Group Information Sharing

Recent works in the area of personal and group information management (PIM/GIM) have also recognized the need to improve people's ability to control who sees

what in their personal information space. In this section, we examine the interdependency between PIM/GIM, usability, and personal privacy.

Erickson (2006) explores the concept of personal information management in group context, by arguing that when personal information is to be shared with a group, the way it is used and managed changes. He uses the phrase GIM, Group Information Management, to refer to how personal information is shared within a networked group, the norms of personal information sharing within groups, and the way those norms are negotiated in the group.

Erickson observes that there are many issues involved with the simple scenario of creating information and sharing it with a group, including how to define and negotiate norms of sharing according to various factors, the definition of imaginary audience and how that affects sharing, and the level of control that is required over the information after it is shared. He suggests that these issues arise from a model of information sharing that is far too simple in the sense that it is based on the assumption that factors are static, whereas in reality, they are not. Erickson thus calls for more sophisticated models that take these issues into account.

Lutters et al. (2007) define GIM as "the practice and the study of the individual actions performed to support group activity." Their article on GIM explores the issues surrounding both leveraging others' information and sharing one's own data with a group. Lutters et al. emphasize that people control their information in flexible and nuanced manners, and that within a group, people's goals and understandings regarding information sharing may vary. They suggest that GIM tools must account for this varying and sometimes conflicting incentives and motivations. They also emphasize that preserving individuals' control over disclosure of their information to other group members (including whether it is shared at all) is an important factor in the adoption and use of GIM tools, and that GIM tools must support this in a seamless and non-invasive manner.

Karat et al. (2007) discuss the importance of usability in the design of tools for management of personal information disclosure. They argue that sharing personal

information for social, education, or professional purposes is an essential necessity in today's world; however, for users to be able to make informed decisions regarding what they want to share in various groups and properly manage selective disclosure of their personal information, the tools must provide usable mechanisms for managing privacy in which users can trust. Trust in the system only happens if users understand system's actions and feel in control. That includes understanding default privacy settings and how to modify them. Karat et al. thus propose that for designing systems with usable privacy management mechanisms, researchers must start from understanding potential users of the system, their tasks and goals, and the context in which they will be interacting with the system.

2.5 Privacy and Usability

In chapter 7, we attempt to validate our proposed framework for end-user privacy in social systems in terms of usability as well as utility. In this section, we provide a review of existing literature on usability for privacy and security mechanism (the HCI-SEC literature). We begin with an overview of suggested guidelines and principles for designing usable privacy and security management mechanisms/tools, and proceed to the works that suggest evaluation criteria for usability of such mechanisms.

2.5.1 Guidelines and Principles for Designing Usable Privacy

While today there is enough awareness that usability must be considered from the early stages of privacy and security development cycle, this has not always been the case. The first mention of usability and security having to work hand in hand is generally accepted to be in (Saltzer and Schroder 1975)'s paper "*The protection of Information in Computer systems*". The authors proposed eight principles to guide the design of security products, the last of which was "*psychological acceptability*". It described the interface design to be of essential importance, to enable users to routinely use the security mechanisms in the correct way.

Whitten and Tygar (1998) performed one of the first and most well known experiments to test the usability of security software. Through a cognitive walkthrough analysis, they evaluated the Pretty Good Privacy (PGP 5.0) encryption software against a set of four usability guidelines. They discovered a number of security risks and usability issues,

which they seek to confirm through a 12 participant user study. This case study demonstrated how a user interface that appeared good by traditional standards, failed to make public-key based electronic mail security manageable for experienced e-mail users due to usability problems. Whitten and Tygar thus suggest that computer security differs from other kinds of consumer software as a problem domain for usability engineering, in that unlike ordinary software, the usability of security software is not entirely based on the interface design. This work is remarkable for both providing a working definition of usability for security, and deriving a set of design principles for building and criteria for evaluating usable security software (to be discussed later in this chapter). It also uncovers how the poor matching between users' need and the technology that is provided to meet those needs can result in failure.

Other researchers have identified various design requirements for privacy management systems in other domains. Lau et al. (1999) observe that despite unanimous recognition for the importance of interfaces for privacy management, they are often inadequate in the sense that users with particular privacy needs and policy often lack the means to articulate and apply them in the system. They suggest that privacy interfaces should make it easy to create, inspect, modify, and monitor privacy policies and that the policies should be applied proactively to objects as they are encountered.

De Paula et al. (2005) discuss three design principles for enhancing the usability of systems with a security and privacy component: visualization mechanisms, event-based architecture, and integration of configuration and action. These principles are intended to help users better understand the consequences of their actions and make informed decisions regarding privacy and security. Through design and development of two technical infrastructures that make the configurations, activities, and implications of available security mechanisms visible, De Paula et al. attempt to create conditions whereby users not only recognize issues as they arise, but also understand them well enough to make informed decisions and take appropriate actions. An important aspect of this work is debating the traditional notion that transparency of privacy and security mechanisms leads to improved usability (because of hiding technical complexity from users). De Paula et al. argue that transparency could actually lead to users being unaware of the privacy and security

implications of their actions, while providing more visibility can support informed decision making.

Lederer et al. (2004) suggest that users' ability to manage and maintain personal privacy is determined by their ability to *understand* the privacy implications in a specific social situation, and to take appropriate *action* to respond to them. As such, system designers need to incorporate these two human-level processes through the limited technical mechanisms of feedback and control. Lederer et al. introduce five pitfalls that must be avoided when designing an interactive system with a personal privacy component. These pitfalls include:

1. Obscuring potential information flow: users need to know and understand what information is to be revealed,
2. Obscuring actual information flow: users need to know and be able to control who gets to see what,
3. Emphasizing configuration over action: users should not have to perform excessive configuration *a priori* to create and maintain privacy, since configurations are de-situated from the contexts to which they apply. Rather, they should be able to manage privacy within their normal interaction with the system,
4. Lacking coarse-grained control: management of disclosure and privacy has to be as nearly effortless as possible and users' normal interaction with the system should not be hampered by the actions they must take to preserve privacy, and
5. Inhibiting established practice: users' normal mechanisms of preserving privacy (e.g. taking advantage of plausible deniability) should not be weakened by the technology; users should not be expected to deviate from normal social practices just because the current technology works differently.

Hawkey and Inkpen (2007) propose five design guidelines for the management of visual privacy within web browsers (i.e., traces of prior browsing or search activity that can be made visible through the browser's convenience features such as AutoComplete). The guidelines are based on the results of three studies of users' web browsing behavior (one survey and two field studies), and include clutter reduction (providing control over what traces are stored, while not interfering with the browser's revisitation functionality), enabling nuanced privacy classifications (supporting more options for classifications of browsing

activities, rather than only public/private), supporting multi-tasking (support concurrent windows containing content of varying privacy sensitivity), supporting users' varying privacy concerns (providing options for personalization, rather than a generic, unified approach), and reducing the burden of privacy classification (i.e., semi-automate privacy classification by leveraging browser-window based patterns). Among these, guidelines two and four (support for nuanced privacy classification and varying privacy concerns) confirm findings from other studies of privacy in other domains. The other three guidelines are specific to the context of incidental information privacy in browsers and are not applicable to our domain of study.

2.5.2 Usability Evaluation Criteria

In order to determine whether our proposed privacy framework is viable, clear success criteria must be defined. Different metrics have been suggested in the literature to determine usability of the security and privacy mechanisms.

Whitten and Tygar (1998) define the four high level usability objectives of a security system as: to reliably make the user aware of the security task that needs to be done, to rely on user's expected knowledge and training for figuring out how to perform the task, to prevent the user from making dangerous errors, and to provide users with a user interface they feel comfortable enough with to continue using it. Based on a critical user study of two password managers, Chiasson et al. (2006) expand on Whitten and Tygar's work by suggesting two additional criteria: that the users must be able to tell when their task has been completed, and that they should have sufficient feedback to accurately determine the current state of the system.

(Yee 2003) proposes ten guidelines for the design and evaluation of user interfaces for security systems, including *expressiveness* (enabling users to express safe security policies in terms of their current task), *clarity* (conveying the results of an action to the user before it takes effect), *visibility/self-awareness* (maintaining accurate awareness of users' own/others' access rights), *foresight* (clearly indicating the consequences of decisions that the user is expected to make), and *revocability* (flexibility and ease in dynamically changing access restrictions). Also, maintaining agreement with users' mental model has been emphasized as an important factor in achieving usability for security.

(Bellotti and Sellen) also propose eleven criteria to be used as a basis for systematic evaluation of alternative design solutions for privacy. Their criteria mainly include suggestions on the characteristics of the mechanisms of feedback and control, plus some general criteria for the system as a whole. For feedback, Belloti and Sellen propose that quality feedback must be provided at the right time, be noticeable and non-intrusive, and should not distract or annoy user. For control, Bellotti and Sellen argue that mechanisms should be flexible, personalizable, lightweight, meaningful, and learnable. And finally, they suggest that any system with a privacy component must be technically reliable and instill confidence in users, and should minimize information capture, construction and access by default. It is also desirable that privacy solutions be lightweight to use and don't incur high costs.

2.6 Summary

The research presented in this chapter provides an overview of the related work in various areas related to end-user privacy, including privacy definitions and theories (section 2.1), factors influencing information sharing behavior of users in various domains (section 2.2), existing privacy frameworks (section 2.3), privacy issues surrounding sharing of personal information in a group context (section 2.4), and the interconnection of privacy and usability (2.5). This body of research helped establish the background to situate the stance of this work (end-user privacy) among the various dimensions of the privacy problem, and identified a gap in the literature that motivated the research presented in this thesis.

Investigating privacy research in various domains showed that identifying users' attitude and concerns is the first step towards building usable privacy management tools in any domain. However, this literature review also showed that while there are commonalities, privacy risks and users' privacy needs and issues are different in each domains (i.e., while identity theft and stalking are the prominent privacy threats in social networking systems, personal preferences and information relevance are the important factors to consider when designing for privacy in SPIM), and that user cluster models developed in one domain are not directly applicable to other domains (i.e., Westin-Harris privacy segmentation model for

e-commerce are not a good predictor of users' behavior regarding their personal privacy). Further examination of the existing privacy literature revealed a gap since none of these works address privacy issues surrounding selective disclosure of user-created content (on which users have vested interest) in SPIM. We also learned a lot about limitations of existing methodologies for studying user behavior that motivated some of the methodological choices in this dissertation.

Studying privacy issues in the context of Group Information Management gave us some perspective as to the issues that must be considered when managing privacy in SPIM, including the changing nature of users' motivations and incentives which calls for flexible privacy management mechanisms, and the importance of usability and understanding users, their information sharing tasks and goals, and the context of information interaction.

Learning about existing design guidelines, frameworks, and evaluation criteria for privacy tools in various domains shaped our holistic approach to the problem of usable end-user privacy in SPIM and taught us how to move from identifying users' needs, to system requirements, design guidelines, and developing frameworks, and how to choose appropriate criteria for evaluating our results.

The next chapter presents our grounded theory study aiming at identifying users' information sharing behavior in a social-personal information management system. Findings of this study clarified users' perspective on the privacy of the information they dispose in a SPIM application, and offered some ideas about how to create privacy management mechanisms for SPIM that are based on users' mental model of information privacy.

Chapter 3
A Study of Information Sharing Behavior in SPIMS

The literature review in chapter 2 indicated there is a lack of research examining end-users' specific privacy needs in the context of social software systems. As a first step in this research, we performed a grounded theory study to understand end-users' information sharing behavior in a social-personal information management tool and identify specific privacy needs and concerns in this domain. This chapter presents the research questions and goals that motivated the study, followed by a description of the grounded theory methodology and the data collection and analysis techniques that were used. Specific challenges of locating the study are also discussed, along with strategies employed to address them.

3.1 Research Questions and Goals

In order for privacy management mechanisms in SPIMS to be usable and effective, those mechanisms must be based on mental models of information privacy that reflect users' experience. As such, the main goal of this study was to gain an understanding of users' perception of personal information privacy in SPIMS. To that aim, we designed this study to investigate several questions with regard to information sharing preferences in an environment that allows sharing of personal and social artifacts of different degrees of sensitivity with various audiences. We were primarily interested in exploring the following aspects:

RQ1. Is privacy management in SPIMS considered important, and why?

RQ2. What factors affect users' decision regarding sharing a particular artifact with certain audiences?

RQ3. Are there any commonalities in the way users arrange their information with regard to sharing?

RQ4. What are the limitations of current systems in terms of privacy management, and how could those shortcomings be rectified?

By trying to find answers to these questions, we were aiming to identify fundamental concerns with privacy from users' point of view. Our main goal was to understand how users abstract the details of sharing into high-level classes of information and recipients that they treat similarly, and to incorporate those abstractions in a conceptual model of information sharing behavior in SPIMS.

3.2 Research Methodology

The research methodology that was employed in this study was grounded theory (Glaser and Strauss 1967, Glaser 1978, 1992, 1998); a primarily inductive investigation process in which the researcher aims for formulating a small-scale, focused theory that is derived from the continuous interplay between data collection and analysis. Grounded theory allows the researcher to develop a theoretical account of general features of a topic while simultaneously grounding the account in empirical observation or data (Glaser and Strauss 1967). More succinctly, it is the "discovery of theory from data". The grounded theory method has been suggested in the literature as the appropriate method for discovering behavioral patterns that shape social processes as people interact together in groups (Morse and Richards 2002). The intent is to develop an account of a phenomenon that identifies the major categories, their relationships, and the context and process, thus providing a theory of the phenomenon that is much more than a descriptive account. The goal of this process is to understand the action in a substantive area from the point of view of the actors involved (Glaser 1998, p. 115). The purpose of grounded theory method is building theory, not testing theory. The method is called "grounded" because a theory systematically is obtained from a broad array of data through a rigorous process of constant comparison. Rather than starting with a preconceived theory and systematically seeking out evidence to verify or prove it, the researcher begins with a general area of study and then systematically develops a mid-level substantive theory derived directly from the data (Glaser

and Strauss 1967, p. 4). The results of this research method are propositions and conceptual hypotheses, not facts or findings. Theory concepts are suggested, not proven.

The process of grounded theory development generally involves simultaneous collection, coding, and analysis of data, adopting a framework that is systematic, emergent, non-linear, and without researcher's pre-conceptions. More specifically, the process often involves the following steps:

- Gathering and *coding* data, which is the process of sorting data in preparation for analysis through labeling (assigning code words to phrases), linking (both data to ideas and ideas to all data pertaining to it), and fracturing (breaking data records to identify categories)
- *Theming*, which is the process of identifying a pervasive thread (e.g., patterns or common sequences) that runs through data; by sorting and analyzing data and recording reflections and insights in memos
- Gradually elaborating a small set of generalizations that cover the consistencies, which leads to the emergence of the *basic social processes* (BSPs). BSPs are the core concepts around which the grounded theory is built, and,
- Confronting these generalizations with a formalized body of knowledge, which leads to the *construction of the theory*.

3.2.1 Rationale Behind the Research Methodology

The nature of the research question is often the determining factor in deciding which methodology would be most suitable to answer that question. From the start of this project, it was quite clear that the nature of our research question demanded a qualitative approach: While quantitative methods are concerned with attempts to quantify social phenomena and collect and analyze numerical data, qualitative methods emphasize personal experiences and interpretation over quantification, and are more concerned with understanding the meaning of social phenomena. Whereas quantitative methods focus on the links among a smaller number of attributes across many cases, qualitative methods on the other hand, focus on links among a larger number of attributes across relatively few cases.

The privacy phenomenon is a highly complex and nuanced research topic. Qualitative methods offer several distinct advantages over quantitative methods in capturing the essence of this complexity, including more opportunities to explore the concept in significant depth, flexibility to detect and discover unexpected phenomenon during the research, ability to expose rather than impose meaning, and ability to investigate processes more efficiently (Strauss 1987, Strauss and Corbin 1990, Morse and Richards 2002). Qualitative research methods are known to be more credible and trustworthy compared to quantitative methods for research topics that are highly contextual and their effectiveness depends on a broad understanding of human behavior (Nielsen et. al 2002). Nielsen notes that translating users' perception into design recommendations (which was the focus of this work) requires a qualitative analysis that pairs observations with interpretive knowledge of usability principles. He also suggests that in order to incorporate usability into system design, designer are often required to combine and trade-off design guidelines, which depends on some grounded understanding of the rationale behind the principles that only qualitative methods can provide. Qualitative studies are also known to be less brittle and thus less likely to break under the strain of a few methodological weaknesses, and more likely to lend good results that rely on understanding users and their observed behavior (Nielsen 2004).

Among various available qualitative research methods (e.g., case study, field study, ethnography, action research), grounded theory has been recommended in the literature for situations where the goal is understanding a complex area, revealing user experience, or constructing a theoretical framework based on reality (Glaser 1978, Strauss and Corbin 1990, Morse and Richards 2002). Recently, grounded theory has been successfully applied in a number of information systems studies (e.g., Orlikowski 1993, Scott 1998, and Pandit 1998, Pace 2003), as a method that has significant implications for the design in the substantive area of study, since it emerges from the experience of participants and is directly aimed at finding their main concerns and generating a theory of how they behave to resolve that main concern. (Orlikowski 1993) suggests that because the grounded theory methodology is inductive, contextual, and processual, it suits those studies that are seeking to develop "a context-based, process-oriented description and explanation of the phenomenon", rather than "an objective, static description expressed strictly in terms of causality". Our goal in this study, too, was to learn the *process* of how people understand and deal with information

privacy in social systems, to identify their privacy concerns in this environment, and to propose design guidelines for privacy management mechanisms that support users in addressing those concerns. Based on the above discussion, the grounded theory methodology seemed appropriate for the purpose.

3.2.2 Glaserian vs. Straussarian Approach

Although the grounded theory methodology was initially co-discoverered by Barney Glaser and Anselm Strauss in 1967 (Glaser and Strauss 1967), a split between Glaser and Strauss started to evolve over their view in 1978 after Glaser published (without Strauss) *Theoretical Sensitivity* (Glaser 1978). Today, grounded theory researchers are expected to state whether they are taking Glaser's approach (Glaser 1978, 1992, 1998), or Strauss and Corbin's (Strauss 1987, Strauss and Corbin 1990, 1998). Rather than disputing which version generates a more valid account of users' experience, the main concerns in clarifying which approach has been adopted is to acknowledge awareness of the two different philosophical views behind the method, to be clear which method is chosen and why, and not to mix methods.

The Glaserian version represents the original version of the grounded theory, with an emphasis on looking for patterns of behavior that is relevant and problematic for those involved in a process. On the other hand, the Strauss/Corbin rendition is a more densely codified structured operation: The Strauss school suggests a *coding paradigm*, consisting of open coding, axial coding, and selective coding, while Glaser school suggests the less structured process of *theoretical sensitivity*. The Strausserian approach has been criticized by Glaser for putting emphasis mainly on *forcing* data into a predetermined paradigm, thus resulting *a full conceptual description,* instead of a theory that is grounded in data (Glaser 1992).

Our study employs the Glaserian approach, because the main objective of this study is building substantive theory to explain a pattern of behavior rather than full conceptual description. The Glaserian version is also better suited for studying complex entities because of its ability to capture the complexity by creating a multifaceted account of a process in action (Locke 2001, p. 95). Furthermore, this version links well to practice and is more likely

to be usable to those involved in the process (Locke 2001, p. 97), because the resulting theory is closer to the context and concerns of the participants.

3.3 Locating the Study

Since a grounded theory method looks for emergence of theory from the data, grounded theory researchers are advised to choose samples in a way that maximizes access to the phenomenon under study by selecting cases in which it is most evident (Morse and Richards 2002). Informants chosen for study must be expert participants with rich, extensive prior experience with the phenomenon in order to be able to provide the researcher with a valid account of their experiences. For these reasons, we needed to adhere to three criteria in locating our study:

1. **Finding the right tool:** We needed to choose a social software tool that provides some form of privacy management, preferably at an advanced level
2. **Finding the right users:** We needed to find a situation where the tool was used extensively, preferably over a long period of time, so that users were properly familiar with it and were not novice users, and,
3. **Finding the right context of use:** We needed to locate a context of use where both concepts of information sharing with various groups and privacy were paramount.

In other words, not only did we have to carefully select the appropriate tool, but we also had to locate advanced users and proper context of use for it. Together, these criteria imposed serious constraints on the process of recruiting subjects for the study. We next explain how we approached the problem of locating the study and how we overcame the aforementioned obstacles.

3.3.1 Finding the Right Tool

We initially focused our efforts on finding the appropriate tool. For that, we needed to find a social software application designed to support sharing of personal, professional, and social information of various degrees of sensitivity with various audiences. After an extensive review of the existing tools, we came across a tool called Elgg (Tosh, and Werdmuller 2004, Tosh 2005) that was designed to facilitate sharing various personal

artifacts with non-uniform groups of audiences. Elgg is an open source social utility designed to enable effective connection of people and information resources and facilitate creation of communities. The main components of Elgg include tools for online journals (weblogs), file sharing utility which enables users to store and showcase their work in a range of formats, collecting of news using feeds aggregation, tagging, and social networking. Elgg is primarily aimed at educational contexts, where it is used as a Personal Learning Space (PLE) to improve the experience of the education community with assessment, accreditation, career tracking, and enhancement of the learning process. However, it has also been used in organizations ranging from France Telecom R&D to the University of Brighton to MIT (O'Hear 2006) as a knowledge management tool that facilitates connection of expertise, resources, and social ties.

One key feature that motivated us in choosing Elgg was its strong emphasis on its permission architecture, which had resulted in reasonable support for privacy control at a fairly granular level that other tools simply didn't have. Elgg enables users to restrict access to many of the artifacts they dispose in the system, including their files, weblog posts, and profile items. The current privacy management mechanism in Elgg restricts access to resources through the creation of *groups* and *communities*. Elgg only supports closed groups, where group members are invited by the group owner/creator from his/her friends' list and need to accept the invitation to join the group (in Elgg, a user can mark any other user as his/her friend without the need for confirmation; however, friendship is not reciprocal). Elgg also supports communities, which are created and/or joined around topics that interest their members and are open for everyone to join. Users often find and join communities that interest them through either keyword search or exploring tags. One's groups and communities then become an option in the "access restriction" dropdown menu for each of his/her resource, beside the two default public and private options. This way, the user can categorize his network into different access groups and communities with different access privileges based on his sharing preferences. The model is equivalent to an RBAC (**R**ole **B**ased **A**ccess **C**ontrol) model (Sandhu et al. 1996) administered by the user (the concept of closed groups is the equivalent of the concept of roles in the RBAC model). Together, providing the functionality to dispose various artifacts in the application and to share them in groups of different dynamics created a suitable environment for the privacy issues to

surface, which made Elgg the right choice for the purpose of this study. Our next step was to locate the right users.

3.3.2 Finding the Right Users

The right users for this study would ideally come from a community where Elgg was used on a regular basis for the purpose of information sharing in various contexts. While exploring various options to identify such user community among the general user population of Elgg, we came across information regarding activities of a UBC entity named UBC Office of Learning Technology (OLT). We found out that at the time, OLT staff members were in the process of implementing pilot deployments of various PLEs in different UBC programs, with the goal of finding a winner to recommend to the whole UBC community. We contacted OLT's Community of Practice Coordinator to introduce Elgg and suggest they try it in one of their pilots, with the goal that once the tool was deployed in a program, we approach them for interview. Interestingly, our consultation with OLT revealed that Elgg was already one of their candidate PLEs and in fact, they already had an Elgg pilot running in a UBC program called "Transition" for over a year.

The UBC Transition Program (or simply "Trans") is a highly competitive, two-year pre-university program for high school students in grades 9, 10, and 11. Funded by the Vancouver School Board, the University of British Columbia, and the Ministry of Education, the goal of the program is to prepare academically gifted students to achieve their goals of early entrance to university and successful studies thereafter (VSB Web Site). Over the course of two years, students aim to complete the required high school curriculum, along with some university level coursework. When they graduate from the program, the students usually have an early acceptance into UBC. Students in both years are required to complete 30 hours of community service, and complete a PLE as part of their Individual Educational Planning. An extra course known as "Self and Society" accounts for the Planning 10 credit, and is taught by the coordinator. Involvement with other extra-curricular activities is also highly encouraged by teachers and staff (VSB Web Site).

Finding that the students in the Transition Program were using Elgg for Individual Educational Planning as part of their curriculum was a lucky coincidence, since it provided

us with a user community for our selected tool that were using it on an ongoing basis for a reasonably long period of time. The next step was to make sure they were using Elgg in a context where the issues of privacy and selective information sharing were important.

3.3.3 Use of Elgg in the Transition Program

Transition students were required to build and maintain a repository of digital artifacts on Elgg for demonstration of competence and reflection on their learning. Upon being accepted to the program, each student was asked to sign up for an account on Elgg, create a personal profile, and build a social network with other Trans students and the program coordinator. All students were then added to a group called "Transition", owned and maintained by their coordinator, where news and information addressed to all students would be posted. As part of their environment, each student was provided with a blog, and students were encouraged to regularly post their reflections on the topics covered in the classroom and on their learning process in their blog. Students also needed to join and participate in the "Self & Society" community, a special community created for their Self and Society course to share ideas and post opinions related to the weekly presentations that was part of their curriculum. For each of these artifacts disposed to the tool (e.g., weblog posts, profile items, and personal reflections posted to the community blog), they had the option of regulating access (i.e. making it visible to only oneself, all or part of their social network as categorized into different group, members of a specific community, or everyone in the open Elgg community). Additionally, students needed to create various teams for doing collaborative projects, and were encouraged to use Elgg for the purpose of communication with team members and posting project-related material. The option to make these groups public or private was left to the students. However, they were required to make their final submission accessible to the course instructor, for the purpose of evaluation.

In addition to the activities related to the Transition Program (which was a requirement for fulfilling their curriculum), students were also a part of the large and growing open community of Elgg users with over 300 user population at the time of the study. As a community of gifted students, many of these students were also blessed with creative talents such as creative writing, poetry, and drawing or animation, and were using

Elgg to post their creative artifacts to showcase their work to this community and get feedback.

Since the combination of the artifacts along with discussion/reflection functionality provided a rich view of a students experiences and skills, students were asked to include their Elgg as part of their application package for early admission to UBC after graduating from Trans. As such, many were encouraged to also include personal life experiences, awards, non-academic activities, and other character/learning revealing artifacts in their profiles. All these factors made use of Elgg a serious activity for Trans Students, and made them carefully target different artifacts to different audiences, which was an essential requirement for the issue of privacy preferences and selective disclosure of information to emerge.

Confirmation of the suitability of the context of use was the final step in the process of locating this study. It must be noted, though, that even though the study is situated in the context of Elgg, constant effort has been made not to limit the discussions to the specifics of the application. Instead, we treated Elgg just as a focal point to ensure that the subjects had the experience with a system that allowed them to manage their privacy directly. Otherwise we were careful to focus our investigation on the more general area of information sharing behavior in the context of a social-personal information management system, as described in the following sections.

3.4 Data Collection

The main characteristic of the grounded theory methodology is theoretical sensitivity; meaning that data gathering and data analysis must go hand-in-hand and each piece of data must be viewed as theoretically rich, linked to other data, and linked to the emerging ideas. As such, the data gathering process in the grounded theory method is considered active cognitive work. An important factor to consider when deciding on a data collection approach for a qualitative study is the issue of *methodological congruence* (Morse and Richards 2002). Methodological congruence suggests that the research method (i.e., phenomenology, action research, grounded theory, ethnography), research strategy (i.e., observation, questionnaires, interviewing, videotaping), and research technique (i.e.,

statistical analysis, data coding, data abstraction) must be selected in accordance with each other for the research goal to materialize. In keeping with the concept of methodological congruence, we selected semi-structured, in-depth interviews for our data gathering strategy, suggested as one of the best fits with the grounded theory methodology (Glaser and Strauss 1967, Morse and Richards 2002). Unlike structured interviews, semi-structured interviews have a flexible and dynamic style of questioning directed towards understanding the significance of experiences from the informants' perspectives. This strategy is primarily suitable for situations where the researcher knows enough about the domain to develop questions, but not enough to anticipate answers (Morse and Richards 2002). Our interview strategy involved asking open-ended questions about key topics that covered the research ground, to allow informants to discuss what is important from their perspective. We then used both planned and unplanned probing to uncover details and specific descriptions of the informants' experiences. The interviews were structured around a list of topics based on the research questions that needed to be answered, including questions about sharing preferences with regard to the type of information, the person or group with whom the information was shared, and the purpose behind sharing. The core questions that the theory aimed to find answers to were as follows:

TQ1. What is the incentive behind using Elgg for information management?
TQ2. What are the categories of objects that users share? What is the incentive behind sharing?
TQ3. What leads to the need for privacy? What categories of objects are perceived to need protection? What factors shape this perception?
TQ4. How is information shared with various audiences?
TQ5. What are privacy needs and challenges? What strategies do they employ to deal with these challenges?

It is recommend that theory questions should not be asked directly in the interviews with the participants; rather, they should just govern how the interview questions are designed (Wengraf 2001, p. 61−62, Morse and Richards 2002, p. 38). To quote directly from (Morse and Richards 2002 – p 39):

"…Deciding on a topic ***locates*** *your research… Framing a qualitative question is harder, because it requires you to think about* ***what needs to be asked*** *as well as* ***what you can ask and reasonably expect to have answered*** *given your resources and skills…"*

Our research questions were initially broken into a set of questions that the theory must answer. Likewise, for each theory question, a corresponding set of interview questions was developed. While theory questions are formulated in the language of the researcher and the research community, interview questions are formulated in the language of the participants. Here is a list of our interview questions and how each relates to the theory questions described above:

IQ1. Tell me about your experience with Elgg; how do you use it? (TQ1)

IQ2. How much time do you spend on Elgg during a typical week? (TQ1)

IQ3. How has Elgg served you during the course of your study in Trans? Do you see any benefits in using such a tool in Trans? (TQ1)

IQ4. Do you think you would continue using Elgg after the end of your Trans program? (TQ1)

IQ5. Give me an example scenario when you would use Elgg to share information with others; e.g., peers, instructors, program coordinator. (TQ2)

IQ6. How sensitive are you about privacy of information you put on Elgg? (TQ3)

IQ7. Can you think of examples of information you would NOT share on Elgg? Why do you consider them private? (TQ3)

IQ8. What other online tools do you use besides Elgg? (TQ4, TQ5)

IQ9. How is your use of [the other online tools] different from Elgg? (TQ4, TQ5)

IQ10. Have you ever used [the other online tools] to share artifacts? (TQ2, TQ4, TQ5)

IQ11. Can you think of a situation when you regretted sharing information on Elgg or other online tools? (TQ3)

IQ12. What factors are important to you when sharing something online? (TQ2, TQ4)

IQ13. Tell me about the groups and communities you are a member of (whether on Elgg or not) (TQ4, TQ2, TQ3)

IQ14. What privacy management feature do you consider necessary/nice to have that Elgg does not provide? (TQ5)

IQ15. Have you ever refrained from putting something on Elgg because of lack of adequate support for privacy? (TQ3, TQ5)

All interviews were carried out in the Transition Program office and were about an hour long. All of the interviews were tape-recorded with the informants' permission, and later transcribed to provide accurate records for analysis. The interviewees were informed of the voluntary nature of participation and of their right to skip any questions and to stop the participation at any point. Standard procedures were followed to maintain the confidentiality of the interview data and the anonymity of the informants.

3.5 Participants

After getting approval for the study from UBC's Behavioral Research Ethics Board (BREB) and Vancouver School Board, we recruited nine students of the transition program as our initial set of participants. Recruitment was done through distribution of flyers by the program coordinator. We then sorted students based on how active they were on Elgg, and started the interviews from the most active ones. The participants' ages ranged from 15 to 17, and the gender balance was almost evenly split with 5 females and 4 males. All participants were quite confident with using the Web. The average web usage reported by the participants was between 20-30 hours per week.

The procedure that led us to select the Trans students for the study is in line with a procedure called *purposeful sampling*. In this process, participants are selected based on their knowledge of the phenomenon under study and their willingness to reflect on their experience. Since active use of the environment was part of the their curriculum, these students had in fact a rich experience in using Elgg for information sharing, and were quite articulate in expressing their opinions about it.

Data transcription and analysis started as soon as the first interview was conducted, and emerging code words and categories resulted from previous interviews were considered when conducting new interviews. Annotations and memos were used throughout the interviews to record context, descriptions, impressions, and reflections. The analysis of the data gathered from our interviews with this initial set resulted in identifying the basic social processes (BSPs) of the grounded theory. Identifying the core concepts was a crucial step in providing an understanding of the phenomenon. After this stage, came *theoretical sampling*, a procedure through which we consciously selected participants according to their potential for developing new insights or refining the insights that had already been gained. Unlike statistical sampling, which aims to be representative of the population under study, theoretical sampling aims to maximize opportunities for exploring emerging concepts and relationships. For this process, we interviewed 3 more students who had extensive prior experience with various sorts of open online environments in addition to Elgg, including participating in social networking sites, online forums, or keeping a personal weblog. We also redirected the interview questions in a way to reflect our new goal of verifying the emerging theoretical themes and their relationships. The experiences of these 3 participants particularly helped in identifying places where the current privacy mechanism was falling short and users felt the need to employ strategies to overcome the insufficiencies, such as switching to other platforms

It has been emphasized in the grounded theory literature that there is no rule on the sample size (Morse and Richards 2002, Glaser and Strauss 1967). The number of participants recruited for a grounded theory study is determined by the quality of the participants' experience, their ability to reflect and report on their experience, and the requirement of further theoretical sampling. Data collection ceases when the researcher finds indicators pointing to *theoretical saturation*, the point at which he/she could identify interchangeable examples showing the same phenomenon in different instances. After interviewing our theoretical samples, we were convinced that additional data was no longer adding to the concepts and relationships being developed. As such, our data collection was ceased at this point.

3.6 Data Analysis

As with most qualitative research, data collection and data analysis occurred simultaneously in this study. The main component of the data analysis process in grounded theory is the *constant comparison* method where concepts are generated through comparing incidents with incidents, incidents with category, and category with category (Glaser 2001). Our theory was derived from the data using a constant comparative method of analysis with three stages: open coding, theoretical coding, and selective coding. This process allowed us to move from code words, to concepts, to categories, and finally, to the theory. The next following sections explain each stage in detail.

3.6.1 Open Coding and Initial Category Building

Initially, the aim is to generate the basic categories from which to build the emergent theory. This stage of analysis involved going through the interview texts and applying code words to sections that identified pertinent concepts, following (Glaser and Strauss 1967) description of open coding. A list of the code words for all transcripts was then compiled and compared against the original transcripts to ensure that a code word was used consistently throughout all the transcripts. Patterns, common themes, and differences were identified and assigned to categories. Notes were taken of emerging concepts, the ideas they represented, and the relationship between codes. The whole categorization process was done by one person to further ensure the consistency of code words.

(Glaser 1978, p 57-58) advises that researchers must follow four rules in the process of open coding:

- They must analyze the data line by line
- They must do their own coding
- They must interrupt coding to write memos of their ideas, and
- They must continually question what category (or property of a category) each incident indicates.

By following these rules, the open coding process enabled us to:

- Raise the empirical level of data to a conceptual level suitable for theory generation

- Stay theoretically sensitive
- Focus on patterns among incidents which yield codes, and
- Identify direction for theoretical sampling.

The qualitative analysis software named NVivo was used to label incidents in the data with code words and to write theoretical notes that captured momentary thoughts. NVivo allows the user to code a document simply by highlighting a passage of text and assigning it a code name. The code may be newly created or selected from a list of existing codes. The software not only helped with the abstraction process, but also with the analysis of the emerging concepts and ideas by providing tools for indexing them in trees. Table 3.1 provides an overview of our initial open codes after the first four interviews and examples of how they were extracted from the interview data.

Table 3.1 Open codes

Open Code	Examples
Benefits of Using Elgg	Keeping track, integrating various personal artifacts, organized, easy to share info
Examples of Personal Artifacts	profile items, project-related material, reflections, creative writings, poems, animation, blog posts, evaluation results, educational affiliation, awards
Reason for Sharing	convenience, anonymous feedback, peer feedback, presentation, collaborative work
Reason for not Sharing	contain images or a personal reference, weird feeling that someone knows something, don't want people to know I am in a group, losing control, no credit, competition, not appropriate, self-critical, affecting people's attitude, vulnerability, interpreted out of context, relevance
Moving Between Public and Private	critical reflections, project material, samples of creative work
Reason for the Transfer	outdated, not ready yet, being asked/encouraged to share something, after due date, changed my mind
Examples of Target Audiences	instructors, coordinator, peers, parents, potential schools to apply, applying for scholarship, public, project groups, online communities
Small Communities	focused, better sense of citizenship
Closed Communities	better contribution, better credit, reliable, better trust
Open Communities	more comfortable after some interaction, introduced by someone, sense of familiarity (due to knowing people opinions), less control/feedback over usage, anonymous feedback , anyone can join so not considered safe/reliable
Elgg Insufficiencies	can't hide that something is not shared, can't limit a conversation to relevant people, can't make a part private, can't share with only instructor
Dealing with Insufficiencies	multiple tools, no link, code word, writing anonymously, hiding private stuff

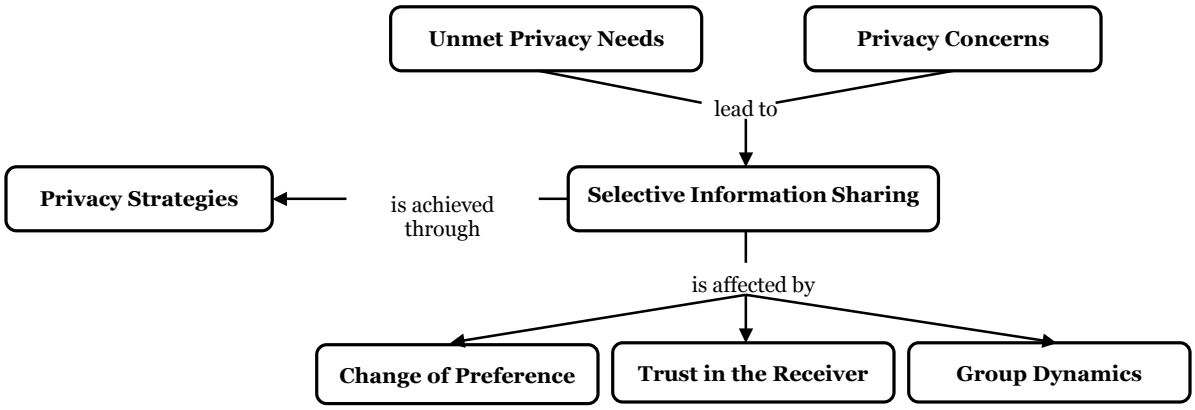
3.6.2 Theoretical Coding

The purpose of the theoretical coding stage is to help develop the emerging categories and to make them more definitive and useful. This stage of analysis involved following the lead of the data to find related instances and increase the degree of abstraction, by taking the concepts that emerged during open coding and reassembling them with propositions about their relationships.

As the coding process continued, new categories emerged and new incidents fit the existing categories. As categories were generated, the next incidents were compared to the categories. This constant comparison of the incidents soon started to generate the relationship between the categories. The relationships, like the concepts, emerged from the data through a process of constant comparison. By means of constant comparison, categories were refined, basic properties of each category were defined, the relationships between categories were identified, and the identification of patterns was facilitated. These emerging propositions then formed a theoretical framework, which served as a guide to further data collection and analysis.

Like many grounded theory researches, theoretical coding in this study was not a distinct sequential step following open coding. Rather, the two phases proceeded quite naturally together. Theoretical coding can only begin after identification of the categories, but often a sense of how categories relate to each other shapes during open coding (Glaser 1978, p. 56). In this study, our focus was more on open coding when discovering codes within the data, and more on theoretical coding when sorting open codes into categories and integrating memos. Figure 3.1 shows our initial idea of categories and their relationships after the process of theoretical coding.

Figure 3.1 Initial categories after theoretical coding



Both the emerging categories and their relationships were repeatedly compared with other incidents in the transcripts to ensure validity. Also, themes and relationships were kept only if they were reflected in the experiences of many participants. Using this constant comparison method, some codes were subsumed under broader or more abstract categories. At the end of this stage, our core category (selective information sharing) and the relationships between the categories were shaped.

3.6.3 Selective Coding

As the theory developed, it evolved around categories and their relationship that reflected the main theme of the study, selective information disclosure. The identification of the core category led to selective coding, the process of delimiting coding to only those concepts and relationships that relate to the core category in sufficiently significant ways to be used in creation of theory. Selective coding helped us define each of the emerging categories in detail, and also made the process of data collection and analysis more focused within the overall context developed through open and theoretical coding, resulting in a more focused theory with a smaller set of higher-level concepts.

3.7 Summary

This section has described the research questions, goals, and methodology that have been used for our foundational study of information sharing behavior of end-users in a social software designed for personal and social information management. A detailed description of the study is presented, including methodology, participants, and the process of data collection and analysis, along with the unique challenges of locating the study and how they were addressed. Reasons for selecting the grounded theory and adopting the Glaserian version as the methodology of choice for the purpose of the study have been discussed. An overview of the defining characteristics of the Glaserian grounded theory is presented, including the processes of collection, coding, and analysis of data. This sets up the context for the presentation of the resulting grounded theory of information sharing behavior of end-users in SPIMS, which is presented in the next chapter.

Chapter 4

A Grounded Theory of Information Sharing Behavior in SPIM

This chapter explains the grounded theory that was developed during the course of our foundational study. The theory is presented through a concept map of major categories and their relationships, which are explained in detail. Our findings suggested that privacy is a main concern for users of a social-personal information management tool, and illustrated challenges users face in ensuring privacy of their information and strategies they employ to achieve the desired level of privacy. Results also identified factors that affect users' decisions regarding disclosure of their personal artifacts to various people and groups in a SPIM tool, including change of privacy preferences in various stages of the information life cycle, the nature of trust between the owner and the receiver of information, and the dynamics of the group or community within which the information is being shared. Together, these themes portrayed a clearer picture of users' perspective on the privacy of their information in the SPIM domain.

4.1 The Concept Map

The concept map in Figure 4.1 represents the theory that emerged from the data in this study. The diagram consists of a collection of boxes connected by arrows. The boxes represent concepts and the arrows represent relationships. Together, they represent a process that leads to selective information sharing. The numbers in square brackets that appear alongside the concepts represent the number of participants whose comments provided empirical indicators for that particular concept. However, these numbers do not represent the absolute importance of one concept over another. Given the semi-structured nature of the interviews, not all participants had an equal opportunity to discuss all of the concepts and relationships that are represented in the diagram. The text that follows in the

coming sections is supplemented with direct quotes from participants in support of each concept or relationship under discussion. It should be noted, though, that the presented quotes are not an exhaustive list of all of the participants' comments and are only meant to serve as selected examples of how the data analysis process has led to the particular concept or relationship. Where applicable, findings from existing literature that are relevant to the concept or relationship under discussion are also discussed. Where appropriate, interpretive commentary is provided to clarify speculations that go beyond the available evidence. In the next sections the concepts and relationships represented in the concept map are explained in detail.

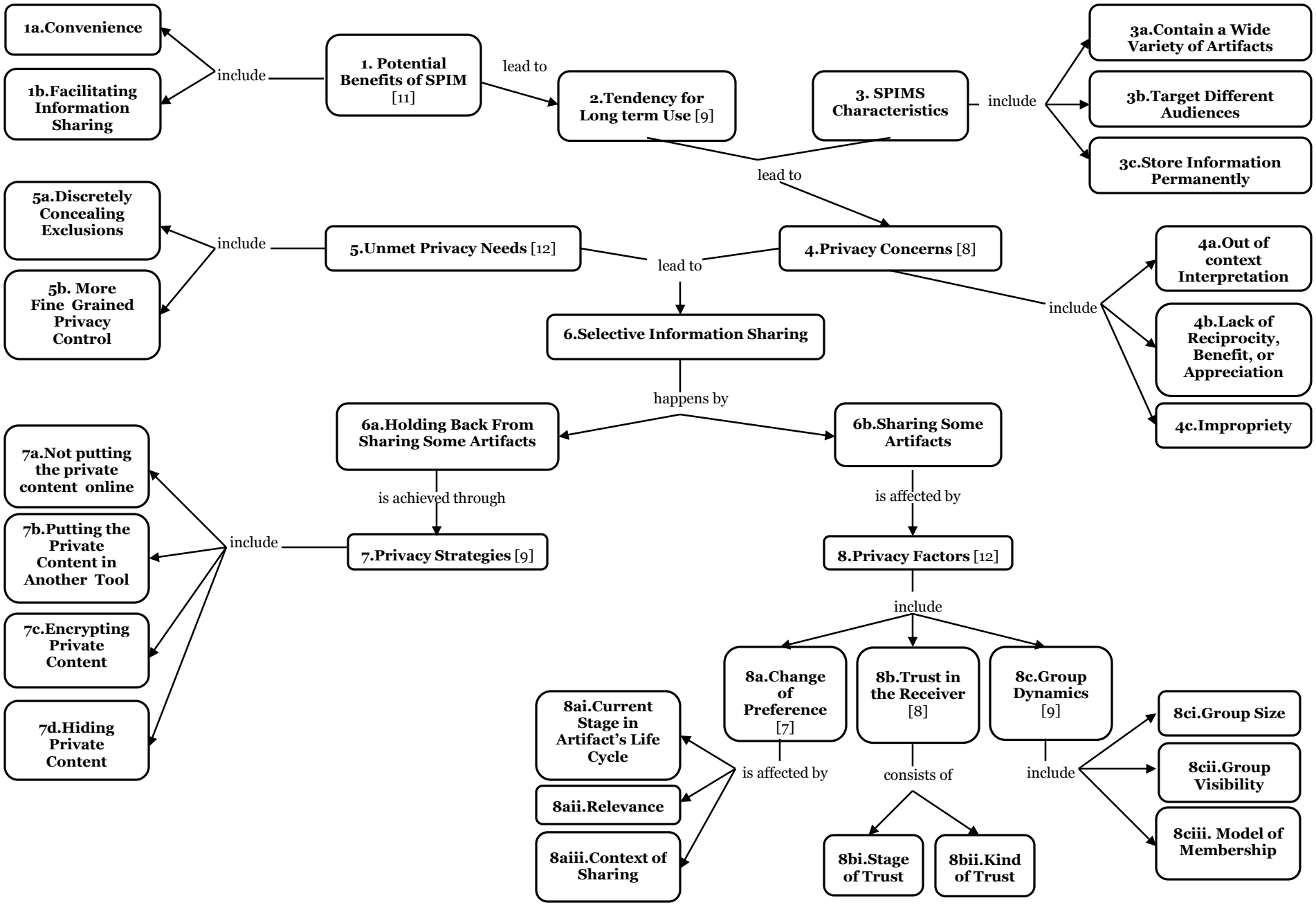


Figure 4.1 A concept map of information sharing behavior in a social-personal information management system

4.2 Benefits of Use of Elgg for Information Management and Sharing

Our results indicated that for most of the participants, use of Elgg was envisioned as a continuous process. 9 out of 12 participants stated that they would continue using Elgg after the end of their program (where they were required to use it), because of the potential benefits the tool provided them. The perceived benefits of the tool as reflected in participants' comments included the convenience of having all their information in one central place and over the Internet, where they could easily share with potential audiences rather than having to send their artifacts to them individually; ease of information management and organizations, enabling "keeping track" and reflection on their improvement over time, and getting unbiased feedback on their creative artifacts from a community of people who share the same interest. The following sample quotes reflect participants' perspective on the benefits of Elgg that lead to tendency for long-term use:

> "It is very useful to have everything in my Elgg, because then other people can see what work samples I have. Like say, if I am applying for scholarships, then I want them to have samples of my work without me having to send them things, so they can access the site whenever they want to. It's like I have an online resume." (participant #4)

> "It's better to keep things in Elgg. You can then get feedback or keep track so that you don't lose anything. Everything is there and is in one place. If I keep things on my computer then after a while there are too many things and I can't find things easily. So I keep things on Elgg and I update it whenever a new milestone is achieved, or an educational goal changes. It's more organized this way." (participant #6)

> "Sometimes getting feedback from your peers, it's only a confined set of people, so sometimes you want people outside your group who don't know who you are [to comment on your work]. Because sometimes if you know who your friend is, it sort of censors your real comment, so sometimes if you get outside the box [feedback], it's sort of a more open feedback. I think that's more reliable." (participant #7)

"It's a good way to keep organized. You know where your things are if you ever need them." (participant #5)

"There are two benefits to using Elgg; first, it has good integration potential, meaning, I can access my course materials all in one house. I can even present from where my data is. Second, it makes it easy to connect with classmates because they are all on Elgg, too. So, if I want to, say, organize an event or something, I know they will all see it." (participant #1)

"For me, the main use of Elgg would be after Trans: It is my ePortfolio; it will be part of the application package I send out to schools. That's the reason I am using it in the first place. Besides, all my Trans friends are using Elgg now; so we can stay in touch if we continue using it after Trans." (participant #3)

"I found [on Elgg] a community on digital photography, and it's going great; I will continue to participate in that". (participant #12)

As the quotes show, even though Elgg was primarily introduced to students for educational purposes, they were also using it for a number of other purposes, including information management (participants 6, 5, and 1), demonstrating competence (participants 4 and 3), making new connections and keeping up with current connections (participants 1, 3, and 12), and interacting with others in communities of interest (participants 7 and 12). This identifies convenience and ease of information sharing as two major benefits behind using an online tool for the purpose of personal and social information management. It is also a confirmation of findings of previous studies in the literature (e.g. Whittaker and Sidner 1996 study of email) that as with other computer-mediated social technologies, when given a rich environment that provides support for both work related and social activities, user communities will adopt it for more purposes than was initially conceived. Support for privacy may not have been initially considered in the tool for all these areas. That's when privacy concerns emerge.

4.3 Privacy Concerns

As a social-personal information management tool, Elgg is designed to support both social activity and engaged work practices and as such, personal artifacts included in Elgg may span a wide range of types, from scholarly work, to personal opinions expressed in a weblog, to bookmarks and personal collections. Various pieces of information are likely to be targeted to different groups of audiences that are not necessarily static. As a system that is designed for collaboration and communicating with others, the artifacts included in Elgg are used during various stages of life for different purposes. For example, while students were using Elgg to discuss course project with classmates while they were in the Trans program, they anticipated to later present their ePortfolio to potential schools they planned to apply, where they needed to expose a different view of their space.

Another attribute of social systems is the persistency of information disposed in them. While this is often considered as a benefit since it enables information re-use (as reflected in participant #4's comment above, for example), over time the aggregation of ones' online information creates a persistent and searchable online identity for the user, to which he may wish to expose different views to various audiences in different situations. The combination of these factors (variety of artifacts and audiences that are characteristics of a social software environment, information persistency, and the tendency for long-term use which implies that artifacts will be used in different contexts) sometimes leads to the emergence of privacy concerns. For the purpose of this study, a privacy concern is defined as follows:

> *Definition 4.1:* A privacy concern is the anticipation of an undesirable outcome that negatively affects a user's intention to share certain information artifacts.

Our participants expressed concerns over a variety of issues including affecting people's attitude by disclosing certain information about themselves, fear that some information might be interpreted out of context (mostly for personal opinions and reflections), lack of reciprocity, appreciation, or benefit, and sometimes just the simple

feeling of "awkwardness" at the thought of exposing certain information. Here are some of the participants' comments regarding some of their privacy concerns:

> "…even though that is an important part of my identity [referring to a certain interest], I just decided to take that off my ePortfolio, because although I don't mind my fellow Trans students know that, I don't want to find people who don't know me think I am weird…. " (participant #5)

> "…I don't always share information on scholarships; because they are highly competitive, by sharing I would just put myself in a less advantaged position. I know other people don't share such stuff, either; so…" (participant #1)

> "I usually prefer people not to know that I am coming to this program because that sort of affects the way that people think about me. By keeping my educational and social information from certain people who really don't know a lot about me, I am treated more like an equal." (participant #12)

The last quote is an example of nuanced nature of privacy and how a user's attitude towards privacy affects privacy preferences: while most people would be comfortable with sharing their educational affiliation (even more so if the affiliation is a prestigious highly selective program), this participant preferred to keep it private for a very personal reason.

4.4 Privacy Needs

Although all of our participants reported using Elgg's access groups to have control over disclosure of their various pieces of data, they were not always getting the results they were looking for. 8 out of 12 participants mentioned certain privacy needs that were not supported by the tool. For the purpose of this study, a privacy need is defined as follows:

> *Definition 4.2:* A privacy need is expectation of the existence of certain privacy related functionality in the tool.

Among the privacy needs that were brought up in participants' comments were the need for hiding the fact that something is not being shared, and need for the ability to

control privacy at a more granular level, including a more fine-grained categorization of audiences as well as resources. Here are samples of users' comments with regard to their various privacy needs:

> "The problem I have with that [current privacy mechanism in Elgg] is that when I let some people see something, other people can see that there is something, but they don't have access to that. So they are like: oh, can I look at it? and then sometimes, you just don't know whether you want to share with them or not, and it's kind of weird to say no right away. So, then sometimes, I just rather keep it all private or all public so not to have to make that decision." (participant #9)

> "What would have been nice to have, is for people who don't have access to it to see a blank page instead of a message like, sorry, you don't have access to this." (participant #8)

These two quotes represent an interesting privacy need: both cases emphasize the fact that once the existence of an artifact is known to an audience (whether an individual or a group), the ability to deny its complete exposure is of little use due to social implications. This implies that in order for a privacy management to be successful, it must support discrete concealment of exclusions.

> "I would rather keep my reflections private, but for example, for [a particular course], we need to write down our reflections so that [the instructor] could see what we took out of the sessions. Then I need to move my reflections from private to public." (participant #11)

This was a universal need among our participants, as all of them at some point needed to share something with only an instructor. Like many other social tools, the only way a user can do this on Elgg is by creating a new access group with only the instructor in it, and then choosing this group as the access restriction group for the particular resource. This is not a practical solution from the instructor's perspective (since it requires instructor's membership in one group per student). It is also too labor-intensive on the part of the student, since it requires creating one group per each instructor, sending group invitations,

and managing group memberships. However, further discussion with the students revealed that the reason none of them had tried this exercise was that they didn't realize they could. In other words, it was not the inherent impracticality of the solution that hindered its usage, but rather, the inability of users to perceive such possibility. As a result, they opted for making their reflections public in order for making it visible to the instructor, even though some stated they preferred to keep the artifact (or at least part of it) semi-private (e.g., shared only with the instructor).

This situation also drew our attention to the insufficiency of the friends model, in which all "friends" are created equal and all relationships are reciprocal. In this particular case, although instructors were also part of each student's "friends" network, they needed different treatment in terms of information exposure, and this was not a reciprocal need (i.e., while to the student, the instructor was a special "friend" in terms of information sharing, the opposite was not true). Other studies have also discussed a similar issue, emphasizing that although most social tools support a uniform model of one's social network, the nature of user's information sharing habits with various members of this network could very well be non-uniform. A recent study of use of Facebook for educational purposes (Adam 2009) showed that while students were using Facebook to communicate with their peers, they were not willing to use it for sharing information with their instructor. Even though the instructor had originally joined Facebook to communicate with her students in a space that was comfortable for them, she discovered that students were not communicating with her on Facebook, but were instead using other forms of communications such as email, telephone, and MSN Messenger. The study concluded that this happened because students were not able to reconcile the notion of a "professional-student relationship" with the notion of "peer-to-peer relationship", which emphasizes that a model that accounts for the differences in interpersonal relationships might be better than the uniformed friends model for modeling information sharing in SPIMS.

4.5 Selective Information Sharing

A grounded theory explains how people resolve their main concern by employing a certain *process*. That process is called the *core category* of the grounded theory. Glaser defines

the core category as "a pattern of behavior that is relevant and problematic to those involved" (Glaser 1978, p. 93). For our Elgg users, we found out that the combination of privacy concerns and privacy needs leads to the process of *selective information sharing*; meaning, users employ selective information sharing practices to maintain privacy. This happens by sharing certain information artifacts, while holding back from sharing others. Selective information sharing was the core category that emerged in this grounded theory. Next, we set out to find *how* users employed selective sharing practices to manage information privacy:

- For artifacts that users shared, we aimed for identifying factors that affected their decision regarding sharing. These factors are discussed in section 4.6.
- For artifacts that users did not share, we aimed for identifying strategies that users employed to control the exposure of the artifact, and whether their decision to hold back from sharing was affected at all by the lack of sufficient support for privacy in the tool. These strategies are discussed in section 4.7.

4.6 Privacy Factors

For the purpose of this study, a privacy factor is defined as follows:

> *Definition 4.3:* A privacy factor is a condition that affects a user's decision regarding sharing a certain information artifact.

Privacy factors can be viewed along three main lines: the resources to be shared or protected, people/groups with whom the resource is being shared and to whom access privileges are granted, and the perceived usage of the shared resource. As the diagram indicates, one substantial factor was found in each dimension: our results indicated that users' perception of the sensitivity of a shared resource might change in various stages of the artifact's life cycle. We also found out that the kind and stage of trust that the user has in the person/group with whom the resource is being shared plays a strong role in user's sharing attitude. Finally, our results showed that users share differently in groups and communities of different culture and characteristics. A more detailed description of each element in the

diagram plus summaries of how each element was derived from the accounts of participants' comments is provided below.

4.6.1 Information Sensitivity: The Effect of Change of Preference

Our results show that as in other contexts, there are commonalities in the way users arrange their information with regard to sharing in the SPIM domain: there is a category of artifacts, such as critical reflections and personally identifiable information, which most users prefer to keep private or semi-private. There is another category of general information, which users feel safe to reveal publicly with a remarkable consistency. However, there is a significantly large set of artifacts that does not fit in either group. For this group of artifacts, one particular criterion that continually appeared in the accounts of users was changing sharing preferences. Rather than a simple classification of artifacts based on a binary scale of public/private, analyzing participants' reports showed that for this group of artifacts, most of them apply a more sophisticated and flexible scale of information privacy that reflects a *transition* between private, semi-private/restricted share, and public. In terms of interaction models, the two broad categories of information whose privacy settings probably don't change over time (those that are made completely public, and those that are either private or shared only with a limited, trusted audience) seem to be well served by the current "set on creation" models for assigning privileges. For the third category, however, there is need for more fluid techniques that can carry a resource through cycles of changing preferences. As the following comments show, this transition can happen due to several factors, including the state of the information, relevance, and change in the context of information sharing.

4.6.1.1 The Effect of Content Life Cycle

Our results indicated that for a certain category of artifacts, such as in-progress projects and samples of creative work, users often apply a *"privacy life cycle",* meaning they require different recipients (i.e. teachers, peers, public) to be able to perform different actions (i.e. read, modify, comment) on the artifact based on the current stage in the artifact's *life cycle*. Life cycle refers to the path that a particular piece of information artifact takes from its creation or receipt (e.g., creation or acquisition of knowledge), to distribution and use (e.g., publication and sharing of knowledge), and further to maintenance (e.g., storage and archival of information) and possible disposition (Tallon and Scannell 2007).

Our study showed that for dynamic artifacts (i.e., artifact owned/created by user that are subject to modifications) users' perception of artifact's privacy varies in different stages of the artifact's life cycle. For these artifacts, users seemed to dynamically match privacy and control to an artifact's state of "completion". For example, many users considered samples of their creative works private at the time of creation when descriptions, goals, and personal reflections were included with the artifact. However, during the work-in-progress stage (distribution and use), they would share the artifact with a restricted audience to obtain feedback, and would share it beyond the group who created it (often a larger/more public audience) once it was completed (moved on to maintenance stage).

> "Sometimes, I would put it [referring to samples of creative writing] on private because it has too much information about me that I don't want sharing over the Internet, or sometimes it has more private things "to me" to go public. But then sometimes they become "outdated" or I need to put them up as samples for assignments, or examples for a question." (participant #8)

> "My reflections are usually private, but for example, for [a particular course], we need to write down our reflections so that [the instructor] could see what we took out of the sessions. That's when I need to move something from private to public: It's because I need [the instructor]'s comments on it." (participant #3)

4.6.1.2 The Effect of Relevance

In some cases, the reason behind moving artifact between various classes of exposure stemmed from the fact that users expected to use the tool in various stages of their life for various purposes, and they felt that some information may not be relevant to the new audiences who would access their space:

> "My critical reflections are public for now; but I will change them to private when I want to provide my Elgg for scholarships. They don't need to see all my critical reflections. Not that they are self put-down or anything; in fact, criticism is the way we make progress, right? it will just be irrelevant for the purpose;" (participant #2)

4.6.1.3 The Effect of Context

One of the salient features of all social software systems is the co-presence of multiple groups that are relevant to an individual. The nature of user's interaction with each group might change over time. In our particular case, one issue that appeared in participants' comments was the issue of frequent move between competitive and collaborative context. Whereas they needed to be cautious about the degree of exposure for their artifacts to groups they had a competitive relationship with, they shared more freely when the context of sharing changed from competitive to collaborative:

> "We have created a group for our [a course] group project in the past. There was this [...] assignment that we had and everyone needed to contribute by writing in the journal. So we uploaded the file into Elgg file repository and initially, gave access to it to only the group. Then when it was done, we also let [the instructor] see it, like we added her to the friends in the group. She was quite happy with the work, so she suggested we make it public so that others can see it, too." (participant #12)

> "I move things between private and public in my ePortfolio, which is mostly schoolwork. For example, say we have a lab assignment due on Thursday; I would post it up for me to look at in the private one, just to check that everything is completed before I submit. Only after the due date I post it in the public one, because of copying." (participant #1)

4.6.1.4 Other Factors

Besides the above cases, we encountered a few incidents of changing preference that could not be attributed to any of the above factors. In most cases, the participants' rational for moving something from private to public or vice versa was simply stated as "I changed my mind":

> "I used to have only a public one [blog], so I would put my critical reflections on public. But then sometimes I re-read it later and I was like, oh, I really don't want so and so to know that, and you kind of feel a little weird, knowing that they know that too. So I changed it to private" (participant #8)

"Most of them [blog entries] aren't public. Because when I first started everything was public, and I found that some strangers leave inappropriate comments. That was kind of annoying and unnerving. Because you don't want random people leaving you things. So I just changed it to friends-only." (participant #2)

4.6.2 Information Receiver: The Effect of Trust

One theme that was repeated in users' comments when reflecting on their information sharing behavior in various communities was the issue of trust. Our data showed that as expected, trust is a significant predictor of an online community members' desire to exchange information, and especially to share information.

The trust issue was mainly mentioned with regard to Elgg communities (as opposed to groups). The main difference between the two was that unlike Elgg groups that were often formed around a specific purpose and/or with known members, Elgg communities were typically emergent and often not task-oriented, and were open to everyone. Members of Elgg communities were typically strangers to one another and could be relatively invisible (i.e. if an individual only took information and did not contribute, it would have been hard for other members to know anything about him/her at all). Moreover, members often joined the community mainly based on a personal interest in community's purpose/goal, and there could be no further assumption on members' social characteristics or attitudes besides that. Our study confirmed existing studies of trust in online communities (i.e., Ridings et. al 2002, Hsu et al. 2007) that in the absence of a shared physical space and prior knowledge on the existing and potential members, the process of information sharing depends to some degree on the formation of interpersonal trust.

Our results indicated that users' trust in the persons or groups who will be the receivers of information plays a strong role in deciding about information sharing: users tended to share less with people/groups with whom they were in the initial stages of trust, and as their trust moved towards a more mature level over time, they began to feel more comfortable and share more. 7 of our 12 participants confirmed that they usually start cautiously regarding information sharing when they join a new community, but after participating in the community for a while their trust moves into a different level, and they

share more freely as a result. This is very much in line with the way face-to-face trust is shaped in the real world and between real people, suggesting that the notion of communities in social software systems must reflect the way human communities behave and work. The following quotes summarize some of the participants' comments that refer to the concept of trust:

> "Right now I am on a forum and I remember in the beginning I was really careful about exposing personal information, such as where I go to school or posting a picture. I would just ignore and leave myself out of it. After a while, you sort of trust them a bit more. I haven't been as far as putting a picture on, but I would say oh, I would get my license in a couple of years or something like that. But I won't make a reference to the fact that I am not old enough - I would just say I will get it in a couple of years. So, I am still pretty cautious about it; because after all, my trust just comes from interacting with these people over time. I mean, I just "feel" more comfortable after being in the group for a while." (participant #11)

> "After interacting in a group for a while I would feel more comfortable sharing with the group but its not always very comfortable, just more comfortable than before; Like, from "not very comfortable" to "sort of comfortable". I am not the kind of person who gets too comfortable over the net" (participant #2)

> "If you participate in an online community and you talk to people and they begin to give their opinions about something, you feel you begin to know who that person is by what they say are their ideas and what they like, and you develop a sense of knowing who they are, and they are no longer unknown; because we fear what we don't know and so if we get to know what that person stands for, maybe we can trust them some more." (participant #6)

> "I am not the kind of person who makes friends over the Internet easily and I don't really connect with forums well; but once that happened, though, I actually had my friend who had visited the forum for a long time. So, it was easier to connect because I had a really strong connection there." (participant #10)

Our findings go one step beyond mere confirmation of the importance of interpersonal trust in information sharing as acknowledged in existing studies, by revealing a key factor that has been sparsely addressed before: the *developmental* nature of trust and the effect it has on information sharing behavior of users.[1] Our results showed that users' sense of trust (either in another person or in the community as a whole) moves along a progressive path (i.e. from less trust to more) rather than following a binary trusted/not trusted pattern. Scholarly research on trust has supported the assertion that trust is multidimensional, and various dimensions for trust have been suggested in the literature (i.e., Butler 1991, Jarvenpaa et al 1998, Mayer et al 1995, Corritore et. al 2003). While a comprehensive discussion of trust is beyond the capacity of this work, we should note that our work mostly aligns with Corritore's conceptualizations of the trust construct into the four dimensions of generality, kind, degree, and stage (Corritore et. al 2003): comments from our participants mainly reflected the effect of kind of trust in the receiver, i.e. cognitive (#10) or emotional (#11), and stage of trust with the receiver, i.e. initial, intermediate, mature, and the transition from one stage to another (#2 and #6).

4.6.3 Information Usage: The Effect of Group Dynamics

Our results indicated that users' willingness to share something they have vested interest in also depends on their perception of how it will be used, with the dynamics of the groups or communities where the information is going to be shared being the most influential factor in deciding about information sharing. Our study revealed that users often hold back from sharing information in anticipation of lack of reciprocity, benefit, or appreciation, and loss of credit for their work. The theory suggested that when group/community dynamics are clear enough to convey to the users how their information will be used within the group, users may be better equipped to make informed decisions regarding how much they want to share within the group. Moreover, this predictability may be critical to making the decision to share information in the given context at all. One drawback of online information sharing in general is that although most tools provide facilities for creating and participating in groups and communities, there is no indication of

[1] In an integrative review of eleven empirical studies of online trust (Grabner-Kräuter and Kaluscha 2003), the authors found that in only three studies the authors explicitly discussed the dynamic nature of trust by pointing to the phase of trust they were investigating.

what the information sharing manners are in a particular group or community. While in real life, implicit group cultures play a strong role in information sharing attitude of the group members, these norms and cultures are not usually clearly specified for online groups and communities. The following quotes summarizes some of participants' reflections on the notions of group and community in Elgg:

> "Right now, we have a group for each course whose members are the students in that course; or maybe that's a community? yes, it must be a community, because the instructor and the students use the [community] blog for updated course information, upcoming deadlines, exam, etc., we also post questions and share thoughts and ideas on the subject" (participant#8)

> "At first we were not sure whether to use a group or a community for our course projects, but we finally decided that group is more appropriate, because it seems to be a more controlled environment [compared to community] in the sense that membership is restricted and no one outside the group has access to the material intended for the group members. So in a sense, it is a "more private" environment than community. But communities on the other hand have the blog; which is good for sharing information." (participant #9)

> "It took me a while to realize that moderation for a community only serve the purpose of placing a border between who can read and who isn't allowed to read the community profile and files, and to contribute to community blog; meaning, material on the public community blog still needed access controls to stay private between me and the community members" (participant #4)

While these participants discussed group and communities in general terms, a few were more specific in describing the effect of a particular group characteristic on information sharing. Applying open coding to these comments revealed the effect of various group dynamics on information sharing, including size, membership model, and visibility. For example, the following quotes show participants' perspective on the effect of community **size** on information sharing:

"[What I share in a community] also depends on the size of the community. Because some communities are really popular; there are lots of people; so you can't really get to know everyone. I am usually more comfortable when it is small, like say ten people. That's a bit more personal, and I get better credit for my contributions." (participant #11)

"I once created a community for [...], which was a closed community. My experience with that community was actually very positive: everyone would contribute actively and give others feedback on their work. But then, we all sort of knew each other, so it was more like chatting with friends... It was a small community, though. (participant #4)

And more on the effect of community **membership model**:

"The problem with anonymous communities [where providing real information is not a requirement for membership] is that you have no way of knowing who the comment is coming from… you can't trust them with their judgment: it could be a grade one kid or it could be a Ph.D. so it's not worth anything." (participant #2)

"[What I share in a particular community] would really depend on who else is in there. In Trans [the particular community they have for all Transition students] I know the students [who the community consists of], so I would share my opinion on certain things that I wouldn't mind sharing with them in person; but for some stuff, I would definitely not share." (participant #3)

And finally, on the effect of group **visibility** and the importance of getting credit for shared artifacts:

"To me there is a strong distinction between private and public groups. Private groups are invitation only, so I would appear with my real name and share practically everything. The public ones are open to everyone though; so I usually use a pseudo name and I am cautious not to reveal any personal information." (participant #2)

"There are different choices [of communities]; There are some that are more discussion-based, where you have to be a more active participant; There are some communities that offer stuff, like you can go there and take it if you want; There is

"everyone can join", and there is this thing like, you can only join by invitation and only members can take away stuff. for example, I like to post [my creative work samples] to a community but I don't want people to take it for free, so this is the place for me because there is the copyright thing. I prefer to share my [creative work] in those communities; because otherwise anyone can just come and take your stuff without giving you credit for that. Usually the discussion-based ones are pretty open, and that's OK." (participant #5)

"[When sharing stuff in a community] I'd like to know what they are doing with it, but they don't have to tell me. I mean I am offering it, so they can use it if they want to. If they want to tell me what they are doing with it, I would like to know that, too. I don't mind as long as they give me credit for it." (participant #12)

As the quotes show, one problem that our participants seemed to particularly suffer from was lack of clarity on the definition of a group vs. a community; especially as to how they differ and which one should be used for a certain information sharing purpose. Moreover, both the group and the community concepts as presented by Elgg are rather inflexible: only a very simple membership moderation mechanism is provided for the community concept, with options for "public" community with no membership moderation, and "moderated" community where membership must be approved by the community creator. The options for the group concept are even more limited: all groups are closed and membership is by invitation only. While all communities are publicly visible, groups are only visible to the group members. Even for moderated communities, the community blog is publicly visible unless separate access restrictions are used for each blog entry, just like the individual blogs. An analysis of users' comments, however, shows that they expect to have more control over the various aspects of their groups and communities, or at least need those aspects to be conveyed to them in a clear and unambiguous way.

In real life, users definitions of groups and communities vary more continuously than the simple private group vs. public community as defined in Elgg. Groups and communities can vary along several dimensions including size, visibility, and membership model. The relationship dynamics of the group/community are then determined by variations of these dimensions; i.e. open/limited number of members, public/private visibility, and open/moderated/closed membership. In that regard, the concepts of a self-defined *group*

that is only visible to members and controlled by the owner himself (as provided by Elgg) and the *community* that is necessarily visible to everyone and whose membership is poorly controlled are simply two ends of a broad spectrum as shown in table 4.1.

Table 4.1 Group and community continuum

	Group Visibility	Membership	Member List Visibility
Community	*Public*	*Open*	*Public*
Club	*Public*	*Moderated*	*Members*
Team	*Public*	*Fixed*	*Public*
Clique	*members*	*Closed*	*Private*

The theory thus suggest that SPIM tools in general can benefit from the addition of a more powerful group/community support where users have better control over administration issues, and the distinction between different groups and communities and their purpose is clearly defined based on size, public/private visibility, and open/controlled membership. We believe that a clear model of group characteristics can highlight the potential trade-offs between risks and benefits of information sharing in a particular group or community and give users' a high-level overview of the effects of their sharing decisions. Knowing how a particular resource could potentially be used in a particular community, users can then tune their sharing decisions accordingly.

4.7 Privacy Strategies

For the purpose of this study, a privacy strategy is defined as follows:

> *Definition 4.4:* A privacy strategy is an activity that a user employs to withhold sharing certain information artifacts.

Privacy strategies initially emerged from analyzing the data gathered from our first 9 participants. 6 of the users in this initial set reported using other online tools in addition to Elgg. Based on this fact, we decided to pick our theoretical samples from the pool of users who actively used other online tools in addition to Elgg, with the goal of clarifying whether

this exercise was motivated at all by lack of appropriate means for privacy management. Further analysis of data collected from our theoretical sample group (the next 3 participants) provided evidence that switching to a different tool was in fact one of the strategies that users employed to achieve their desired level of privacy when the tool failed to support it. In Total, 9 out of our 12 participants reported using other platforms with better privacy management mechanisms for their more private content. In more extreme cases, 4 had even decided to forego deploying certain stuff in an online environment because of lack of acceptable privacy levels, while 3 had decided to have their private content somewhere (i.e. a web page or weblog), but not to provide a link to it from places where her real identity was known. 5 had anonymous blogs for their more private reflections in addition to their Elgg blog, while 2 had chosen to stick to the same platform, but write their more private content in some sort of a "code language" so that it was meaningless to anyone other than themselves. The following quotes show some of the statements regarding strategies for maintaining privacy beyond what is provided by the tool:

> "Besides Elgg, I have two other ePortfolios, and a couple of weblogs. One is private and one is public. On the private ePortfolio, I have things that are actually more private, like it has information about me, that sort of stuff. The purpose of that is that I just want to write some stuff down, so that it is sort of "said" somewhere. Sometimes I don't want to keep stuff in my mind, like for example, a journal or something, I would put it on the private one." (participant #11)

> "I use [another platform] for more private stuff because there are settings for public or friends-only or you get to choose who gets to see it. If it is something you want the teacher to see but not anyone else, you can just set it that way." (participant #10)

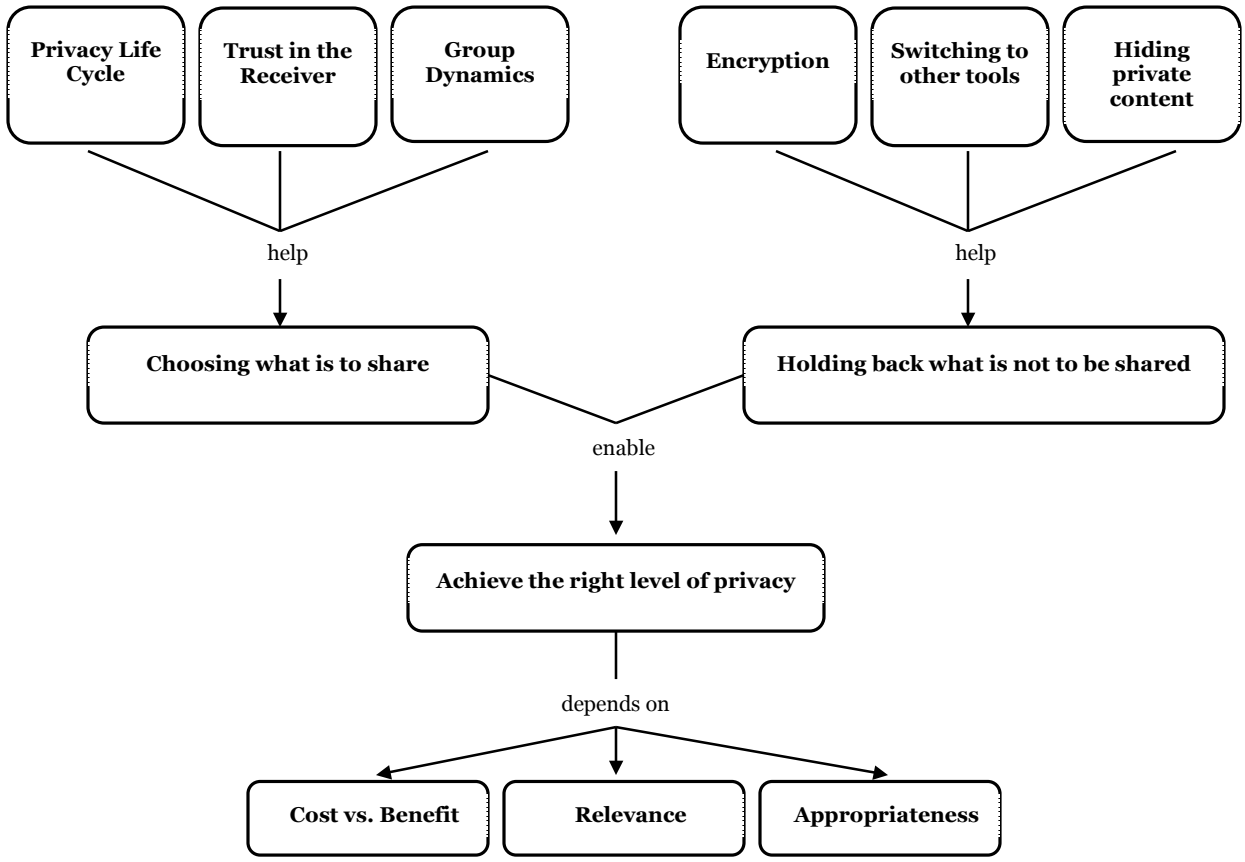
> "[my private blog] is open, but it's sort of hidden, it's not obvious how to find the page. I have not provided a link to it from anywhere. So, it's open, but it's sort of hard to find." (participant #9)

> "I use LiveJournal for stuff that I want to share only with my closest friends; for things that I consider really private, however, I wouldn't write them down anywhere online; because the easiest secret to keep is the one that is never told." (participant #12)

These strategies pointed to the fact that there are certain privacy needs of the users that the tool fails to support. 8 of our 12 participants mentioned that a better privacy management mechanism would improve their experience with Elgg.

Figure 4.2 provides a summary of the concepts described in this chapter, with an emphasis on causality.

Figure 4.2 The privacy management process in a social-personal information management system



4.8 Summary

The purpose of this study was to generate a theory of information sharing behavior of users in a social-personal information management system. The goal of the study was to discover users' main privacy concerns in this domain, and to explain the basic social

processes that users employ to resolve their concerns. The presented grounded theory shows that both the purpose and the goal of the study have been adequately met. Overall, the theory provides a relatively comprehensive answer to the research questions. The representations of the data that emerged from the grounded theory analysis provides a set of propositions for understanding information sharing behavior of the users in a social-personal information management system: the factors that shape users' perception of information privacy in such an environment, some of the challenges they face in ensuring privacy of information, and strategies they employ to achieve the desired level of privacy. Some of the concepts and relationships that emerged from data during this study (like the role of trust) support findings of other researchers. Still other elements of the theory (like the effect of information life cycle and certain group characteristics on information sharing attitude) may be considered new insight into information sharing behavior in social systems. Even for the concepts that have been studied before, their role in the particular context of social systems had not previously been explored. Another important distinction between this study and previous investigations is how it goes beyond speculation to propose explanations as to why certain factors are important: our results are grounded in data gathered from users' opinions based on their long-term experiences and as such, they give valuable insights into the *processes* entailed in information sharing in social-personal information management systems. In the next chapter, we present the process of validating the theory and discuss the limitations of our study. We also propose several design heuristics and a framework for privacy in SPIM domain based on the results of the grounded theory.

Chapter 5
Towards a Framework for Privacy in SPIMS

This chapter describes the process of moving from the findings of the study towards a framework for privacy in the SPIM domain. We start by an analysis of the validity of the study based on the four principles that have been proposed by Glaser as appropriate criteria for judging validity of a grounded theory. We also discuss the limitations of the study and present a meta-analysis of the findings. Finally, we present a framework for privacy in SPIM that we developed based on the results.

5.1 Validating the Theory

It is important to use appropriate criteria when assessing the quality of a grounded theory study. (Glaser and Strauss 1967, p. 224) observed that some of the critics of the grounded theory methodology stemmed from using "the canons of rigorous quantitative verification" as criteria for judging the credibility of "theory that is based on flexible research". The criteria for evaluating quantitative studies (i.e., validity, reliability, replicability, and generalizability) have historically evolved in the context of varificational research, with an emphasis on testing and measurement. When applied to qualitative research, however, whose purpose is generating rather than testing theory, they are highly problematic as they lose much of their relevance and significance.

(Glaser 1978, p. 4−5, 1992 p. 15, 1998 p. 236−238) has addressed the need for appropriate means for evaluating the quality of a grounded theory by suggesting that it should satisfy four essential criteria: fit, work, relevance and modifiability. These criteria engender trust because when a theory fits the data, works, has relevance to people in the area of study, and can be readily modified, it has "grab" without pressure to force it on data

(Glaser 1998, p. 237). According to Glaser, a grounded theory is never right or wrong, it just has more or less fit, relevance, workability and modifiability. In this section, we will compare the theory that was developed in this study against each of these criteria.

5.1.1 Fit

This criterion is equivalent to internal validity and represents the degree to which a study's findings correctly map the phenomenon in question. The fit criterion implies that a credible theory must fit the data it purports to represent. It suggests that data should not be forced or selected to fit pre-existing or pre-conceived theory, nor should it be discarded to keep an extant theory intact (Glaser 1978, p. 4). The fit criterion is often met in a grounded theory study by the fact that theory is generated directly from the data, and that the constant comparison method has been properly applied to the emerging concepts and relationships. As such, compliance with this criterion can be verified by reviewing the process of theory generation: if the correct procedures of theory generation have been followed (i.e., constant comparison, memo writing, theoretical sampling, sorting, and saturation), a high level of fit is expected.

We believe our theory satisfies the fit criterion as indicated by the details of the study in Chapter 4. This chapter shows how the open codes that were directly derived from users' answers to interview questions were analyzed and reassembled into categories that effectively explained how users manage the process of information sharing; including what leads to the need for selective information sharing, how it is handled, and what factors affect an information sharing decision.

Also, the semi-structured in-depth interviews that were used for data collection in this study allowed us to stay close to empirical evidence in order to ensure a close fit between the data and how our users actually managed information sharing. Through applying the constant comparison method of analysis, emerging concepts were compared with other incidents for verification, and with other concepts for establishing the best fit with the data, thus ensuring that concepts and relationships were systematically generated from the data. The numerous direct quotes from study participants that serve as exemplars of concepts and relationships in the previous chapter can also serve as evidence of such fit.

5.1.2 Relevance

Relevance is defined as the degree to which a study's findings contribute to the understanding of the phenomenon under study. This criterion is about how well the emergent concepts represent the true issues of the participants. Theoretically, grounded theory is supposed to automatically arrive at relevance because it allows core problems and processes to emerge from the data. This should lead to trust of truly getting at what is really going on in the substantive area of inquiry, which will thus have practical impact beyond academic value (Glaser 1998, p. 236).

The basic social process of selective information sharing that emerged in this study is highly relevant to users of social and personal information management systems, as it explains factors that lead to the need for selective information sharing and strategies users employ to ensure proper level of data privacy. By avoiding preconceptions and bias, we were able to generate concepts of importance and value to participants that were not identified in previous theories or similar studies (such as the frequent change in privacy preferences), thus ensuring compliance with the relevance criterion. The descriptive commentaries in Chapter 4 that provide analytic explanations of how important problems and processes emerged in the context of the study also provide evidence of relevance. Furthermore, in the next section we present several heuristics for designing a usable privacy management mechanism in SPIM that are directly derived from the grounded theory, which provides evidence that our theory has practical benefits in addition to theoretical benefits, thus ensuring compliance with the relevance criterion.

5.1.3 Workability

The Workability criterion implies that a grounded theory must be able to explain what happened, predict what will happen, and interpret what is happening in an area of substantive or formal inquiry. In other words, the theory must be able to explain how the main concerns of participants are continually resolved while accounting for all variations. By diligent application of Glaserian processes of further theoretical sampling, we were able to account for the variations in users privacy needs, concerns, and strategies with no

unexplained exceptions. We believe that the resulted theory will be able to explain the individual cases if provided with raw data or new instances.

Also, the subsequent phases of this research involve proposing a privacy framework for SPIM based on the result of the theory and instantiating it in a real world SPIM application (presented in chapters 5 and 6 respectively). Experimental data from empirical evaluation of this privacy model (as presented in chapter 7) shows that it has positive impact on users' behavior when interacting with the system: the resulted privacy management mechanism provides users with more control over privacy of their information while maintaining adequate support for usability. This shows that our grounded theory has been successful in predicting the privacy needs of users and how they should be addressed, since incorporated findings of the theory into design effectively improves users' experience, which provides evidence of compliance with the workability criterion.

5.1.4 Modifiability

Modifiability ensures that a theory is flexible and can accommodate changing circumstances and additional data. A modifiable theory can be altered when new relevant data is compared to existing data (by integrating new categories or modifying existing categories, for example), so that if the new data reveals variations in certain concepts or relationships, one is not forced to discard the theory entirely.

The theory that was developed in this study satisfies this criterion because a key element of the constant comparative method of analysis is the progressive modification and refinement of the evolving theory. Whenever new data presented variations in emerging concepts during analysis, the theory was modified to accommodate those variations. As a result, the theory is intimately linked to the data and can be expanded by incorporation of new data even if new data reveal necessary modifications. Furthermore, the main categories of the theory present a high level description of concepts that are further explained through their corresponding sub categories. This provides flexibility in incorporating modifications based on new data. For example, one can imagine that participants of different age or working in a different context (e.g., corporate environment) would probably have different privacy concerns, needs, factors, or strategies. Nevertheless, the information sharing

behavior of such group is probably still explainable in terms of these high level concepts. However, this claim can only be verified through further extensions of the study in other directions.

5.2 Study Limitations

Glaser notes that resource limitation is often the determining factor in dictating the boundaries of the domain of a grounded theory study, the extent of theoretical sampling, and in general, how far any one researcher can go with the study (Glaser 1998, p. 199). Resource limitation was certainly a factor in the case of this dissertation, as we were bounded by the constraints of both time and budget. These limitations were instrumental in confining the research to a particular group. A common exercise for enhancing the results of a grounded theory study is to follow the "site spreading" approach, which is the practice of pursuing wider theoretical sampling avenues outside of the study's original context. However, we did not attempt to pursue theoretical sampling outside our original sample pool, for two main reasons. The first reason was the practical difficulties of locating other appropriate samples for investigation due to the protective nature of most organizations towards their data, especially since the study was about privacy and information sharing habits. As explained in Chapter 3, considering our research questions, the process of locating the right sample would be considered non-trivial at best. The second reason was that the standards of practice for the engineering discipline (which was the context of this thesis) required extending the results of the study into system design, development, and validation. As such, the objective of keeping the study "unit bound" was a deliberate strategy in order to focus directly on the main purpose of the research: developing an understanding of users' information sharing behavior within a particular setting, namely social-personal information management systems, in order to propose guidelines for the design of privacy management mechanism for such systems.

It should also be emphasized that a grounded theory is not a formal theory and as such, should not be expected to provide general explanation in the area of study. Findings of the grounded theory process are an integrated set of *contextual hypotheses,* rather than universal facts. In a taxonomy of information system theories (Gregor 2002), the author categorizes

grounded theory as *explanatory theory*, a theory that is for explaining and understanding how a phenomenon happened. Explanatory theories are not formulated to enable formation of predictions or generalization beyond their area; rather, they are meant to identify a range of possibilities that can provide guides to action. Explanatory theories are a sub-category of *substantive theories*, which have much lower level of generality than formal theories and can only act as building blocks for a formal theory. (Glaser 1978, p. 153) mentions that formal theories are *extensive*, as opposed to substantive theories that are *intensive*. In order to elevate to formal theory, a substantive theory needs extensive further comparison with similar theories from diverse contexts (Guthrie 2000, p. 177). One disadvantage of this formalizing process, however, is loss of the intensiveness in favor of extensiveness of theorizing (Glaser 1978, p. 153). In simple words, by attempting to elevate a grounded theory to formal theory, the researcher runs the risk of making the theory too general to be of practical use.

Although this study is situated in the context of an educational environment, our participants used it for managing and sharing many more varieties of personal and social information in different contexts. As such, their experiences reflected diverse information sharing habits in various contexts, as evidenced by the fact that none of the privacy factors that emerged in the study (the changing nature of users' privacy preferences, the effect of trust on information sharing decision, and sharing differently in group of different dynamics) were specific to the educational context.

The unified demographics of our particular sample might raise a question as to whether the results were affected by the specific characteristics of this group. Particularly, an argument can be made as to whether the fact that this group of participants was under age had any effect on the results. While there were certain categories of artifacts that our participants would not share because of their age (e.g., none of them were allowed to post a real photo of themselves on their profile), we did not focus on this group of artifacts. Rather, we tried to stay away from these obvious cases, and focus on the more general area of sharing information they had vested interest in. As such, we believe that even though our study has well-defined boundaries in terms of the user set, types of information artifacts, the intended audiences, and the context of use, it does provide meaningful insights into users' privacy needs and strategies in SPIM domain.

The next section describes the process of translating social requirements of information sharing (as identified by the study), into technical requirements of a privacy management mechanism for SPIM systems. This is attempted through proposing a series of design heuristics that lead to a framework for usable privacy management mechanisms in social-personal information management systems.

5.3 From Grounded Theory to Design Heuristics

As an explanatory theory, the main objective of our grounded theory study was to improve understanding of a phenomenon (information sharing) and to construct an evidence-based theoretical framework describing the phenomenon, where the framework can be used to inform the design. As in other grounded theory studies, the focus of this research was not to achieve statistical validation and universal generalizations, but to discover patterns and develop theories for a better understanding of the process of information sharing from users' perspective, in order to identify guidelines that can inform the design. As such, we next move on to the description of a set of heuristics for the design of privacy management mechanism that we developed based on the results of the grounded theory.

One of the observations that followed from our study was that even though our participants were actively using Elgg's privacy management features, this didn't mean that the tool's support for privacy management was adequate or satisfactory: Tools often "seem" good enough because users are good at adapting their behavior to the tool, by employing creative strategies to "gel" what the tool offers with what they wanted to do. We believe this is because users think at task level, not conceptual level: Users are not expert enough to report what is missing or needed or even nice to have. However, the fact that our users tried to address lack of desired level of privacy with certain strategies pointed out to the fact that our users had nuanced ideas about what they wanted to share with whom and in what context, and they considered it a shortcoming of the tool when their desired level of privacy was not supported. In some cases, privacy even determined their choice of tool or their level of engagement with it. We thus suggest that a privacy management mechanism that is built

on a different conceptual model might be better at supporting users' needs. The following design heuristics are proposed to address such goal.

5.3.1 H1: Privacy Control Must be Available on a Fine-grained Basis

The first heuristic follows directly from concept 5b in the theory (need for fine-grained privacy control), and reflects the need for control of the privacy of information in terms of individual artifacts as well as their collections. Although this is a confirmation of the long-standing model that access rights should be associated with individual objects (e.g. files) and collections (e.g. folders), the higher granularity and incremental object creation model in SPIM suggests that the way in which these rights are managed to protect privacy and facilitate sharing needs to be different in some essential ways: The diverse nature of content and audiences in the SPIM domain implies that different artifacts in the same category might have different privacy requirements (i.e. landscape photos made visible to public, but family photos restricted to friends). Moreover, often times users need to grant or deny access rights other than just the read action to their artifacts (i.e. colleagues may view, but not modify), which suggests that SPIM systems need to support fine-grained privacy management not only for resources, but also for target audiences and actions.

5.3.2 H2: Privacy Preferences Must be Defined in Context

While our first heuristic calls for fine grained privacy management, the wide variety of artifacts that exist in SPIM (3a) and the fact that various artifacts in this large pool of data are targeted to various audiences at different times (3b) suggest that such attempt might in practice become too labor-intensive for the users: Research shows that while non-technical users seem to have a good idea of what their personal privacy preferences are, often times they have difficulty articulating them in terms of a set of rules (Egelman and Kumaraguru 2005). Personal preferences are also context sensitive, which makes it even harder to enumerate specific privacy rules. Enabling privacy preferences at a fine granular level makes this problem even bigger. This is evidenced by the fact that while the need for managing privacy at a fine-grained level has been recognized by other social systems as well, it has often found to pose a trade-off with usability. In Facebook, for example, users have to go through more than 50 knobs just to set privacy preferences for their various profile items

and public search visibility. Privacy-related options for individual applications are found with the application and users have to be aware of the features to find the options and visit separate privacy pages for each. Although fine-grained, the result is a completely unintuitive system where non-technical users are highly unlikely to be able to set sensible privacy preferences or understand the ramifications of their choices. Interestingly, all this effort is needed for just regulating *visibility* of one's various artifacts (i.e. the *read* action). Facebook currently does not provide any mechanism for regulating other types of action; e.g. who can edit an artifact, leave a comment, tag user in a photo, etc. To the best of our knowledge, neither do any of other existing heavily used social systems.

Thus, our second heuristic is a direct follow up to the previous one and suggests that a privacy management approach that requires users to indicate their privacy preferences to the system *a priori* (i.e., through a privacy setting page) may not work; rather, we propose that any attempt to support fine-grained privacy management must be paired with enabling users to express their privacy preferences *in context* (e.g. at the time an artifact is created or modified) when they have a better idea of whom they want to share the artifact with.

5.3.3 H3: Privacy Mechanisms Must Provide Control Over Ownership

Another deduction that followed from the grounded theory was that users have a fundamental assumption that when they put something in the tool, they should have control over its ownership as well as its visibility. In other words, users should not have to give up control over ownership of their artifacts or dissociate from managing it as a result of sharing. This followed from the observation that one reason behind users' reluctance to share information was the tool's inflexibility in providing them with the ability to control the transactional aspects of knowledge sharing activities (e.g., getting proper credit for their contribution (8c) or ensuring reciprocity (4b)). The persistent nature of information contained in SPIMS (3c) makes this problem even more important, as it creates the need for supporting issues such as reconsideration. Consequently, our third heuristic suggests that in order to build information systems that truly support personal and social information needs, they must provide a complete, persistent sense of the degree to which information that an individual creates or consumes is his/her own, the amount of control he has over the use of that information, and the ability to properly assess or exploit its value. In other words, in

addition to providing the means for users to control access rights at different degrees between the extremes of private and public, tools need to also allow users to maintain personal ownership control over their shared information.

5.3.4 H4: Privacy Mechanisms Must Support Various Group Models

The next three heuristics all deal with the concept of groups and communities in SPIM, and are based on concepts 8b (trust), and 8c (group dynamics). Our forth proposed heuristic suggests that privacy management mechanism in SPIM needs to support various group models, and is derived from the observation that both trust and group dynamics were major contributing factors to users' information sharing decisions: From a user's point of view, the primary concern in managing information sharing is in the ability to define and/or understand the audience that will have access to a particular information artifact. Generally, the choice of audience for a particular artifact or personal attribute is expressed in terms of a group of others who one trusts with that particular piece of information, so we suggest that a privacy mechanism must enable users to understandably model their trust into groups in a flexible and dynamic way. In essence, this means group definition and management mechanisms must provide the means for categorizing one's network in terms of *groups* of others that one *trusts* with various pieces of information. We thus propose that group management in social software must support various group models rather than a generic unified form, and that groups must be defined and controlled by users, rather than the system.

5.3.5 H5: Privacy Mechanisms Must Provide Control and/or Awareness Over Group Dynamics

Privacy in SPIM is also affected by the semantics of social network relations. For example, membership in a group with public membership visibility may thereby disclose interests, preferences, or other personal information regarding group members. This means that if a group member discloses information about him or groups including himself, he (whether willingly or inadvertently) also discloses information about someone else. In other words, one member's treatment of his/her privacy has a direct effect on another member's privacy. This suggests that awareness of group dynamics is an essential need for a privacy management system; meaning, such dynamics must be both controlled by and clearly

articulated to users. We believe that addressing such need will also contribute to alleviating the privacy concerns (4a, 4b, 4c) that emerged from the theory.

Also, our grounded theory clearly delineates that the trust one has in a particular group with which one might share information depends critically on the model by which the membership in that group may change over time. We thus propose that in addition to providing support for the definition and manipulation of various group dynamics, tools must also provide means for controlling changes in those aspects.

5.3.6 H6: Privacy Mechanisms Must Allow Definition of Groups that Reflect Interpersonal Relationships

This is a follow up to the previous two heuristics, and suggests that one group model that must be supported by SPIM is the egocentric group that is defined based on users' interpersonal relationships. This follows from the observation that in SPIM domain, one's personal and social information are not always shared with identifiable, accountable individuals or groups, and sharing may happen in a variety of contexts, for example competitive as well as collaborative. Moreover, people may act simultaneously in several contexts, holding multiple potentially conflicting relationships simultaneously. As such, a lot of users' information sharing needs is better described in terms of the relationship that exists between the owner of the artifact and the person or group with whom the information may be shared, specially since new intricacies have blurred the boundaries between public and private (Boyd (2006) for example, points out that US teenagers feel strongly about preserving a certain form of privacy: they want to be visible and searchable for their friends but not their parents). In terms of rights management, these observations strongly imply that the potential audience for some artifacts or attributes is likely defined in the user's own terms, based on a variety of kinds of relationships that more closely resemble real-life privacy boundaries (e.g., one-sided and short-term relationships) and not in terms of any organizational "roles" or groups. Thus a privacy system must enable users to control the release of their personal information in the same manner they would control it in the real world, based on their relationship with the data receiver, rather than the receiver's organizational role.

5.3.7 H7: Privacy Mechanisms Must Easily Accommodate Changes in Preferences

The next heuristic is derived from a combination of the theory concepts, including 8a (change of preference), 8b (dynamic nature of trust), and 4c (impropriety). What these concepts have in common is that they all emphasize the dynamic nature of users' privacy preferences in the SPIM domain. In general, any act of information sharing can be defined as "a user sharing an artifact with a receiver based on their relationship". In the SPIM domain, however, the information sharing act is often about establishing and maintaining a dynamic sharing relationship: Over time, the nature and state of personal artifacts might change (i.e. research results getting published, patented ideas getting approval, personal opinions reconsidered), the receiving group with whom the information is shared might change (i.e. competitors joining a group or collaborators leaving), and the relationship between the owner and the receivers of information might change (i.e. people switching to a different project groups or changing affiliations). In short, all the contributing factors to a user's information sharing decision can change over time, and all the three privacy factors that were identified by our study reflect users' needs for support for privacy management when these kinds of change happen: The privacy life cycle factor emphasizes the effect of the dynamic nature of the artifact; the trust factor reflects the effect of change in the relationship between the user and the receiver; and the group factor shows how users try to deal with these changes by organizing their network into various groups as a way of compartmentalizing trust and audience, rather than having to deal with it on an individual basis.

We thus propose that a privacy model that statically assigns access rights based on these factors at the time of an artifact's creation or modification will be insufficient. Rather, privacy mechanisms need to be flexible enough to accommodate frequent changes in users' privacy preferences in a non labor-intensive way.

5.3.8 H8: Action Possibilities and Their Consequences Must be Clearly Presented to Users

Our last design heuristic emphasizes the importance of interface clarity and is derived from the usability problems that were observed during the study: our data confirmed the intuition that users can be reluctant to share personal information when they are not sure how exactly to do things, or when the consequences of a sharing decision are unclear. A counterintuitive consequence of this was that some users were more willing to share personal information in a space that afforded virtually no privacy control (e.g. blogs or Myspace pages) than one which offered them an unclear set of privacy management tools. In our study, users were made aware that they could have some control of privacy and should manage the audience for their personal information by the promise of an access control system. However, many found it inadequate because either they could not perceive how to do something they wanted to do (i.e., users didn't know they could make something visible only to one person, even though such functionality was supported by the tool), or they were not sure what the consequences of a sharing decision were (i.e., even though the concepts of groups and communities supported different information sharing models, users were not clear on how they differed). As a result, they were not able to take advantage of certain aspects of the privacy management mechanism, because of the inability of the tool to convey their existence or consequences.

This observation is inline with the findings of existing HCI literature on the importance of perceived affordances in the design of a mechanism to support end-users. As defined in the HCI field, perceived affordances are "action possibilities which are readily perceivable by an actor (Norman 1999)". Simply put, the concept emphasizes that suggested interactions with a tool/interface must be in accordance with the ability of the users to perceive those interactions. Compatibility with user's understanding and expectations of interaction with a tool/interface is known to be a contributing factor in improving its perceived affordances (Hartson 2003, Norman 1999). Additionally, perceived affordance has been identified as a major contributor to enhancing usability of a design (Norman 1988, McGrenere and Ho 2002), which emphasizes the importance of the compatibility of the privacy model with users' mental model. We will revisit this issue in chapter 7.

5.4 A Framework for Privacy in SPIMS

The overall goal of the proposed design heuristics was to identify a minimal set of technical requirements for privacy management in social-personal information management systems. While heuristics H1 through H7 describe *what* kinds of privacy control are necessary for managing and sharing personal artifacts, H8 pertains to *how* these controls must be built and incorporated in order to be usable. We now consolidate these requirements into a framework for user-centered privacy in SPIM domain that describes privacy in terms of required controls over artifacts, audiences, relationships, and change; with an emphasis on the clear presentation of those controls to users.

5.4.1 Artifact Control

The principle of artifact control is a consolidation of H1, H2, and H3, and essentially reflects the finding that privacy management in social software must be defined on a per-artifact level as opposed to per category; and that the access rights need to be applied in-context (meaning, at the time of artifact creation or modification) as opposed to a-priori (through a privacy setting page). Furthermore, in addition to control over visibility (the read action), users also need the ability to control other rights over their artifacts, e.g., modification or deletion (the write action), and over further delegation of such rights.

5.4.2 Audience Control

The principle of audience control is a consolidation of H4 and H5, and reflects the need to restrict both the visibility and ownership of artifacts to certain user-defined groups. Although most existing social systems support some group functionality, we suggest that social software systems must provide the means not only for creation of these user-defined groups, but also for definition and control over various aspects of these groups (sizes, membership models, and visibility), and for controlling changes in those aspects. Furthermore, these controls need to be in the hands of users, rather than pre-defined by the system.

Audience control is largely concerned with the issue of identity construction and maintenance (Boyd 2006). The greater risk of exposing identity attributes to a worldwide

audience in open online environments results in the need for deployment of a great deal of audience control for one's personal information artifacts. We see some of this control currently being expressed in certain social software systems, notably the social bookmarking system Del.icio.us[2], and the social networking systems Orkut[3] and Facebook. Del.icio.us was originally completely open (i.e. anyone could see anyone else's complete set of bookmarks), but due to user demand and competition from other social bookmarking services (notable Magnolia[4] and Bluedot[5]) it added the ability to mark bookmarks as "private" in the Spring of 2006. A private bookmark in del.icio.us is essentially invisible to anyone else but the user himself. Another simple example of audience control is the case of contact management, in which users selectively choose which of a variety of different categories of "friends" will be allowed to contact them in a particular way (e.g. who do I give my phone number, address, or AIM id to?). Without aid of technology, we either publish them for all to see or hand them out individually or in particular contexts (e.g. a prof might be willing to give his cell phone number to students he teaches, but not other students).

5.4.3 Relationship Control

The principle of relationship control is a reiteration of H6, and reflects the need for the ability to define information sharing based on a user's self-defined relationships with others. In essence, this emphasizes that users need to control their information sharing in the online world in the same manner as in the real world: based on the relationship between the user and the person or group with whom the information is to be shared, and not in terms of externally imposed constraints such as the receiver's organizational role. In other words, users need the ability to define groups of friends or collaborators in their own terms, and to use this model of their relationships with others as the basis for audience control.

Again, we look at Orkut and Facebook for examples. In Orkut, a user is able to define an audience for identity attributes in terms of his/her self-designated "friends" and a limited transitivity of that friendship network (e.g. I'll let my friends and any friends of my friends see my phone number), as well as arbitrary groups created by users to group their

[2] http://del.icio.us
[3] http://www.orkut.com
[4] http://ma.gnolia.com/
[5] http://bluedot.us/

friends. In Facebook, the relationship categories are much finer and reflect a variety of different kinds of relationship (e.g. we worked together on a project, we "hooked up"). The consequences are similar, however, in that I can then choose to allow access to particular posts or personal attributes based on these relationships, but without the transitivity of the Orkut model. Of course, Facebook also has more traditional groups that are formed by users explicitly joining them as well, but the audience for user attributes and personal posts is controlled completely in terms of the ***relationship control*** that the system allows.

Relationship control is clearly a manifestation of the need to define trust in terms of *egocentric* groups of users. In essence, it is very likely that the best match for the assignment of visibility and ownership rights to an artifact might sometimes be these relationship groups rather than the traditional groups. Given that an egocentric relationship model naturally aligns with patterns of trust and information sharing for personal information, it is essential that visibility and ownership rights be assignable based on these user-controlled relationship models.

Some of the existing social software systems have attempted to provide users with the ability to categorize their social network in terms of their relationships. In Facebook, for example, relationships are classified with a set of standard assertions represented by the dialogue in Figure 5.1.[6] However, we believe that there are two problems with this model: 1) the categories are clearly incomplete (e.g. how do I indicate that I "taught" a student in a particular course?), and 2) I can't designate that individual photos, notes, etc. are to be shared with only a subset of these relationship groups. We thus suggest that the fixed, traditional, application-defined *taxonomies* may not be the best way of categorizing one's social network in terms of relationships. Rather, we propose that the *folksonomic* model of information organization (tagging) be adopted for this purpose.

[6] This screen shots reflects the state of relationship groups in Facebook as off April 2006. Since then, Facebook development team has decided to remove such classification of one's friends network. Instead, users are now able to define "lists", which are arbitrary user-created groups used for categorizing one's social network into subgroups based on a user's own terms. However, in current implementation of Facebook user-defined lists are only used for selective reception of news feeds, as opposed to selective sharing of information, which is the focus of our work. As it stands, it is still not possible in Facebook to share an information artifact with a selected *user-defined* audience.

Figure 5.1 Facebook friend categories

The folksonomic information organization model allows a user to associate a set of personal keywords (tags) with a particular piece of information (an artifact). Each such keyword then automatically becomes a category term that can be used to select collections of artifacts for recall or comparison, using both individual keywords and certain Boolean combinations of these collections (as sets). Since this model was first introduced by Del.icio.us and the photo exchange system Flickr[7], it has been adopted in a wide variety of uses (e.g. blogging and RSS syndication), has become widespread in its exposure to the Internet community, and has been the subject of a body of research. While much of this research has been focused on the social aspects of the model, our interest is primarily on its usefulness as a model for organizing information for completely *egocentric* purposes. We will further expand on this topic in the next chapter when we discuss our implementation of the concept of relationship control in an experimental SPIM application.

5.4.4 Change Control

The principle of change control is a reiteration of H7, and is something of a cross-cutting concern within the other three controls. This principle reflects the observation that in the SPIM domain, the artifacts, the audiences, and the relationships used to define privacy and sharing patterns are all dynamic. A privacy and user interaction model must thus take into account that artifact life cycle and categorizations will change, that a user's requirements to share classes of artifacts with certain audiences will change, and that a user's relationships

[7] http://www.flickr.com

and trust patterns within those relationships will change, and that users come to expect their tools to provide flexible support for these changes in their privacy preferences when the social parameters that define the sharing model change.

Change control is about the need to fluidly manage the evolution of privacy over time, either at the artifact or relationship level. It contrasts with the demonstrated inadequacy of the set-on-creation privacy management models for static artifacts when used in the social software domain. In other words, any adequate access rights management system must track changes in users' privacy preferences and apply them dynamically and visibly.

5.4.5 Clarity

While the other heuristics focus on what kinds of control of privacy are needed, the last one focuses on how those controls must be presented to users in order to be usable. As such, it must be considered in parallel with the other four controls as presented in figure 5.2. We use the term clarity to represent this heuristic; meaning any functionality to incorporate artifact, audience, relationship, or change control must be compatible with users' mental model, to ensure that in practice, the average, non-technical users would be able to take advantage of the extra control over privacy that these user-centered controls are supposed to provide.

Figure 5.2 provides an overview of the elements of the framework. Table 5.1 shows a summary of how theory concepts, design heuristics, and various privacy controls in the framework correspond to each other.

Figure 5.2 A framework for usable privacy control in SPIMS

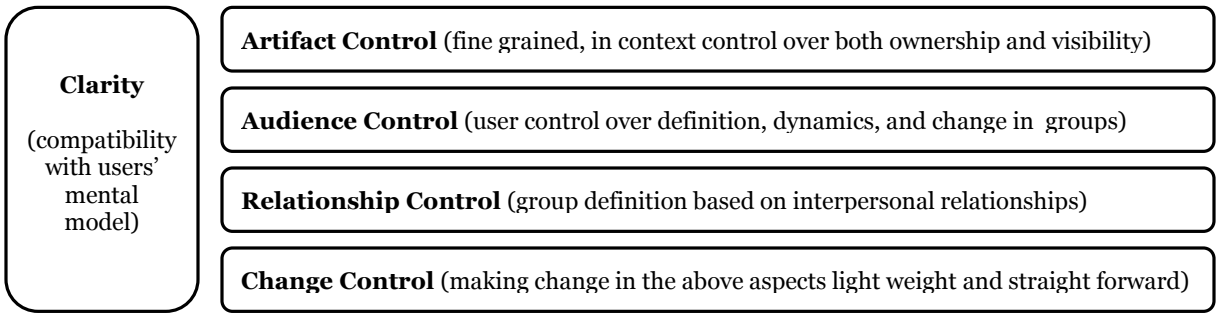


Table 5.1 The correspondence between theory, heuristics, and the framework

Theory Concept	Corresponding Heuristic	Corresponding Control
5b	H1	Artifact Control
5b, 3a, 3b	H2	
4b, 8c, 3c	H3	
8b, 8c	H4	Audience Control
8b, 8c, 4	H5	
8b, 8c	H6	Relationship Control
8a, 8b, 4c	H7	Change Control
Various usability problems	H8	Clarity

5.5 Summary

This chapter provided validation of the grounded theory study that was conducted as the foundational research for this work. We evaluated our study against the four criteria for measuring validity of a grounded theory, and provided evidence for meeting each of the described criteria. We also discussed the limitations of the study and the rationale for keeping the study unit bound.

Based on the results of the grounded theory, we then proposed eight design heuristics for designing privacy management mechanisms in SPIM. The heuristics included hints on both the kinds of control of privacy that would be required in various dimensions, and how those controls must be presented to users to ensure usability. Finally, we consolidated these heuristics into a framework for designing privacy management mechanisms for SPIM.

Use of grounded theory methodology allowed us to devise a privacy framework for SPIM that is grounded in users' understanding and perception. To show that the model is of practical benefit (e.g., it improves users experience when interacting with the system by yielding a usable privacy management mechanism), our proposed framework needs to be systematically tested under experimental conditions. For that, we implemented our proposed

framework in an experimental SPIM tool and performed empirical evaluation through user studies. The next chapter explains the design and implementation of our test bed, OpnTag. The evaluation study and its results are presented in chapter 7.

Chapter 6
OpnTag

In order to illustrate the suitability of our proposed privacy framework for design, our next step was to design a system that instantiates the framework, in order to provide an environment in which we can test these principles in action. This chapter describes OpnTag, an experimental SPIM system developed by our group for which we have built a privacy management mechanism based on the four user-oriented privacy controls as described in the previous chapter. We present the technical structure of OpnTag, along with a discussion of how our framework as embodied in this system supports each of the four user-oriented privacy controls.

6.1 OpnTag's Conceptual Model

OpnTag[8] is an open source web application for note taking and bookmarking that we developed to facilitate creation, organization and sharing of information and knowledge for an individual operating in various social contexts. The fundamental unit of information storage in OpnTag is the *memo*, a tagged textual annotation that may optionally link to a web resource. Users create memos to save notes or bookmark URLs, browse and tag other users' shared memos to mark their interest in them, and reply to other users' memos to create a conversation. OpnTag utilizes tags as a lightweight and flexible way to organize, contextualize, and represent memos. Tags are descriptive terms, keywords, category names, or metadata that can be assigned to memos to describe them. They provide meaning, context, and categorization and can be used to support search, representation and organization. Another important component of Opntag are *groups*. Opntag allows its users to create groups and invite other to join these groups in order to define various communities to collaboratively create and manage information and knowledge. Fundamental to this

[8] sourceforge.net/projects/**opntag**

implementation is providing functionality to restrict the visibility of one's memos to a certain group, including the *private* group consisting only of oneself (Figure 6.1). In the following sections, each of the key concepts in the design of OpnTag's is described in detail.

Figure 6.1 Memos in OpnTag: public, private, and selectively shared in a group

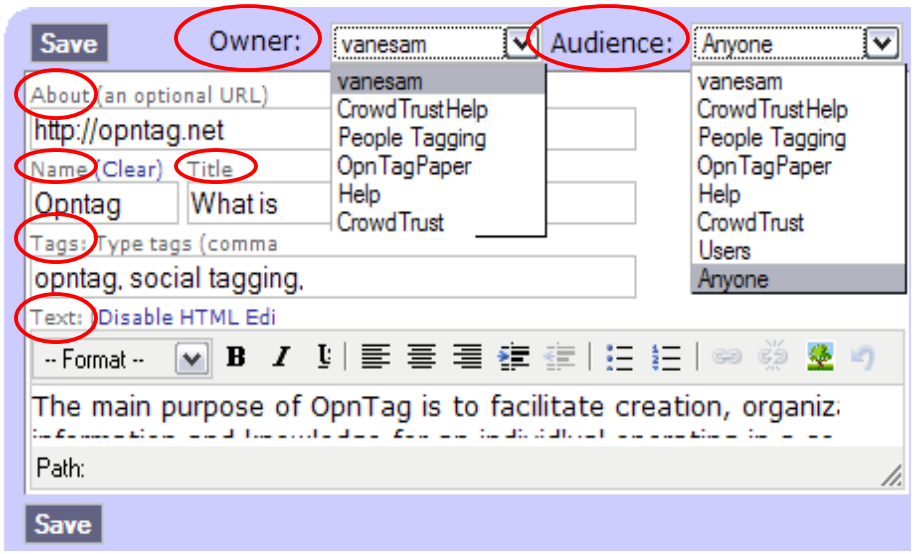
6.1.1 Memo

Memo is the basic unit of information in Opntag. It has a *name* or *title*, an optional *link* (URL) specifying what it is "about", a set of *tags*, and optionally, some *text* representing its content. Memos can function as bookmarks, notes, or wiki pages and are organized based on their intrinsic metadata (e.g. who owns or created them and when) and tags applied to them by various users. Memos have globally unique system-assigned IDs and may have a user-assigned name that is unique among all memos owned by the same user or group. This unique name can be used to refer to that memo in a more meaningful way than the ID, either when linking from another memo (using a "named reference" shorthand), or when providing a URL. For example, to refer to a named memo within OpnTag, one only needs to provide a reference to the memo's owner and memo's Name (e.g. a memo named "Privacy" created by Maryam can be referred to as [[Maryam:Privacy]], which will create a hyperlink to that named memo). This gives memos a Wiki-like functionality, as pages can be referred to by name. Like a Wiki page, a memo does not have to exist to be referred to;

following a memo's name link actually opens a dialogue to create the memo if it does not exist. With this in place, OpnTag enables its users to easily create both individual notes and bookmarks and networks of cross-referenced information units.

Each memo has an *owner*, which controls who owns the memo and thus can edit and delete it, and a potentially restricted *audience,* which controls who can see that the memo exists and read it (in OpnTag, visibility implies readability, so there is no "I can see that it exists but can't read it" issue). Both the owner and the audience can either be set as an individual or defined as a group (Figure 6.2).

Figure 6.2 Various elements of a memo in OpnTag. Owner and audience lists include both the individual (vanesam) and her groups



6.1.2 Users, Spaces, and TagClouds

In OpnTag, each memo is either acquired or created by a *user* and exists within a *space*. A user is an individual who has an account with OpnTag. Each user in OpnTag is provided with a *personal space*, a workspace within which s/he has complete control over how to organize, represent, and share memos. In addition to personal information spaces, OpnTag also allows users to have access to one or more *group spaces,* by the virtue of being a member of those groups.

A user's personal space contains all memos created or edited by the user and a set of tags that s/he has assigned to these memos, which is called *user's tag cloud* (Figure 6.4). A group's space on the other hand, includes all the memos that are either crated in that space or have been specifically made visible to that group, along with the *group's tag cloud*, a collection of tags associated with these memos (Figure 6.3). While within a user's personal space only the user himself can create, edit, and organize the body of information consisting of memos and associated tags, in a group space any member of a group may do so. As such, personal spaces support *personal* information management in OpnTag, while group spaces are there to provide *social* information management functionality.

Figure 6.3 An snapshot of a personal space in OpnTag, with user's tag cloud

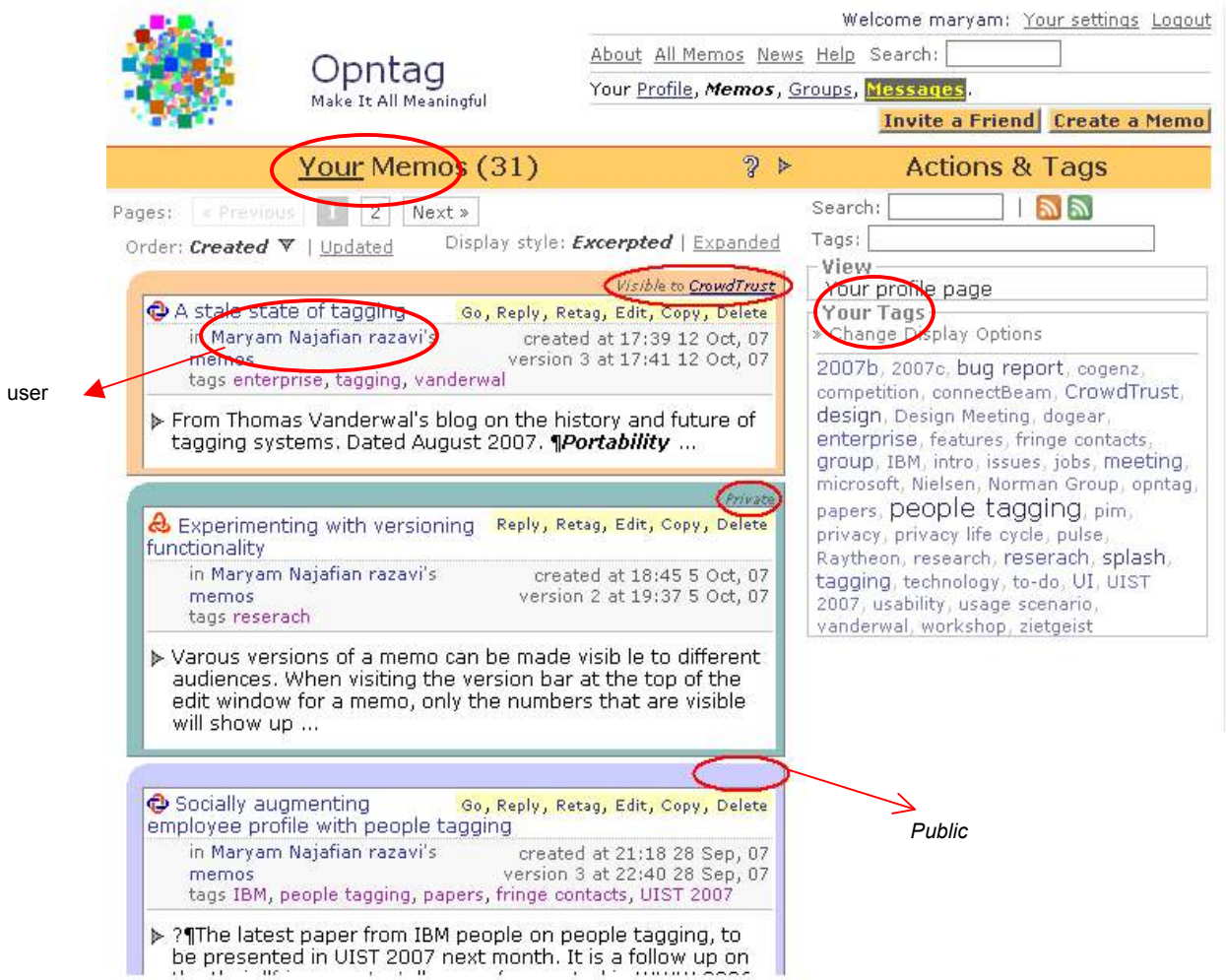
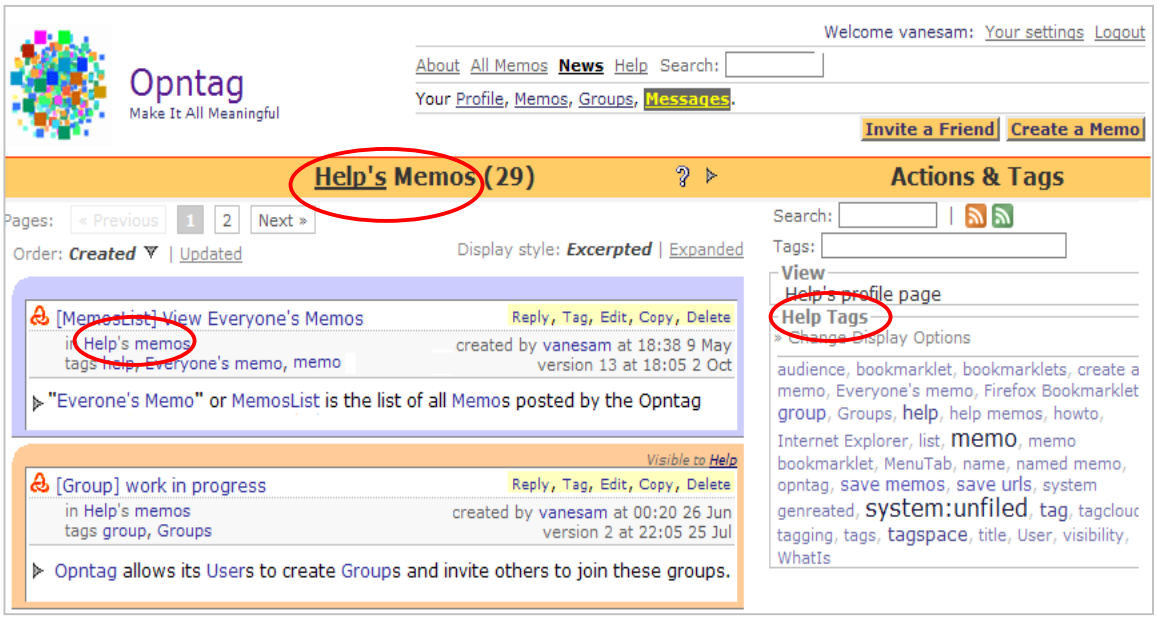


Figure 6.4 An snapshot of a group space (Help) in OpnTag, with group's tag cloud



6.1.3 Navigation & Grouping

As in other tag-based systems, objects in OpnTag are grouped based on ownership and tagging. After logging in, users are initially transferred to a page where they see a list of all memos in the system (depending on memos' various visibility restrictions and users' various group memberships, different users will see different sets of memos when they first login to the system). From there, they can select subsets by filtering based on an ownership space (e.g. all maryam's memos), a tag or set of tags (e.g. all memos tagged "research" and "privacy"), or some combination of those (e.g. all maryam's memos tagged "research" and "privacy"). Thus these attributes of a memo (both the intrinsic metadata and user-supplied tags) are used both for identifying and grouping of memos.

6.2 Groups in OpnTag

A fundamental goal of OpnTag is to provide selective information sharing, which is supported through creation and management of groups and providing functionality to restrict the visibility of one's memos to a certain group. OpnTag supports two different types of groups: *classic groups* and *egocentric groups*. Next, each is explained in detail.

6.2.1 Classic Groups

The primary function of a classic groups is to allow a set of people with a shared interest to create a context for selectively sharing personal information and a collective space within which they can actively collaborate to create, edit and organize information either publicly or in private. Because classic groups have their own group space, including a view of entries created by or assigned to the group and their own tag list, a classic group can be a very convenient way for a set of people to get a more focused view of their data than by searching or browsing through all the memos presented on OpnTag's main page. Two classic groups may have a subgroup relation, where one group is made a member of another, or a nested set relation, where all members of the enclosed group become members of the enclosing group. A number of special classic groups exist: "Users" which includes all individuals registered with the OpnTag instance, "Unknown" which includes the anonymous, unregistered user, and "Anyone" which includes both groups and thus represents truly public access.

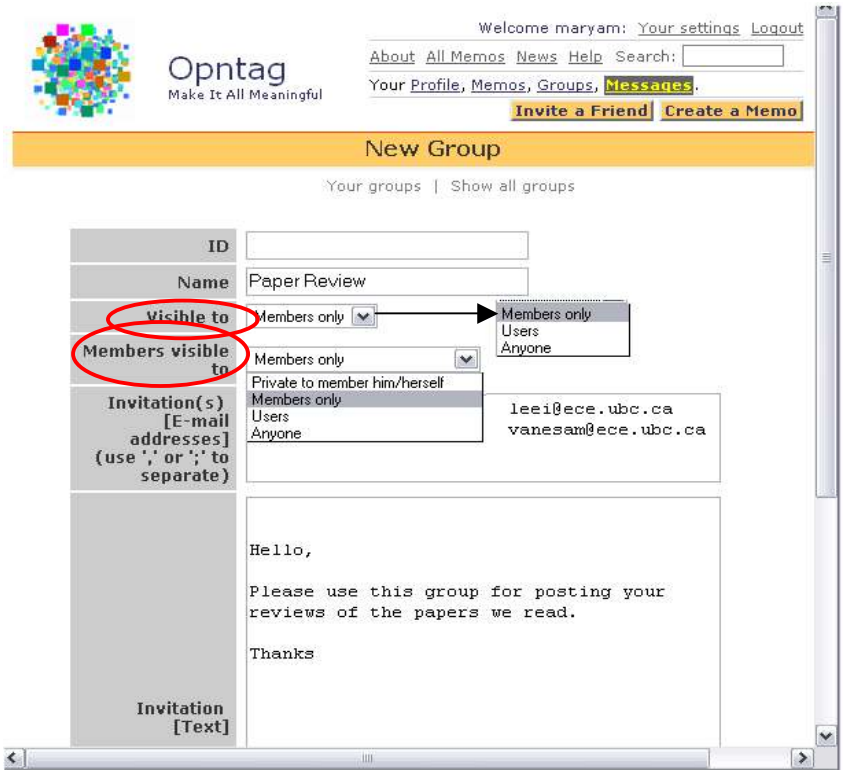
Membership: Membership in a classic group is by invitation: each member of the group may invite as many people to the group as s/he wants by sending an invitation to their email address. Accepting the invitation happens on a voluntary basis. Users can choose to be members of as many such groups as they want, and can create as many groups as they want.

Visibility: For each group created, the creator specifies the group's visibility (one of "Members Only", "Users", or "Anyone"), and the visibility of the member list (same options as group visibility plus "Private", meaning no one would know of user's membership in the group except for the user himself). The visibility of the memos, tags and member list of a group is then restricted by the visibility of the group itself (e.g. it is not possible to make a group visible only to its members, but make its member list visible to the public). By using various combinations of group and members list visibility, users can create groups with different dynamics and then restrict the visibility of their memos to any of these groups, including the "private" group consisting only of oneself (thus making the memo completely

private), or the "Anyone" group, which is the super-groups of all others (thus making the memo completely public).

Administration: Currently, all members of a classic group have equal administrative rights, which include creating subgroups, inviting new members, tagging within the group, and editing, deleting and changing the visibility of any group-owned memo. The only exception is that only the group creator can destroy it, and only on the condition that there are no memos in the group space.

Figure 6.5 Group definition page with group and member list visibility menus



6.2.2 Egocentric Groups

In addition to classic groups, OpnTag supports a different type of group called egocentric group, which enable users to define dynamic, non-reciprocal relationship groups to which they can grant or deny access to various pieces of information in their personal space. Egocentric groups primarily provide support for relationship management and are

handled by *tagging people* through their profile pages. We next describe the general idea and motivation behind people-tagging and explain how the idea is implemented in OpnTag.

6.2.2.1 Motivation

In chapter 5 we discussed that one of the challenges in categorizing target audiences for various personal artifacts in social systems is supporting the subtle and nuanced ways in which users' patterns of trust change over time and the ways in which this interacts with their transactional approach to information sharing and exchange. We also discussed how users view sharing information as establishing and maintaining a *dynamic* sharing relationship, and how this calls for flexible and lightweight mechanisms to enable users to categorize their social network in terms of their (often changing) relationships with others.

We propose that the same folksonomic organization model that works for information categorization can be adapted for organizing relationships, and that the tagging model has certain characteristics that make it a suitable candidate for facilitating relationship management for selective information sharing:

- Many tags (and thus relationships) can be associated with each person
- The choice of tags is entirely in the control of the tagger
- The act of tagging is simple, intuitive, and well-adapted to granular relationships, and
- The collections created by coincidental tagging (i.e. all people tagged with the same word) form natural categories.

We believe that these characteristics make people-tagging a possible approach for relationship management. Our idea is to allow users to use tags on other users and then treat each tagged category of people as a relationship group that can be used as an access control feature for each of the tagging user's memos. We also believe that the visibility of these people-tag groups should be controlled, i.e., one should be able to designate who should be able to see the people s/he has tagged with certain keywords, thus making it reasonably safe to "opinion-tag" others. Tagging people can be used for both signaling assessment of others (i.e. tagging someone as "gifted") and signaling relationships (i.e. tagging someone as "colleague"). Each tagger might assign multiple tags to the same taggee, and multiple taggers

can tag a taggee with the same tag, but only one instance of each tag for each taggee is allowed. We also differentiate between one's *Incoming* and *outgoing* tags: incoming tags are those assigned to the user, while outgoing tags are those the user has assigned to others.

6.2.2.2 People-Tagging in the Literature

While use of social tagging for information classification and annotation has been the topic of much recent research, the concept of tagging *people* (as opposed to tagging web resources) has received relatively little attention in comparison. The two most notable works in this area are Fringe Contacts project from IBM (Farrell et. al 2006, Farrell et. al 2007a) and the Tagalag[9] project.

In Fringe, social tagging of people is used to support contact management and augmenting employee profiles in an enterprise directory. The application puts the same emphasis on tagging *self* as on tagging *others,* which enables employees to place themselves and their fellow employees in a folksonomy of skills, interests and projects. Since the application is targeted to the enterprise environment where tags are traceable to people's identity, only public tags are supported. Tagalag is another project which uses people-tagging for the purpose of self-promotion and finding others with similar interests. Tagalag enables users to tag others based on their email address and can be integrated with web based email systems.

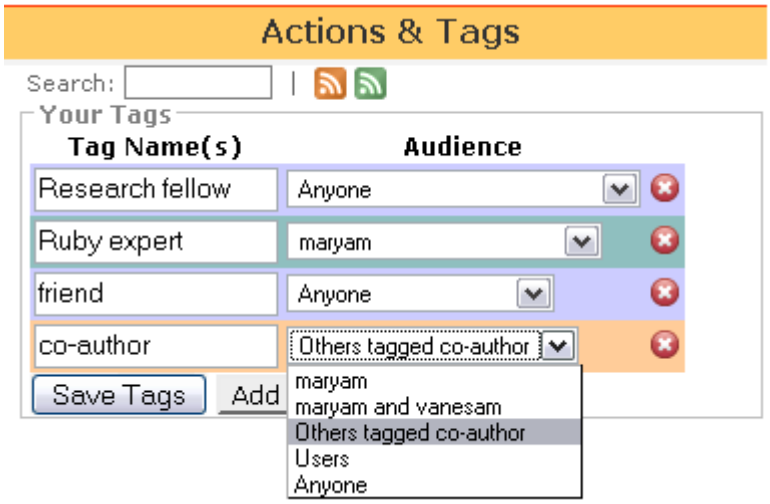
Prior research has also indicated that users use people-tagging to define and socially manage communities. A recent study (Farrell et. al 2007b) showed that users frequently tagged people for the benefit of others, shared links with tag-based groups of people, and used tags to construct ad hoc mailing lists. To the best of our knowledge, our work is the first to consider tagging people for relationship management and by extension, for privacy management. Also, private or restrictedly shared people tags have not previously been addressed in the literature.

[9] http://www.tagalag.com

6.2.2.3 People-Tagging in OpnTag

Figure 6.6 shows how the concept of people-tagging is implemented in OpnTag. When visiting another user's profile page, a user can tag the profile owner with keywords that represent his/her perception of that user (i.e. "ruby expert") or their relationship ("research fellow"). In the same way that tagging resources both identifies and groups them (e.g., all memos tagged "rails" can be treated as a group), each such *people-tag* represents a relationship group that will later appear as an access control option for each of the tagging user's memos (Figure 6.9).

Figure 6.6 Tagging a user profile in OpnTag. Each tag becomes an egocentric group with specifiable visibility



Each new tag applied to a person has a distinctly specifiable visibility, assigned by choosing one of the options in the "Audience" drop-down menu (Figure 6.6). The choices for people tag visibility include only the tagger ("maryam" in the Figure 6.6), the tagger and the taggee ("maryam" and "vanesam"), only the set of people tagged with the same tag (by the same tagger), "Users", and "Anyone", with the default visibility set as private (tagger only). Significantly, a single tag may have different visibility to different taggees (e.g., I may not want to share my assessment of others as "interesting" with everyone I assess as such). However, in no case can the tagger make a tag visible to anyone other than the taggee without also making it visible to the taggee himself. Since all such tags are visibly attributed to the tag creator, this design choice was made to discourage antisocial tagging by forcing

such taggings to be exposed to their subjects (e.g., I can't let my friends know that I've tagged someone as a "jerk" without letting the "jerk" know too).

Unlike classic groups in which membership is voluntary, people-tags are assigned and removed by the tagger without a confirmation process for the taggee (and possibly without even notification). As such, the relationship groups that are created as a result of people-tagging are entirely controlled by the creator; meaning people do not need to agree to be in the group, and they may not even know that they are included in a certain relationship group. An important implication of users being able to assign their acquaintances to different relationship groups (potentially without their knowledge or approval) is the opportunity for handling some social situations that are generally hard to handle in online world. One example of such situation is discretely concealing exclusions; i.e., I may want my "friends" (i.e. those others I have tagged with "friends") to see that they are included and have special privileges to my information store as a result, but non-friends should not be visibly excluded. In OpnTag this situation can be handled by making the "friends" tag visible only to the people tagged as such.

OpnTag also supports the concept of social scoping, in the way that the same tag applied by different users will result in two distinct relationship groups. For example, people tagged "friends" by user A will comprise a different relationship group than people tagged "friends" by user B. Furthermore, both of these relationship groups will be independent of the classic group "friends", if either A or B are a member of such group.

Figure 6.7 shows the incoming and outgoing tag clouds on a user's profile page, typographically modulated based on the tag frequency. Different color codes imply various degrees of visibility (color codes have been chosen in consistence with memo color codes, with green for private, blue for public, and orange for group-shared). It is possible to pivot on the incoming or outgoing tags, both individually and collectively, to have various filtered views; i.e., all people who have used a certain tag (or a combination of tags) on this user, all people this user has tagged with a certain tag (or a combination of tags), all tags applied by a certain user to this user, and all tags this user has applied to a certain user.

Figure 6.7 Incoming & outgoing tag clouds in OpnTag

6.3 Privacy Management in OpnTag

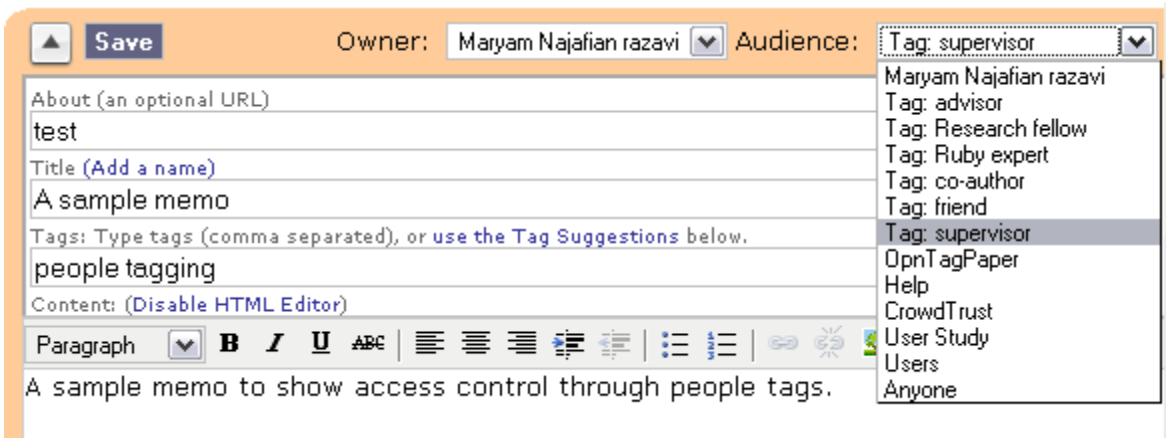
With individuals, classic groups, and egocentric groups, OpnTag's privacy control centers on the joint concepts of ownership and audience management, which together, ensure that individual users retain both control and credit over the artifacts they dispose in the system. Although a group can be specified as the designated owner of a memo, each memo is also visibly attributed to its individual creator, thus ensuring that each group member gets proper credit for the contributions s/he makes to the group's shared information space (Figure 6.8). Only a memo's creator can modify ownership, but any member of the owning group can change a memo's audience. This design choice allows the creator of a memo to decide whether access restrictions of a memo should be controlled

collectively or individually. The latter case supports situations when there is a need to ensure that the access restriction policies *stay* with the shared data; i.e. it is not possible to make a memo visible beyond the intended audience set by its creator.

Figure 6.8 Recognition of individual contribution in the group space

Audience restriction is the fundamental mechanism for selective sharing in OpnTag: at the time of creating or editing a memo, the creator has access to both his classic group memberships and his relationship tags and can thus adjust the audience of the memo to be either the owner himself, a classic group that the owner is a member of (with its collective membership dynamics), or one of the owner's egocentric groups (over which he has complete control). Figure 6.9 provides a screenshot that shows how this choice is made.

Figure 6.9 Selective sharing of a memo in OpnTag, with both classic and egocentric groups



6.4 Supporting the Four User-Centered Privacy Controls

A fundamental goal in the design and development of OpnTag was to instantiate the privacy framework that resulted from our research, in order to create a test bed for its

empirical evaluation. In this section, we discuss how each of the four user-centered privacy controls in our framework is supported by OpnTag's privacy management mechanism.

6.4.1 Supporting Artifact control

Artifact control in OpnTag is supported by providing ownership and audience management at the level of individual memos. For each memo, the user specifies the memo's owner with the default owner of a memo set to the creator (if in individual space) or the group (if in group's shared space). By enabling context-sensitive ownership management at the memo level, OpnTag's privacy system supports definition of *fine-grained* privacy policies *in context* (at the time of artifact creation or modification), as opposed to *a priori* (through a privacy setting page for defining general privacy policies on collections as most current tools do). In addition to fulfilling the fine-grained privacy management requirement, this also allows users to define their own privacy policies when they have a rather clear idea of their sharing preferences. Moreover, with OpnTag's ownership management, privacy policies stay with the data, since no one other than the memo's owner can change it's audience; i.e., it is not possible for someone who is not an owner of a memo to make it visible beyond the intended audience set by its creator.

6.4.2 Supporting Audience control

Audience control in OpnTag is supported through deep visibility management via user-defined groups and relationships. Every memo in Opntag has restricted visibility; meaning, it can only be seen by members of a designated group. OpnTag enables users to control various aspects of the classic groups they define, including group visibility and member list visibility, and also supports definition of egocentric groups which have totally different dynamics in terms of visibility, membership model, and reciprocity. By enabling these variations, we were hoping to accommodate variation in users' trust and sharing behaviors that were shown by our study to depend on the visibility and dynamics of the sharing context. We believe that by supporting groups of varying dynamics and by conveying those variations to users in a clear way, we enable users to reliably predict and effectively control the potential audiences for their various memos. For example, one would be less likely to share sensitive information with an "open" group that anyone could join than in a

"closed" group where new members must be invited. In addition, through people tagging, users can take complete control of the audiences for their artifacts.

6.4.3 Supporting Relationship control

Relationship control in OpnTag is supported through people-tagging, which enables categorizing one's network into user-defined, egocentric groups that may represent various kinds of relationships. Since egocentric groups are controlled entirely by the creator (the acts of creating or deleting a tag on a user, and controlling the visibility of the tag are solely controlled by the tagger), the act of tagging a person via their profile page is equivalent to asserting their membership in a group whose membership is entirely under tagger's control. As such, people-tagging provides a lightweight and flexible mechanism for handling volatile relationships that frequently show up and fade out in natural social environments, but are often hard to manage in online world.

6.4.4 Supporting Change Control

The combination of artifact, audience, and relationship control in OpnTag allows fine categorization of resources, audiences, and actions, provides advanced group functionality, and enables users to define privacy preferences based on their often changing relationships. The last principal, *change control*, is supported by ensuring that ownership and visibility management (modifying owner and/or audience of memos and tags), group management (creating, managing, joining and leaving user groups), and people-tagging (creating, modifying, or removing people-tags or changing their visibility) are all handled in a flexible and straightforward way and that all the settings are modifiable at any time, making it easy for users to make frequent changes to their information space.

6.4.5 Supporting Clarity

Our last design heuristic emphasized the importance of clarity in presenting action possibilities and their consequences to users, as in our study, lack of awareness of supported privacy functionalities or the consequence of information sharing decisions proved to be a major problem: our participants expressed a desire for functionality that already existed in the tool, which suggested that they may not have been able to use the privacy features

effectively. In OpnTag, we have followed some of the known principles of usable design to achieve clarity.

One such principle is to make privacy features (as secondary functionalities) highly visible and seamlessly available to users in the context of their primary actions (De Paula et al. 2005). This is derived from the fact that privacy features often act as barriers to action, while usability principles aim to remove such barriers (Dourish et al. 2004). Supporting proper visibility can thus help achieving the right balance between the two seemingly conflicting goals and ensure that privacy management features complement existing actions rather than inhibiting it. In OpnTag, relevant action possibilities for each privacy control are graphically clustered and presented together, while irrelevant or rarely used information are omitted in order to reduce clutter: owner and audience selection for a memo (OpnTag's privacy controls options for artifacts) are visibly presented at the top of the memo edition page, and are the only functionalities presented at this level, separated from other functionalities that deal with the content of the memo (Figure 6.2). Likewise, group visibility, member list visibility, and people tag visibility (OpnTag's privacy control options regarding audience and relationship) are presented in the same context that classic or egocentric groups are created or modified. While this makes these features visibly and seamlessly available to users in the context of defining information artifacts, audiences, or relationships, it still gives them the option to skip such configuration and accept the defaults, if they would rather focus on their primary task.

Another usability principle is consistency (Nielsen 1992). In OpnTag, various choices for each privacy control feature (owner and audience for memos, group and member list visibility for classic groups, and visibility of people tags) are consistently presented by dropdown menus, as a learned convention that is successfully and frequently used by people at all skill and experience levels. Designers of privacy systems are also advised to use feedback mechanisms to help users understand the implications of their privacy decisions (Bellotti and Sellen 1993, Lederer et al. 2003). OpnTag supports this through the combination of providing visualization and pairing configuration with action. For both memos and people tags, the choice of audience causes the background color to change accordingly. Such immediate visualization of visibility and the choice that triggered it will

help users understand how to generate rules that reflect their preferences, or to notice when the results of their actions do not correspond with their intended goal.

The clarity principle also suggests that users must be provided with clear explanations regarding feature functionalities, so that they can understand the consequences of their configuration actions. In OpnTag, this is supported through extensive use of tool tips (small labels that appears beside a button if user pauses over it). OpnTag tool tips are enhanced with added text description to make users aware of what each feature is for and what the results of each choice would be. Figure 6.10 shows examples of tool tips used for ownership and audience management, and Figure 6.11 shows examples of tool tips used for various actions on people tags.

Figure 6.10 Tooltips for owner and audience

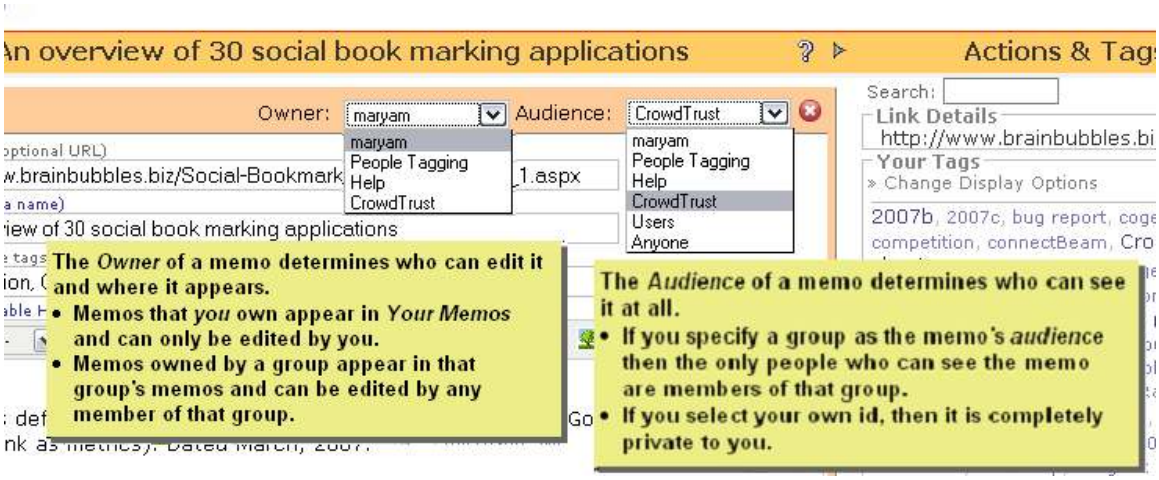
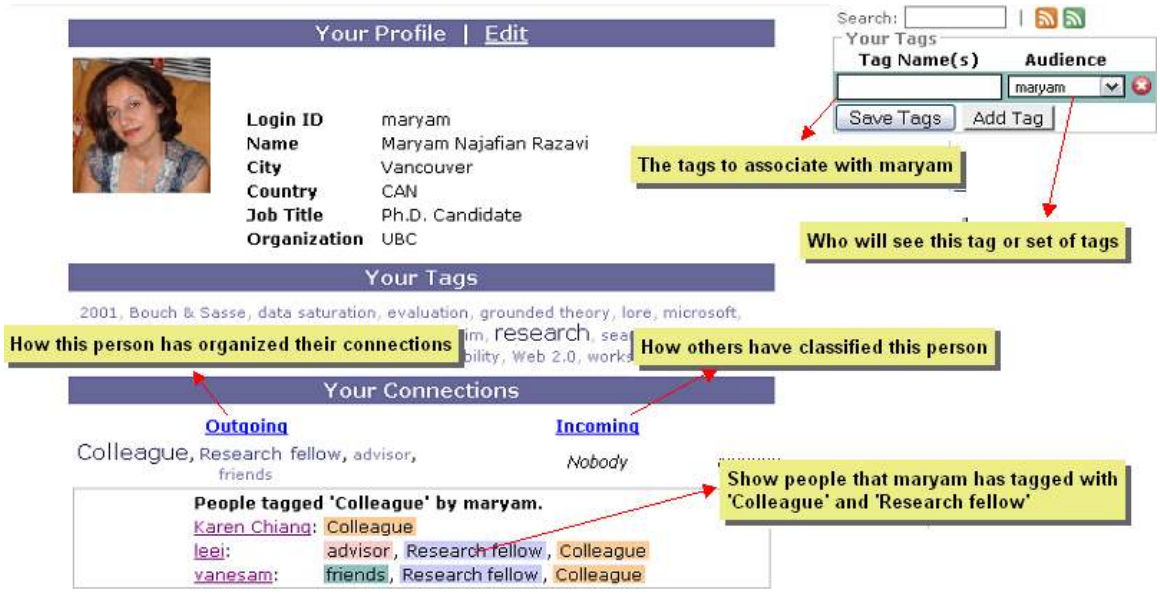


Figure 6.11 Tooltips for people tags



6.5 Summary

This chapter presented OpnTag, a social system we developed as an instantiation of the design guidelines for privacy management in social-personal information management systems as identified by our study. We described various key concepts of OpnTag, with an emphasis on the notion of classic and egocentric groups as OpnTag's main structures in support of selective information sharing. We explained how OpnTag allows definition of groups of various visibility and membership models, and explored the concept of tagging people for relationship management and by extension, privacy management, which has not been explored in the literature before.

We showed how people-tagging enables users to create one-sided, egocentric, user-defined social relationship groups and categorize their social network into groups of target audiences for their information in terms of their (often changing) relationships with them. Finally, we discussed how each of the five elements of our proposed framework is supported by OpnTag's privacy management mechanism.

Having explained our test bed and how it supports the design heuristics suggested in the previous chapter, our next step is to show that our proposed framework as instantiated in this particular implementation in practice yields a usable privacy management system that provides users with more control over privacy. We now move on to a discussion of an empirical lab evaluation of OpnTag in terms of both utility and usability, which is presented in the next chapter.

Chapter 7
Evaluation

This chapter describes the last stage of our research, which involves evaluation of OpnTag's privacy management mechanism. We start by clarifying the goals of the evaluation process and proceed to the description of the process of validating our approach. We also explain the challenges we faced in the evaluation process, and the limitations of the evaluation study. The study also indicated areas for improvement, which will be discussed in the next chapter.

7.1 Objectives

The two main features that distinguish our privacy framework from existing models of privacy management in current social systems are the introduction of per-artifact ownership management and audience control (as opposed to per-category access control supported by most current tools), and use of people-tagging for creating nuanced relationship groups (as opposed to equal-weight, reciprocal relationships supported by the network of friends model). The result of these two innovative features is the promise of a more fine-grained, flexible, and dynamic privacy management that is still easy to understand and use. As such, at a high level, the goal of our evaluation process has been to answer the following questions:

1. Are OpnTag's innovative privacy management features in practice successful in meeting users' varying privacy needs?
2. Is OpnTag's privacy management mechanism intuitive and easy to use for users, enough so as to be adopted in their day-to-day information management activities?
3. How is the people-tagging feature understood and utilized by users? Is the feature intuitive and compatible with user' mental model?

4. Does people-tagging enhance OpnTag's privacy management in terms of meeting users' privacy needs?
5. What are the constraints that would make people-tagging a useful feature for managing personal privacy?

Since our primary concern was not to do a comparative evaluation with another privacy management technique, we focused our efforts mainly on investigating two equally important quality attributes of OpnTag's privacy management mechanism from users' perspectives: the utility (does it do what users need?) and usability (is it easy to use?). The methodology we used was an empirical study consisting of three phases: an initial questionnaire, observing users performing pre-defined tasks with the system, and post-task, semi-structured interviews to gain feedback on their experience. This combination of methodologies allowed us to make detailed first-hand observations of how first-time users interacted with OpnTag's privacy management scheme and how they reflected on the usefulness, effectiveness, and usability of it.

The following sections provide a detailed description of our evaluation scheme and the results that were gained from it. Along the way, we discuss the challenges we faced in the evaluating process and the way we dealt with them. We conclude the evaluation section with a discussion of the limitations of the study.

7.2 Participants

OpnTag target users are individuals operating in various personal, professional, and social group environments, meaning, our sample pool was practically general public and there were no salient characteristics to look for in participants to assert their suitability for participation in the study. As such, a participation request was distributed via email to mailing lists and social connections of existing OpnTag users. Ten people (6 male, 4 female) who responded to the invitation were recruited for participation in the study. Our participants came from diverse backgrounds, including both technical and non-technical users, which was appropriate since OpnTag is designed for personal and social information management across a wide variety of usage contexts. All participants had some university

education, with 7 participants having a graduate degree and 3 an undergraduate degree. Table 1 shows a summary description of the participants in our study.

Table 7.1 Participants' demographics

No.	Area of expertise	Age	Expertise with Computers	Social systems usage
1	Medical Science	41	Passing	2-4 h/w
2	Graduate Student in Engineering	34	Expert	8-10 h/w
3	Information Technology	42	Expert	6-8 h/w
4	Graduate Student in Engineering	35	Expert	8-10 h/w
5	Media Production	39	Expert	4-6 h/w
6	Information Technology	37	Expert	6-8 h/w
7	Medical Science	30	Passing	2-4 h/w
8	Online Instructor	41	Knowledgeable	2-4 h/w
9	Accountant (CGA)	43	Knowledgeable	4-6 h/w
10	Information Technology	36	Expert	6-8 h/w

7.3 Procedure

Each of the ten participants was invited to attend a one-hour study session. The online availability of OpnTag allowed us to have the flexibility to carry out the sessions at the place of participants' convenience. 7 sessions were carried out at participants' offices, while the other three were done at researcher's office. After giving informed consent, participants were given an introduction to OpnTag and its privacy management mechanism. After clarifying potential questions, each participant was asked to carry out three consecutive steps:

1. Fill out a survey questionnaire
2. Use a pair of login name/password provided to him/her to login to OpnTag as an imaginary persona, and perform a set of pre-defined tasks
3. Answer a set of follow-up questions

We now describe each step in detail.

7.3.1 Phase One: The Initial Questionnaire

We started our evaluation process by asking participants to fill out a survey questionnaire consisting of 10 questions covering general areas of users' demographics (sex, age, education level), profession, expertise with computers and the Internet, plus specific questions on their familiarity with popular social software systems (i.e. LinkedIn[10], Orkut[11], Facebook[12], or other) and the privacy management features in them. The survey also included questions aimed at identifying users' privacy attitudes, i.e. what information they feel comfortable to reveal about themselves in social applications they use and whether they have ever experienced privacy violation problems in the past.

Figure 7.1 summarizes participants' familiarity with social systems and Figure 7.2 shows a summary of participants' privacy concerns. As the figures show, all participants had some familiarity with social systems (i.e. at least using one other social system for a prolonged period in the past) and some concern for privacy (i.e. at least considering some artifacts private and some sharing uncomfortable). These two criteria ensured that they had an idea of the subject of the study and can relate to it, and that their answers were not affected by lack of knowledge or concern about the topic of investigation.

None of our participants had used OpnTag before. On average, participants reported using various social systems for 6 hours per week. None had any special expertise related to privacy; however, when reflecting on their experiences with privacy management in applications they used, 7 participants said they find the feature useful and have used it at some point. 6 said although they consider privacy management mechanisms necessary, they often find them too difficult and/or time consuming to use. 5 said they don't trust social systems to put their private information there. Only 2 participants said they think social systems provide adequate privacy management. 7 out of 10 participants reported experiencing privacy violation at some point while using social systems, either as a result of their own action or others.

[10] https://www.linkedin.com
[11] http://www.orkut.com
[12] http://www.facebook.com/

Figure 7.1 Participants' familiarity with social systems

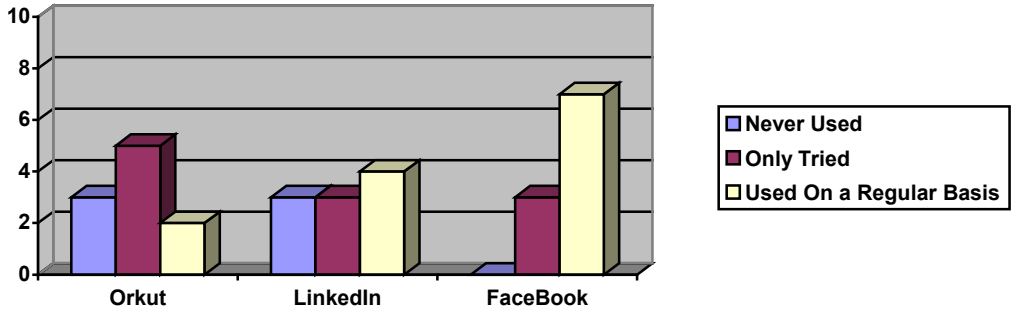
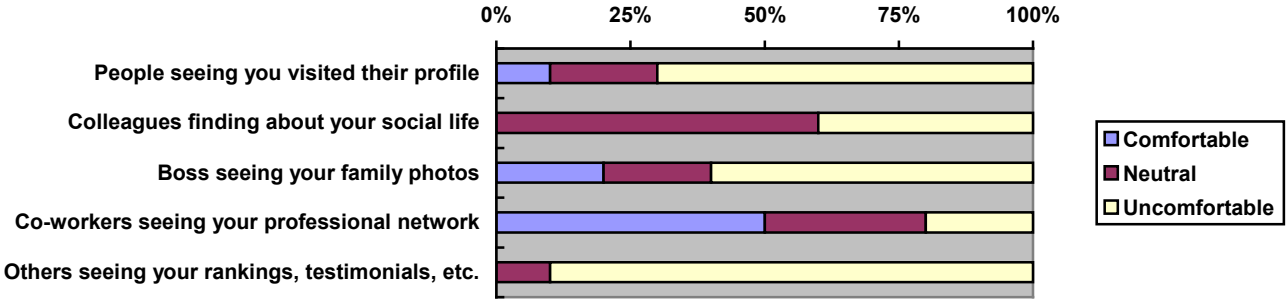


Figure 7.2 Participants' privacy comfort level



7.3.2 Phase Two: Observing Users' Interaction with the System

For the second stage of the study, participants were asked to login to OpnTag as an imaginary persona and perform a set of pre-defined tasks, each involving creating memos of various degrees of sensitivity and sharing them with various people in the imaginary persona's social and/or professional network. The tasks were selected to represent the kinds of tasks real users might perform while using a social system. For each task, participants needed to decide:

1. Whether to use classic or egocentric groups
2. The appropriate owner and audience for the memo they created

Participants were allowed to make any changes to the persona's information space that they felt necessary for the purpose of carrying out the tasks; including creating,

modifying, or deleting groups and/or people-tags, or changing the visibility of their groups/group members/people-tags.

While conducting laboratory studies, it is a challenge to ensure that the participants are representatives of real users, that they are put in settings that match expected usage, and that the tasks are designed as realistically as possible (Iachello 2007). The initial survey helped us address the first two challenges; by ensuring our participants relate to and care about privacy issues in social systems. Another factor to consider was to provide scenarios that had a high degree of realism for our respondents. This posed another key challenge on the process of designing scenarios for evaluation: when a tool is designed for just one security/privacy task (i.e. secure email or password), it suffices to show that users can do that task easily and successfully to prove the system usable. However, the goal of OpnTag is facilitating information management and sharing across a variety of personal, professional, and social contexts, and our goal was to show that our privacy model is capable of supporting users in managing their privacy in all those contexts while they are doing other primary tasks. As such, it was hard to come up with scenarios that showed off the capabilities of the privacy model by covering various privacy management activities in different contexts and yet, resemble real-life scenarios. On one hand, we needed uniform scenarios to have a standard basis for comparison; on the other hand, to have greater confidence on the validity of results, we needed to design tasks in a way to cover a variety of information sharing scenarios in each of our participants' personal, professional, and social lives.

To address this challenge, we developed four sets of tasks to have some flexibility in aligning usage scenarios to participants' varying professional backgrounds. We then used users' demographic data (gathered through the initial survey) to decide which set to present to each participant. Each set included 5 tasks, starting from rather easy tasks towards the more difficult ones. We considered the following criteria in designing the sets:

- That the tasks in each set cover information sharing situations across a variety of privacy sensitive contexts, ranging from inherently private, to semi-private, to public
- That the tasks in each set allow exploration of all four user-centered controls with the same proportion

- That the tasks in each set are meaningful to the potential group the set is targeted to
- That the tasks in each set require a mixture of visibility and ownership control

The first set (set A – personal scenarios) was used as a practice session for all participants to familiarize them with the nature of the tasks that were expected in the course of the study. While performing tasks in this set, participants were encouraged to ask questions if they needed clarification. Upon completion of tasks in set A, participants were asked to complete the tasks in a second set on their own (i.e. without receiving comments or clarifications). The decision on which set to give to the participant at this stage was made based on participant professional background as stated in the initial survey: participants in information technology and media production (P3, P5, P6, and P10; a total of 4) were given set B scenarios, situated in the corporate environment. The two students plus the instructor (P2, P4, and P8; a total of 3) were given set C scenarios, situated in an academic setting. And finally, the two participants with backgrounds in medical science plus the CGA (P1, P7, and P9; a total of 3) were given set D scenarios, situated in a small business environment. The sets were carefully constructed so as to differ only in the nature of the data and relationships described, but not in the actions needed on the part of users or the privacy concerns reflected.

During both the practice and the actual portion of the evaluation session, subjects were directed to think aloud while going through the tasks. This helped us understand why they chose to do things a certain way and assisted in identifying potential sources of problems. Tables 2 through 5 provide a description of the four sets of scenarios used in this study and the expected outcome of each task. In the following descriptions, the term *personal memo* has been used to refer to a memo that is owned by the user, while *group memo* refers to a memo owned by a group. For simplicity, all personas involve in each scenario were made members of OpnTag beforehand, and participants were asked to assume that all personas use OpnTag as their primary communication tool.

As we go through the description of tasks, two points are important to note: first, it should be noted that for most tasks, there were more than one correct way to complete the task with no certain right or wrong. For example, to share something with only one person

(a task presented in all four sets), the user could either create a private group consisted of him/herself and the receiver, or use a tag on the receiver (again either a public, restrictedly shared, or private tag) to create an egocentric group, with both paths resulting the disclosure of the created memo to the right audience.

Second, although the privacy management requirements for the tasks in each set were similar, users had different ideas regarding whether something should be public, selectively shared, or private. Participant 5, for example, thought that the resource on head lice (task A.2) should be shared with parents of children in the same class, while all other participants preferred to keep it private. We considered the outcome "correct" as long as the user's actions resulted achieving the "right" level of privacy; i.e. exposing the created memo to the intended audience as mentioned by the user in the think aloud procedure.

Table 7.2 Set A – Personal scenarios

Task#	Description	Expected Outcome
1	You try a new restaurant and want to recommend it to your friends.	personal memo made visible to group containing friends
2	Your kid has been diagnosed with head lice. You find a useful resource on the subject (www.LiceResource.com) and want to keep it for your reference.	personal memo made private
3	You find a good resource on supporting math activities (www.MathResource.com) and want to share it with all parents in your kid's class.	personal /group memo made visible to group containing all class parents
4	You want to arrange a play date for your kid with a certain kid ('s parent).	personal memo made visible to a specific parent
5	You want to invite four certain kids to your kid's birthday party.	personal memo made visible to four specific parents

Table 7.3 Set B - Professional scenarios (corporate) - The participant is acting as a manager for Team A.

Task#	Description	Expected Outcome
1	Team A is collaboratively working on a new design. You want to create a space for everyone to contribute to design ideas where everyone can get credit for his/her contributions.	Group memo visible to group containing employees in Team A
2	You have been to a business trip and are preparing a report on it for your boss. It is an ongoing draft at this point.	Personal memo made private
3	The report is complete and you want to share it with your boss.	Same memo made visible to the boss
4	You are working on a new development with your colleagues. You find out that a competing company has just launched a new product similar to what your group is working on. You want your colleagues to learn about this and study the new product (www.CompetingProduct.com).	Personal/group memo made visible to group containing colleagues
5	You want to endorse a certain team member to your boss and recommend him for promotion.	Personal memo made visible to boss

Table 7.4 Set C - Professional scenarios (school) - The participant is acting as a graduate student

Task#	Description	Expected Outcome
1	You want to share a paper draft with your advisor.	Personal memo made visible to the advisor
2	You find a link to some job resources (www.JobResource.com) that you think might be useful in your future job hunt.	Personal/group memo; could be public, private, or shared with fellow students depending on users' choice
3	You want to keep notes on ideas for potential papers you are planning to write in the future.	personal memo made private (most probably; again depends on users' choice, though)
4	You and a fellow student are attending a conference and want to record your reflections on it and also give your other lab mates an idea of what's going on there.	Group memo made visible to lab mates
5	You come across a paper at www.Apaper.com and think it might be useful to a certain lab mate.	Personal memo shared with the certain lab mate; again can potentially be made public or visible to other lab mates as well

Table 7.5 Set D - Professional scenarios (office environment) - The participant is acting as the business owner

Task#	Description	Expected Outcome
1	You want to create a space where everyone writes down their vacation time so that there is no conflict.	Group memo made visible to all administrative staff and employees
2	A new procedure/regulation is available at www.NewProc.com that all employees must be aware of.	Personal/group memo made visible to all employees
3	There is a new office policy (i.e., change of rates) that all administrative staff must be aware of.	Personal memo shared with all administrative staff
4	There has been a recruitment issue that you want to discuss with your partner.	Personal memo made visible to the partner
5	A client has complaints about a specific employee. You want to discuss the issue with that particular employee.	Personal memo made visible to the certain employee

7.3.3 Phase Three: The Post-Task Semi-Structured Interviews

We closely observed our users' interactions with OpnTag while performing the tasks, looking for errors as well as signs of confusion and/or frustration. Upon completion of the tasks, we engaged each participant in a semi-structured, in-depth interview to gather subjective data on various aspects of OpnTag's privacy management mechanism, including ease of use, effectiveness, relevance of the tasks to their every day information management activities, and willingness to adopt the tool. We also tried to uncover the strengths and weaknesses of the people-tagging functionality and identify constraints that would make the feature more useful and usable. The general format of the interview was to use participant's actions during the tasks as a starting point and then try to cover a number of main topics. In all cases, we made sure to seek participants' answers to the following questions:

IQ1) Are the scenarios realistic and relevant to users' everyday information sharing practices?
IQ2) Are the additional privacy control features (ownership control, per-artifact control, and people-tagging) useful and usable?
IQ3) Is the difference between classic and egocentric groups clear, i.e., how users decide when to create groups and when to tag others?
IQ4) Is supporting five different visibility options for people-tags useful, or whether in practice, certain visibility options (e.g., public or private people tags) are used most of the times?
IQ5) Are users comfortable with the idea of other people assigning tags to them without explicit permission, or whether adding some sort of control on the part of taggee is considered necessary?

The open nature of the semi-structured interviews enabled us to gather ratings of usefulness and usability of our privacy framework from first-time users, probe for the reasons behind the ratings, and discuss opportunities for improvements. An interesting result of using this methodology was evolution of new ideas from participants that were not part of our interview questions (as discussed in the following sections). However, a discussion guide was used throughout the interviews to maintain consistency and reliability, and to eliminate potential bias on the part of the researcher.

7.4 Results

In this section, we reflect on the utility and usability of our privacy framework, based on both our observations and participants' feedbacks. We also discuss intuitiveness and suitability of the people-tagging feature in enhancing users' privacy management experiences across the various usage scenarios.

7.4.1 Ease of Use and Effectiveness

We asked users to rate OpnTag's privacy management mechanism in terms of both ease of use and sense of privacy compared to the social software systems they were familiar with (Orkut, LinkedIn, Facebook). We used a scale of 1 to 5 for rating, with 1 indicating the worst performance and 5 indicating the best. Ease of use was rated 4.2, with min = 3 and max = 5 (users thought there were just too many steps involved to navigate to a user profile for tagging). Users' gave their perceived sense of information privacy an average rank of 4.0, with min = 3 and max = 5. Although not all of our participants took the optimum path for doing all scenarios, they were all able to navigate their way through the privacy management system to get the tasks done. From the total of 50 tasks that our participants performed unassisted (10 participants each doing 5 tasks), we only witnessed 5 errors. Furthermore, all 5 errors were results of improper understanding of the task in question; for example, making a memo visible to team members rather than colleagues, as mentioned in task description.

Overall, the concept of setting the owner and the audience for access management seemed to be fairly understandable to users: none of the users showed any signs of confusion or frustration. Also, the majority of our participants seemed to grasp the difference between granting "write" vs. "read" access: 9 out of 10 users correctly created a group memo for the tasks that involved some form of collective contribution (tasks B1, C4, and D1). Since this distinction is not supported by most of the existing tools, it was encouraging to see users quickly picking up and using a fairly new feature.

7.4.2 Usability

Traditionally, usability of a software application is measured based on four quality components: speed, accuracy, success rate, and overall user satisfaction (Nielsen 2001). However, researchers have often considered different criteria for measuring usability of a security or privacy mechanism. The reason is that privacy management is often a secondary goal in most systems, and therefore does not get the same consideration that many other aspects do (Egelman and Kumaraguru 2005), which makes it difficult to set particular metrics for usability of privacy aspects (i.e., what exactly should be measured?). Whitten and Tygar (1998) were the first to propose a working definition of usability for security software based on the special characteristics of the usability problem for security, and suggest several criteria for evaluating usability of a security system. A number of other researchers have also proposed similar and/or complementing guidelines for evaluating usability of security or privacy mechanisms (Cranor 2005, Karat et. al 2005, Chiasson et. al 2006, Clark et. al 2007). We found these criteria suitable for the purpose of our study. Here we reflect on the usability of OpnTag's privacy framework based on Whitten and Tygar's four usability criteria, plus the two complementary criteria suggested by Chiasson.

- ***Users must be reliably made aware of the steps they have to take to perform a task***

This is a restatement of the first guideline of (Whitten and Tygar 1998) and suggests that the application must provide user with enough cues as to how to start the process for each task, and to identify the intermediate steps that are required to complete the task. In OpnTag, the acts of setting a memo's owner and audience are fairly straightforward, because those action possibilities are associated and presented with memo creation/modification functionality. Also, the fact that privacy management in OpnTag happens on a per-artifact basis, made it easy for participants to figure out that the owner and audience are the two attributes of a memo that they need to set in order to share something with a certain audience.

- ***Users must be able to determine how to successfully perform the steps***

Whitten and Tygar's second usability guideline suggests that once the user is made aware of what intermediary steps are necessary for each task, she must be able to figure out how to perform these steps. (Wharton et. al 1994) suggest that users develop a mental model of how a system works, and that in order for users to be successful in performing the necessary steps required to complete a task, the model behind the system must match user's mental model.

In our study we designed the scenarios so that the privacy management requirements for the tasks in each set were similar. However, participants employed different privacy management strategies based on their privacy attitude and concerns. For example, participant 2 considered a potential job resource (task C.2) as private data, while participant 8 thought it should be public. Regardless of their privacy attitudes, our subjects were successful in achieving their desired level of privacy, properly disposing the created memo to the right audience. Moreover, OpnTags' privacy management system seemed to have a fast learning curve: after the initial training session, our subjects all seemed at ease with creating memos, defining or modifying groups and/or group members, tagging other users, and choosing the appropriate owner and audience for their memos.

- ***Users should not make dangerous errors from which they can not recover***

Since OpnTag is an information management system, the most dangerous error that can happen is exposing a memo to the wrong audience. However, in OpnTag memos are created in the current work space; meaning the default value for both a memo owner and audience are the user himself (if in user space) or the group (if in group's shared space). As such, even if users miss to set the right owner/audience, having rather conservative values as defaults help decrease the chance of accidentally making a memo visible to a too large audience (e.g., public). Also, the different background color codes that reflect various levels of visibility for a memo provide a powerful visual cue to the user as to whether the memo is set to have the right visibility.

- ***User should know when they have completed a task***

This is the first complementary criterion to Whitten and Tygar's proposed by Chiasson et. al (2006) (also mentioned by Cranor (2005)). This criterion suggests that one of

the essential usability requirements is enabling users to tell when their task is completed, which implies that the feedback provided by the system to users during a task should be adequate to ensure they are aware of its successful completion.

We asked our subjects several questions in an attempt to gauge their perception of the appropriateness and adequacy of the feedback provided by the system (the owner name, color-codes, and group name in the resulting memo). After each task, we asked the participant if they believed they performed the task correctly, and if yes, how could they tell if they have been successful in setting the appropriate privacy level for the artifact they created. 8 of our subjects mentioned the use of at least one of the feedback mechanisms to check their results, while the other two participants just assumed they did it right and had not paid attention to any of the feedback information.

Overall, the consensus was that the combination of color codes and owner/audience in the memo list conveyed enough information to users to form an idea of the current privacy level of their various artifacts at a glance. 6 participants mentioned the color-code as the most useful feedback mechanism, probably because of its visualization power. Interestingly though, our choice of color-codes was not popular with the participants. One participant mentioned that blue (OpnTag's color-code for public) and green (OpnTag's color-code for private) are quite easy to be mistaken, and suggested we use other colors for the two cases that show the distinction more clearly. Another participants thought green is not the right color-code for private, since it implies "green light" in a way. This participant thought a more strong color like red would be a better choice for the case.

- ***Users must be able to determine the current state of the system at all times***

This is the second guideline from Chiasson et. al (2006), and suggests that it should be visible at a glance what is visible to whom. Cranor (2005) advocates the use of "persistent indicators" that allow the user to see privacy information at a glance. OpnTag does this through its use of color-codes.

- ***Users must be sufficiently comfortable with the interface to continue using it***

This is the fourth principle of usable security of Whitten and Tygar (1998), and is an essential part of the principal of psychological acceptability quoted by Bishop (2005). After completing the tasks, we asked our subjects if they can relate to the scenarios and whether a tool like OpnTag with advanced privacy management features would be more likely, less likely, or just as likely, to be incorporated in their daily personal and social information management activities. Most participants (9 out of 10) said that although the scenarios don't resemble their information sharing practices exactly, they could think of similar scenarios in their day-to-day activities where similar selective information sharing activities would be useful or necessary. 8 out of 10 participants said that they would try using such a tool for information management and sharing, and 6 participants said they feel a strong need for a tool with these sort of privacy management in their work place. Table 7 summarizes some of participants' comments regarding willingness to adopt the tool.

Table 7.6 Participants' comments on willingness to adopt the tool

P #	Comment
3	It would be nice to have something like this that limits the *contribution* to a certain group, even though it may be *exposed* beyond that group. For example, I would like the contributors to a discussion on board level design to be limited to: X [who is a board designer], Y [who is his boss], Z [my boss], and me. That's all people who are knowledgeable enough to contribute to this discussion, although the whole group may read it.
5	I strongly feel that something like this is needed in our workplace, even though personally I am against using something like this because of privacy reasons: I don't think one should put his ideas in a system on the Internet; unless he can control who would have access to it.
6	One thing I like about this [OpnTag] is the control; I like that it can act as an integrated environment; because you can separate your personal and work-related stuff. I think that's very important. The idea of having my desktop online sounds cool.

7.4.3 Understanding the Difference Between Classic and Egocentric Groups

One innovative aspect of OpnTag's privacy management mechanism is the addition of egocentric groups for relationship management through people-tagging. The two main

differences between the people-tagging functionality and the classic group functionality for privacy management are that:

- People-tagging is user-centered; meaning the tagger solely decides to assign taggees to various egocentric groups created by people tags, whereas with classic groups, users decide to join or leave
- People-tagging is lightweight and dynamic; people tags can be assigned or removed easily and frequently, which makes people-tagging suitable for occasional information sharing, as opposed to continuous or archival sharing which is better suited to classic groups

In order to find out how users compared people-tagging with classic group functionality, we asked our participants how they decided which feature to use for each task. Participants' answers to this question showed that overall, people-tagging fared well as a new concept: although all participants said they needed the practice session to grasp the concept, they were all able to perform tasks in set B, C, or D on their own, properly using either group or people-tagging functionality to create the right audience for the memos as described in each task. The following quotes are some of participants' comments comparing people-tagging and group functionality, which show that they understood the distinction and the added functionality and control provided by people-tagging.

Table 7.7 Comparing group and people-tagging functionality

P#	Comment
2	Group is for longer-term communication when you know what the focus is, like school. Tags are short-term, and for people you may only have a single commonality in a specific context.
6	Group is for ongoing communication, whereas people-tagging is about how to expose my information to the people.
9	People-tagging is on-the-fly, temporary, dynamic.
4	My groups are like forums or communities; with the same definition in all social systems; whereas people-tagging is a personal and opinionated categorization of my network.
1	I say I classify people this way and I can put my classification into the system. It may not necessarily be used only by me; others can use it, too, or they can combine it with their (different) ideas; But my definition is that memo tags are the way I look at information and people-tags are the way I look at people.
3	I will use group if I want to share with everyone; I will use people-tagging if I want to share with an individual or a subset of a group.

7.4.4 Usability of the People-tagging Interface

We observed that the path to tagging a user was not completely intuitive to all our subjects. While all ten participants eventually figured out how to use it, they consistently reported two aspects of the interface problematic:

1. The fact that users needed to navigate to a user's profile page to tag him/her was confusing: some spent a lot of time the training tasks before they realized they needed to click on a users' name to navigate to his profile in order to tag him. participants rather expected to be able to "add a user to an existing tag" (rather than "adding the tag to the user"), much like the way people join a classic group.

2. Even after figuring out how to do it, our participants found the path to navigate to a user profile page cumbersome in the system in its current shape, and expressed the need for an easier, more obvious way to do so, i.e. a user search option or a button to navigate to a page with all users' information and the ability to add tags.

Other desirable options mentioned by participants included the ability to tag a number of users at the same time (mentioned by 6 participants) and having a pull-down menu at the audience button to enable "tagging in place", similar to Gmail labels (mentioned by 3 participants).

7.4.5 Suitability of Visibility Options for the People-tagging Functionality

There were some variations among participants on their choice of the visibility level for people tags they created. Overall, we observed that users were more inclined to make their tags private or semi-private rather than public. We asked participants to reflect on their choices and explain how they decided to assign a certain visibility to a tag.

8 out of 10 participants said they think that in general, people tags should be exposed to the minimum audience possible (i.e. just tagger or tagger + taggee). 2 participants said they might occasionally disclose a people tag to others similarly tagged, but only if the situation requires it. The rationale behind this argument was that participants mainly considered people tags as their personal opinion about others. As such, did not see any reason or need to share this with others, even with the subject of that opinion. An exception to this argument came from two of our enterprise participants: although this group agreed that people tags should not be shared unless there is a certain need to do so, they could envision situations in the workplace where that need would be present. Examples included to mark someone as the "goto" person for a certain job, and to specify who is currently in charge for tasks that are handled by different people at different times (e.g. who is writing automated tests right now?). These two participants thought public people tags would be extremely useful in such cases, providing valuable, time-saving information, specially to someone who is new to the corporation.

7.4.6 Taggee's Control on Incoming Tag Cloud

Next, we asked participants to imagine using the people-tagging feature in a social system they are familiar with (i.e. Facebook): would they feel comfortable with the idea of

others tagging them without their confirmation or even knowledge? 9 out of 10 participants said they think some sort of control on taggee's part would be necessary if they are going to adopt people-tagging as part of their daily information management activities. One participant, however, thought that since people tags don't give the tagger any privilege over taggee's data, it is OK to leave the feature as it is. It is important to note, though, that this particular participant was envisioning the feature in an enterprise setting, where transparency (the ability to link the tags to tagger's identity) would be a natural barrier to anti-social tagging.

We then followed up with the other 9 participants by asking which of the following control options they consider most appropriate:

- Asking for taggee's confirmation for incoming tags
- Giving taggee the ability to delete his/her unwanted incoming tags
- Giving taggee the ability to change the visibility of his/her incoming tags

Among these, the confirmation option was most popular, voted for by 8 out of the 9 participants as having the least social implication. Only one participant said she would find it awkward not to confirm an incoming tag, especially if it is from someone she knows, and would prefer the ability to delete unwanted tags at a later time. The 8 participants who preferred the confirm option said they find the delete option "too reactive" and "rather pointless" (because the tag has already been exposed). None of the participants thought the ability to change the visibility of an incoming tag (i.e. to make it more restrictive) would be useful. They thought there is no point in changing visibility considering that the tagger can change it back. Also, they all found it natural for the tag visibility to be tagger's choice, "much like the CC option in email" as one participant put it.

Although in the trade-off between keeping the current lightweightness and flexibility of the feature vs. having more control over tags applied to them the majority of participants opted for more control, some were able to come up with compromises (control options other than the ones we suggested) to keep the benefits of both worlds. One participant mentioned that she only thinks confirmation is necessary because currently in OpnTag anyone can tag anyone. She suggested limiting the ability of tagging someone only to his/her

social network, since one's social network would probably know him/her enough not to use any inappropriate tags. Another participant said he would find a simple notification of the incoming tags good enough and wouldn't need a confirmation mechanism, an option that would keep the flexibility of the feature intact while providing the taggee with some awareness.

7.5 Discussion

We believe that while our evaluation study was small-scale, it was appropriate at this initial stage for successfully gathering meaningful feedback from potential users. Our results indicated that people-tagging could potentially enhance OpnTag's privacy management mechanism by enabling new social interactions that are not supported by current tools, such as transient/occasional information sharing and definition of nuanced social networks. We were particularly encouraged by the fact that users seemed to quickly grasp the difference between privacy management through classic groups vs. people-tagging, and to appreciate the added utility that the people-tagging feature provided, which we thought was promising for a fairly new concept.

7.5.1 Compatibility with Users' Mental Model

Previous research has demonstrated the difficulties that users face in completing security tasks when the technology that is provided to meet users needs does not match their mental model (Whitten and Tygar 1999). Perhaps the most interesting outcome of our evaluation process was the indication that the underlying cognitive model behind OpnTag's privacy management mechanism was a suitable match for users' mental model of information privacy. Our participants successfully identified the steps required to make an artifact visible to a selected audience, were able to carry out those steps with very few mistakes, and seemed to understand the implications of their privacy decisions to the point of making sophisticated observations and creative suggestions (other than what we proposed) about the people-tagging model after only half an hour of controlled interaction with the system. In addition, when asked, all said they think people-tagging makes access control easier. We believe these evidences suggest the advantages of this model over the existing models of privacy in current social systems.

Current privacy models for social software suffer from two major problems: they either define access control as a private/public dichotomy, ignoring the various other shades of privacy in between (e.g. del.icio.us), or they are complicated and cumbersome for the average user, to the point that they cannot always predict the consequences of their privacy decisions (e.g. Facebook). Our privacy model seems to address both of these problems by giving users a simple, well-defined framework that works at the level of an individual artifact, and which they can use to confidently control and reason about privacy. However, a deeper study of the expressiveness, understandability and predictability of these different approaches is needed for a grounded comparison.

7.5.2 Study Limitations

There are a few limitations to the study that must be taken into account when interpreting the results. Ideally, we would have liked to perform a field evaluation with real users using OpnTag and its privacy system in real-life information sharing situations. The advantage of field studies is that they report on users in their natural environment doing real tasks, demonstrating feasibility and in-context usefulness. The disadvantage, on the other hand, is that they are time consuming to conduct, and require mass adoption.

Despite the potential difficulties, we did initially aim for capturing real data from real user population of OpnTag and administered two small field trials within two pilot groups (Appendix A). However, in practice creating a virtual community in these setting that can serve the purpose of this research proved difficult, as we noticed that the privacy management mechanisms were not used extensively in either of the test deployments. We believe this might be attributed partly to the context of use for the field trials (one educational, one in a small team of collaborators) and partly to the fact that privacy and security features are usually appropriated late in the learning curve of an application, often after some unexpected security or privacy "incident" (Iachello 2007). This clearly reflects the difficulty of privacy evaluation as a secondary task, which plays into evaluation in that the privacy management features do not often get as much attention as they should as the result of often being a trade-off with performance and/or functionality.

It quickly became clear to us that gathering rich data on the privacy management behavior of real OpnTag users would depend on two essential requirements: extensive time and an active user community; both beyond the affordances of a Ph.D. thesis. At this time, a comprehensive review of literature on privacy evaluation methodologies (e.g., Hawkey and Inkpen 2007, Cranor et. al 2006, DeWitt and Kuljis 2006, Iachello and Hong 2007, James et. al 2007, Chiasson et. al 2006) showed that many times, privacy researchers opt for controlled laboratory studies, especially in the early stages of the work to examine the viability of an approach before proceeding to more fundamental implementation/evaluation schemes. Our goal at this stage was also finding out whether our framework yields a viable approach for improving privacy management, whether it is intuitive enough, and to get feedback on the potential problems users encountered with it so that we can refine the design. As such, a controlled laboratory study of a smaller sample seemed appropriate to gain a profound understanding of these issues, and we believe our evaluation scheme provides sufficient answers to these questions.

However, while our laboratory evaluation was effective at getting users' initial feedback, it also came with some limitations of its own. First of all, even though we tried to recruit participants of varying background to cover usage scenarios across a variety of contexts, our representative sample was a small one and our representative usage scenarios almost certainly did not cover all privacy concerns and contexts of use for all potential users. However, it must be stated that this criticism can be applied to the vast majority of laboratory studies in this area. Different opinions and problems may well be expected for other types of user, which is best to be investigated in follow-up studies.

Second, although laboratory studies provide an appropriate situation for observing users' interactions with the system in a controlled fashion, there are inherent challenges associated with laboratory studies in the privacy domain. One such challenge is that even though the researcher might try to mimic various information sharing situations as realistically as possible in designing usage scenarios (as we did in ours), because participants are not dealing with their own data, they might not be as motivated as in real life and as a result, not make the same effort to protect privacy of their data.

Another issue is that methodologies for studying privacy may themselves be deemed too privacy-invasive, causing users to deviate from normal practice and/or to withhold revealing sensitive aspects. As a result, relying on self-reported attitudes and behavior alone may not provide a valid view of normal practices. Many privacy studies suggest that there is a gap between users' stated privacy preferences and their real behavior (Ackerman et. al 1999, Jensen et. al 2005, Spiekermann 2001); however, such gaps are hard to capture in a lab study and are best addressed in field studies.

Another limitation was that all the tasks addressed sharing information within someone's network, i.e., sharing with *known* people; The nature of the lab study didn't allow us to investigate information sharing behavior with *unknown* people that are also present in a social software system (i.e. open communities) to see whether the added clarity of groups dynamics supported by OpnTag's advanced group features contributed to improving ease of decision making in selective information sharing within groups.

We also believe that while general feedback was positive, there might be social implications that didn't show up in this limited study and may show up after OpnTag has been in use for a prolonged time; i.e. maintenance and space pollution (spam) for people-tags. All points indicate that a longitudinal evaluation in the field might be necessary to ensure flexibility of OpnTag's privacy management mechanism in accommodating a wide range of privacy needs over time.

7.6 Summary

In this chapter, we evaluated OpnTag's privacy management mechanism in terms of both utility and usability. We also explored whether the concept of tagging people for relationship management was successful in enhancing personal privacy in OpnTag. Through a preliminary empirical evaluation of OpnTag's privacy management model, we provided initial validation of the viability of the proposed framework for building usable privacy management systems that provide users with more control over privacy, although we did not directly address its completeness or adequacy. Our evaluation study also provided evidence of the feasibility of the people-tagging concept and helped revealing users' ideas regarding

how to make the feature more useful and usable for their everyday privacy management. Our results also identified a number of areas where improvements might further increase usefulness and usability, including a more obvious feedback mechanism, improving the interface and opportunities for people-tagging, and better navigation mechanism between pages. Further evaluation may be required to validate the model in a natural usage environment.

It is important to note that the evaluation study must be considered in the context of an iterative design/evaluation process: our study presents an evaluation of the first iteration of the design and implementation of OpnTag's privacy management mechanism as an instantiation of our proposed framework, and like any first iteration, it successfully identifies the parts of the design that work, as well as the problems that need to be rectified in a second design iteration. Moreover,since the study focused on a particular implementation of the principles, we are unable to use it to validate the principles themselves. Instead, at this point we can only claim to have demonstrated that exposing this extra expressibility and complexity in the way we have done so in OpnTag does not reduce the usability of the system, and does seem to resonate with our users' own assessment of their privacy needs.

Chapter 8

Conclusions and Future Directions

8.1 Summary

Web 2.0 applications are getting more and more adopted in our daily lives for various purposes, including personal and social information management. The research presented in this thesis has been motivated by the need for usable privacy management support for the non-technical user population of these applications, to enable them to regulate access to their personal and social information they dispose in these tools.

In chapter 2, we presented an overview of the previous research on personal privacy. This review emphasized the importance of empirical investigation of users' needs and perspective in the design of a usable privacy system, and showed that research into users' attitude towards personal privacy issues has been so far limited to the boundaries of specific tool or domain (i.e. ubicomp or eCommerce). This literature review also indicated that there is a lack of research examining specific privacy issues surrounding selective disclosure of user-created content (on which users have vested interest) in SPIM domain.

After identifying this gap, we moved on to the presentation of our foundational research, a grounded theory study aiming at understand end-users' information sharing behavior in a social context, to identify specific privacy needs and concerns in this domain. In chapter 3, we presented a description of the grounded theory methodology and the data collection and analysis techniques that were used. This study was designed with two research goals in mind. The first goal was to develop a better understanding of users' information sharing behavior and needs in SPIM domain, while the second goal was to propose a framework for designing usable privacy management mechanism for SPIM tools that support those needs.

In chapter 4, we discussed the results of the study based on the feedback we received from our 12 participants, such as challenges and strategies in achieving the desired level of privacy, and factors that affect users' decisions regarding disclosure of their personal artifacts to various people and groups in a SPIM tool, including change of privacy preferences in various stages of the information life cycle, the nature of trust between the owner and the receiver of information, and the dynamics of the group or community within which the information is being shared. Findings of this study clarified users' perspective on the privacy of the information they dispose in this domain, and offered some ideas about how to create privacy management mechanisms for SPIM that are based on users' mental model of information privacy.

In chapter 5, we validated our grounded theory by providing evidence that it meets each of the four criteria of fit, relevance, workability, and modifiability that have been proposed in the literature for measuring validity of a grounded theory. We then proposed eight design heuristics for privacy management in SPIMS based on the results of the theory which led to the description of five user-centered controls for privacy in this domain, including artifact control, audience control, relationship control, change control, and clarity. These five user-centered privacy controls constituted the building blocks of a framework for designing usable privacy management mechanisms for SPIM that is grounded in empirical findings.

In order to illustrate the suitability of our proposed privacy framework for design, we next implemented our proposed framework in an experimental SPIM tool to create an environment in which we can test it in action. In chapter 6 we presented our test bed, OpnTag, for which we have built a privacy management mechanism as an instantiation of our four proposed user-oriented privacy controls. We described various key concepts of OpnTag, including classic and egocentric groups that are OpnTag's main structures in support of selective information sharing, and showed how each of the four user-oriented privacy controls are supported by OpnTag.

Chapter 7 reported the forth and final stage of the research, which involved empirical evaluation of OpnTag's privacy management mechanism, in order to show that our proposed privacy model as embedded in OpnTag is both usable and successful in satisfying users various privacy needs. For that, we conducted a laboratory user study with 10 participants to evaluate OpnTag's privacy system in terms of both utility and usability. The evaluation facilitated the assessment of the specific design, and helped to assess the appropriateness of our framework and to identify problems. The preliminary empirical data from these first-time users provided initial validation of the viability of the proposed framework for building usable privacy management systems that provide users with more control over privacy. Participants' feedbacks also provided several suggestions for improvement, including the need for a more clear feedback mechanism and easier navigation between pages to further improve ease of use and usability.

8.2 Extensibility of Results Beyond SPIMS

It is important to reemphasize that while the work in this dissertation is situated in the context of SPIMS, the domain of applicability for the recommendations that were derived from our findings (i.e., design heuristics and the privacy framework) is the much wider general area of social software systems, rather than just the SPIMS. To reiterate from Chapter 1, SPIMS often have heightened privacy issues compared to other types of social software systems (e.g., social networking tools), because they include information artifacts that users have vested interest in. Nevertheless, the various privacy concerns, needs, and factors that were identified in our grounded theory study and make the foundation for the recommendations, are true for other social software systems as well. This is evidenced by the fact that many of the shortcomings of current social software systems with regard to privacy can be explained in terms of not meeting one or more of our proposed design guidelines. As an example, here we discuss current state of privacy management in Facebook (as a widely used social networking system with extensive privacy controls) with regard to our proposed design heuristics, and provide examples of two privacy problems in Facebook that are a result of lack of support for some of those guidelines.

8.2.1 Current State of Privacy Management in Facebook

Facebook provides very limited support for fine-grained privacy control (H1): while users can set different visibility for their various profile items, wall posts, and friends list through the privacy setting page, the only available options are friends and networks or a combination of those. A first level transitivity for the friends option is also supported (i.e., friends of friends), and it is also possible to exclude certain friends from viewing certain artifacts. Exclusion is the only way to create a selected subset of one's network and friends as the target audience for these artifact, which is obviously a labor-intensive and overall, non-practical way for the task. Privacy settings for photos are set at the album level, meaning, there is no way to set different visibility for various photos in the same album.

No support is provided for in-context control (H2), as all the privacy controls are located in the "Settings -> Privacy" page. Ownership control (H3) is not supported either, as all the privacy settings only control the visibility of artifacts. Facebook users currently have no control over who can tag them in a photo, leave a comment on their notes or status, or further share their posted items with others.

With regard to support for various group models and users control over group dynamics (H4 and H5), there is again very limited support: Facebook groups are either global (i.e., anybody can join), or by invitation. Lists are another group model supported by Facebook. Any Facebook user can create various lists and assign people form his/her social network to it. Although this seem like an equivalent of relationship control (H6), Facebook users can only chose to *receive* new feeds from certain lists; that is, the purpose of lists is to support selective reception of information, rather than selective disclosure of information. Currently in Facebook there is no way to limit visibility of one's posted material and other information artifact to a subset of his/her friends or network(s). Finally, any change in users' privacy preferences (H7) has to be administered through the "Settings-> Privacy" page. Furthermore, the change option is only available for the artifacts listed on that page (e.g., profile items, contact information, friends' list).

With this brief description, we now discuss two examples that illustrate how violating these heuristics lead to privacy problems in Facebook.

- ***Example 1:***

During Iran's 2009 disputed presidential election, people relied heavily on Facebook to communicate with the world in a timely manner despite the strict censorship that was imposed on the media by the government. While Facebook enabled people to get information out fast and to share it with a wide audience, the fact that it didn't allow users to select the audience for each of their posted information artifacts created some problems (e.g., while users wanted to inform the world of what was really going on, they didn't want to spam their non-Persian speaking friends with news and videos that were in Persian, or expose the younger audience to graphic images of people getting beaten or shot). This example clearly shows how lack of support for fine-grained control (H1) and relationship control (H6) can hinder the usefulness of the tool in a real world situation.

- ***Example 2:***

Facebook conveys a false sense of privacy over one's shared photos, whereas in reality, it provides none: while a user can set the visibility of a photo album to friends-only, if any of those "friends" leave a comment on any photo, the whole album gets exposed to that friend's network of friends. The mere act of tagging someone in a photo has the same effect (i.e., making the whole album visible to the taggee's friends' network, even if they don't belong to the original audience group set by the album creator). There is however, a workaround for the latter case, which is that the taggee can limit visibility of tagged photos of him/her to certain audiences (through the privacy setting page). However, this solution is often not clear to users. Likewise, if I set the visibility of tagged photos of me to "everyone" and my friend tags me in a photo of hers that she has made visible only to her friends, we have two conflicting policies for a certain photo which Facebook resolves by adopting the less restrictive one (mine), thus violating my friend's privacy policy. While the original problem is the result of Facebook failure in performing appropriate conflicting resolution, the concept of ownership control (H3) would have provided users with some control over the situation (e.g., allowing them to limit who can tag them in a photo or leave a comment on their posted photos). This is also a clear example of ignoring the clarity criteria (H8), as Facebook does not convey to users that the privacy settings on their photo albums can easily get overwritten (as a consequence of someone leaving a comment on them, for example, over which users have no control).

8.3 Research Contributions

The research presented in this dissertation has made several contributions to the fields of Usable Privacy and Security, Social Computing, Computer-Supported Cooperative Work (CSCW), Personal/Group Information Management (PIM/GIM), and Human Computer Interaction (HCI). Some of these contributions have been presented, in whole or in part, in earlier publications. A detailed list of the publications and presentations generated as a result of this dissertation research can be found in Appendix C. In this section, we describe the theoretical, applied, and methodological contributions that resulted from each phase of this research.

8.3.1 Contributions from Phase One: An In-Depth Study of Privacy in the Social Software Domain

Privacy is an inherently challenging area of research in the sense that it can be difficult to elicit participants' privacy concerns in a research setting. Also, while research on social software systems is growing, there have been very few scholarly investigations of selective information sharing and privacy issues in these types of environments. From this perspective, one of the contributions of our study is that it adds to the relatively sparse, albeit growing, set of published studies of privacy issues in social software literature.

Furthermore, we had the rare opportunity of conducting our investigation with informants who had used our system of choice for information sharing over a significant time. As such, our participants were able to provide us with richer insights compared to participants who briefly experience with a system. Employing the grounded theory data coding practices to analyze the large volumes of data that was collected during the study enabled us to go beyond merely reporting the anecdotal evidence of users' experiences, by constantly addressing the concepts that are represented by each piece of data. Drawing on the rich data from the experiences of our long-term users allowed us to generate a grounded understanding of users' information sharing behavior and the factors that affect their decision regarding information disclosure in an open online environment. Such an understanding has been absent from the research and practice discourses on privacy management mechanisms in social software systems.

8.3.2 Contributions from Phase Two: Modeling Information Sharing Behavior in SPIMS

Before proposing design guidelines for privacy management in SPIM, it was important to understand users' information sharing patterns that might impact system design, so that our solutions were grounded appropriately. Studying users' perspective on personal privacy in social software systems allowed us to contribute both to theory, and to the design of systems that support end-user information sharing in social contexts. The resulted grounded theory provided the theoretical basis for a conceptual model of information privacy in SPIMS, and led to the development of design guidelines and a framework for designing privacy management mechanisms in social software systems.

The conceptual model of users' information sharing behavior in SPIM as presented by the concept map in chapter 4 provided a rich view of personal privacy concerns in a social domain. This model constitutes a theoretical contribution in that it is unique in presenting a comprehensive picture of the various factors that impact the process of selective information sharing; including incentives for use of social software for information management and sharing, privacy concerns, needs, and expectations, and strategies to achieve the desired level of privacy. Furthermore, the classification of privacy factors that affect users' decision regarding sharing information with various audiences in the SPIM domain presents an empirical contribution, as it provides a high-level overview of privacy factors that specifically pertain to the social software domain and as such, add substantive content to overall understanding of personal privacy issues in this domain.

The conceptual model can be used as a guide for future studies of personal privacy concerns and strategies in social software systems, and may also be of benefit to researchers investigating other privacy domains; particularly those areas where privacy depends on users' preferences (as opposed to organizational rules) and/or areas where frequent changes in artifacts, audiences, or context of sharing are expected. Furthermore, as shown by our attempts at developing a design framework in chapter 5 and a privacy management mechanism based on the proposed framework in chapter 6, the model can be used to inform practical privacy management solutions for social software domain.

Another theoretical contribution comes from the development of several design heuristics and a general framework for usable privacy in SPIM. The innovative framework constitutes a theoretical contribution in that it provides a high level categorization of types of controls that are needed for supporting personal privacy in SPIM whose structure is informed by empirical data from the research presented in this thesis. The design heuristics are important because they break the complex notion of privacy into tangible technical elements for design and as such, provide great benefit to other researchers studying privacy in various domains; since they constitute a framework on which other researchers and practitioners can build to develop recommendations and implications for the future design, development, and evaluation of tools to support personal privacy in these domains. These types of contributions are often relied upon in many areas of research, such as privacy, HCI-SEC, and PIM/GIM research communities. As evidence, the three previous privacy frameworks that heavily influenced our work (Belloti and Sellen 1993, Adams 1999a, Adams 1999b, Adams and Sasse 2001, Lederer et al. 2003) have over 700 citations combined by researchers studying a wide variety of research topics.

8.3.3 Contributions from Phase Three: Design and Implementation of OpnTag

While many times empirical studies stop at the point of proposing guideline and implication for designs, we went one step further to investigate our privacy framework in action. The results of our exploratory study informed the design and development of OpnTag, a Social and Personal Information Management tool, whose privacy management mechanism instantiates the four user-centered privacy controls as proposed by our framework. OpnTag provides advanced group functionality which allows users to define, control, and understand various aspects of the their groups, including size, visibility, and membership model. An important aspect of this implementation was incorporation of the people-tagging functionality to enable users to categorize their social network into groups of target audiences for their information in terms of their (often changing) relationships with them; by creating one-sided, egocentric, user-defined social relationship groups to which they can grant or deny access to various pieces of information in their personal space.

Together, the combination supports a fine-grained, relationship-based, in-context, and flexible mechanism for selective information sharing.

The experience of designing and implementing a privacy management system for OpnTag based on the proposed framework as presented in chapter 6 provides an applied contribution for future design and evaluation in this area, as it can be used as a practical guide into how the various elements of our proposed framework can be incorporated into design and implementation of privacy management systems in a social software system.

The idea of implementing the relationship control component of the framework through people-tagging also presents another applied contribution. This idea has so far been presented to the research community in two publications (an extended abstract at CHI 2008 and a full paper at GROUP 2009). In both cases, we have received very positive feedback from the anonymous reviewers who commented on the idea as being "novel", "promising", and "a lightweight and user-centered mechanism for managing access rights". One reviewer commented that "The approach of tagging people (as a well known technique for users in social systems) seems to be very useful as it reduces the effort for handling the access privileges and at the same time, allows users to configure the access rights to their information based on their personal needs."

8.3.4 Contributions from Phase Four: Empirical Evaluation of OpnTag's Privacy System

Finally, our investigation of the use of various aspects of the OpnTag's privacy management mechanism as presented in chapter 7 provides an applied contribution for other researchers and practitioners in the field. While support for these functionalities were informed by the privacy requirements we observed during our study, the results of the preliminary empirical evaluation provided evidence that the proposed privacy framework is flexible enough to meet users' varying information sharing needs in terms of both utility and usability, and that the underlying cognitive model behind OpnTag was a suitable match for users' mental model of information interaction.

Our findings also provided initial validation of the feasibility of the people-tagging concept and provided interesting insight into users' ideas regarding various aspects of the feature, including suitability of visibility options for people tags, level of control on the incoming tags, overall usability, and how to make the feature a more viable option for everyday privacy management.

Our evaluation process identified a number of areas where improvements might further increase usefulness and usability, including a more clear feedback mechanism, improving the interface and opportunities for people-tagging, and a better navigation mechanism between pages. The discussing on the challenges of the evaluation process presented in chapter 7 can be helpful to other researchers in recognizing potential difficulties that may arise in privacy evaluation. Addressing the ways we dealt with these challenges can help identify potential strategies to employ to overcome such challenges. During the process of designing our evaluation study, we learnt a great deal from the experiences of other researchers who had dealt with similar issues before. As such, we hope that our discussion of the challenges and limitations associated with the certain evaluation strategy we employed will provide insight to future researchers in the field as well.

8.4 Future Directions

In this section we outline the major areas of future work that are the natural next steps to the research presented in this dissertation. We will explore potential future directions for each of the four phases of the research separately.

8.4.1 Extensions to the Grounded Theory Study: Site Spreading & Comebacks

As discussed in chapter 5, one possible direction for extending our grounded theory study is to apply the practice of "site spreading", by complementing the results of this study with results from the study of a possibly wider and more diverse group. A possible direction would be to approach user communities of different demographics and/or using other social software systems, to see whether our findings are valid in supporting their goals and views as well. Moreover, since the kind of analysis, synthesis and reflection that are so apparent in

these social-personal information management systems are also a fundamental part of any knowledge-oriented work practice, another interesting future direction would be to extend this study into work environments (i.e., moving from Web 2.0 to Enterprise 2.0) where the cooperative-competitive balances are likely to be different.

Another potential future direction to a grounded theory study is to explore what Glaser calls "comebacks" (Glaser 1998: 200). Comebacks are categories emerged from data analyses that have less significance and less relevance to the core category, but are potentially interesting to be investigated in more depth. Since the main focus of our grounded theory study was on the core category of users' information sharing behavior, not all of the sub-core categories in our concept map were explored to full extent and may thus be considered comebacks. This includes privacy needs, concerns, and strategies. Further attempt to extensively densify these categories in terms of sub-categories and properties presents interesting and potentially fruitful topics of research for future grounded theory studies.

8.4.2 Extensions to the Privacy Framework: Investigating Individual Elements & Their Relative Weight

An immediate extension of this research would be to apply the proposed privacy framework to other information intensive environments, to examine whether it improves group and collaborative information interaction. One possible direction would be investigating how each of the various elements of the framework affect individual and group behavior with respect to sharing personal (documents, content tags, bookmarks) and social (profile elements, social networks, people-tags) artifacts. Another research direction would be measuring the relative weight of each factor on users' behavior and examining the inter-relationship of factors.

8.4.3 Extensions to OpnTag: Other Directions for the People-Tagging Concept

One of the main contributions of this work was exploring the idea of tagging people for categorizing one's social network into relationship groups of non-equal weights. In this dissertation, we primarily experimented with the idea of people-tagging for the purpose of

selective information sharing. However, people-tagging is a rich concept, which, combined with the powerful infrastructure provided by social systems, can be used for investigating other issues in the context of personal and social information management. Here are a few directions we believe are interesting to explore in the future to enhance the people-tagging functionality in OpnTag.

8.4.3.1 People-tagging for Groups

One feature that might potentially be useful is enabling people-tagging for groups in addition to individuals, i.e., people-tags applied by all members of a certain classic group. For example, "Corporate X's Customers" represents a situation where such functionality might be useful. All corporate X's members can collectively tag users as their customers, and the people-tag group created as a result could be useful to all team members for spreading information, etc. People tagged with the same tag will probably not have much to share, so it makes sense to make the group tag private to the group. There is a potential problem with this approach, though, as this will create a hole in preventing the anti-social tagging: it will be possible to make a group tag visible to all members of a group without the taggee being aware of the tag.

8.4.3.2 Combination of Two or More People-tags

Another direction to explore in the future is enabling combination of two or more people-tag groups to be used as an access control option; i.e. making memos visible to "friends & family", which would include all people tagged as "friends" plus all people tagged as "family". This feature can be supported by defining rules on people-tags, which is a potential research project by itself.

8.4.3.3 Exploring Transitive Trust

Another issue that would be interesting to study is how the assessment of one's trusted network of a person affects his/her idea of that person. This question can be explored by a field study of people-tags applied by someone's immediate connections to a certain user and their affects on this user's people-tags. The study can be conducted in the OpnTag community, which is open source and publicly available and has a user population of over 100 at this time. The study will theoretically be situated in the literature on transitive

trust, and the results will have implications for the design of systems for identifying trends and/or experts within the context of social systems.

8.4.4 Extensions to the Evaluation Study: Exploring Privacy Evaluation Methodologies

One of the main challenges of this research was evaluating the proposed privacy framework. As discussed in chapter 7, there is recognition within the field of Human-Computer Interaction that there are inherent difficulties associated with the problem of privacy evaluation. First, privacy is a secondary task and often a trade-off with functionality, performance, or convenience. As such, it is hard to define particular metrics for usability of privacy separately from the usability of the tool (i.e., what should actually be measured?). Second, privacy concerns are highly nuanced and contextual and as such, it is hard to design tasks with privacy elements artificially that resonate with users real life tasks. Furthermore, fabricated tasks may represent an incomplete thread model and as a result, incomplete evaluation. On the other hand, privacy revealing data that are generated in real-life situations and can shed some light on users' actual privacy behavior (e.g., data logs), are considered sensitive data and there are often ethical barriers associated with accessing them. A very useful future direction for the work presented in this thesis would be to address some of the question raised by our research in this area, including how to design sample privacy scenarios for laboratory studies of privacy that encompass all aspects of the design and at the same time, realistically evoke users' personal motivations; how to define criteria for determining suitability of a certain privacy evaluation method in a certain setting; and how to combine various evaluation methodologies (e.g., lab and field studies) to get the most reliable results.

Our results also indicated that a more ecologically valid evaluation of OpnTag would be necessary. One future direction would be to conduct a longitudinal evaluation in the field once sufficient data from the active user community of OpnTag becomes available.

8.5 Conclusions

Although the use of social software systems for social and personal information management has moved from leading edge to mainstream over the past few years, it is still in

the early-adopter phase. We believe proper support for privacy management is a key point in enabling incorporation of these technologies into users' everyday information practices. By providing support for selective information disclosure that maps to users' mental model of information privacy, social software systems will become a safe place for sharing ideas and reflections, and for effective collaboration. By providing the users with control over sharing of information of different degrees of sensitivity with a small or large group, we provide the opportunity for easy transition of information sharing process into productive, collaborative work.

We believe there is great potential for improving personal privacy management in social software systems, and this dissertation presents our attempts towards such goal. While our focus was on personal privacy in SPIM, our results are likely applicable to other social software systems as well. Social computing is a growing research area and social software systems continue to become more ubiquitous in our personal lives as well as our work activities. We anticipate that personal privacy concerns will increase as people continue to uses these systems for more personal, professional, and social purposes, moving between various contexts of use. This dissertation has provided several contributions of a theoretical, practical, and methodological nature that may be of use to researchers and practitioners in this area.

References

Ackerman, M. S., Cranor, L. F., and Reagle, J. (1999), Privacy in e-commerce: examining user scenarios and privacy preferences. In Proceedings of the 1st ACM Conference on Electronic Commerce, Denver, Colorado, United States, November 03 - 05, 1999.

Ackerman, M. S. (2000), The Intellectual Challenge of CSCW: The Gap Between Social Requirements and Technical Feasibility, Human-Computer Interaction, Vol. 15, No. 2, pp. 179–203.

Ackerman, M. S., and Mainwaring, S. D. (2005), Privacy Issues and Human-Computer Interaction. In L. Cranor and S. Garfinkel, editors, Security and Usability, pp. 381–400. O'Reilly, 2005.

Acquisti, A. and R. Gross. (2006), Imagined Communities: Awareness, Information Sharing and Privacy on The Facebook, Proceedings of the 6th Workshop on Privacy Enhancing Technologies, Cambridge, UK, 2006.

Adams, A. (1999), Users' perception of privacy in multimedia communication, Extended Abstract, in Proceedings of CHI 1999, Pittsburgh

Adams, A. (1999), The Implications of Users' Privacy Perception on Communication and Information Privacy Policies, in Proceedings of Telecommunications Policy Research Conference, Washington DC, 1999

Adams, A.., and Sasse, M. A. (2001), Privacy in multimedia communications: protecting users not just data, in Proceedings of IMH HCI, 2001, pp. 49 – 64.

Adam, S. (2009) Facebook and Education: The Pros and Cons. The Teaching and Learning with Technology series, January 2009, available online at https://www.elearning.ubc.ca/home/index.cfm

Altman, I. (1975), The Environment and Social Behavior—Privacy, Personal Space, Territory,Crowding, Brooks/Cole Publishing Company: Monterey, CA, 1975.

Altman, I. (1977), Privacy Regulation: Culturally Universal or Culturally Specific? Journal of Social Issues, 33, 3, 1977, pp. 66–84.

Barnes, S. B. (2006). Privacy paradox: Social networking in the United States. First Monday, Vol. 11, No. 9, September 2006

Bellotti, V. and Sellen, A. (1993), Design for Privacy in Ubiquitous Computing Environments. In Proceedings of the Third Conference on European Conference on Computer-Supported Cooperative Work, Milan, Italy, September 13 - 17, 1993.

Bergman, O., Boardman, R., Gwizdka, J., and Jones. W. (2004). Personal information management SIG. In Extended Abstracts of CHI 2004. ACM Press. pp. 1598-1599

Bishop, M. (2005). Psychological acceptability revisited. In L. Cranor and S. Garfinkel, editors, Security and Usability, pp. 1-12. O'Reilly, 2005.

Boyd, D. (2006), Identity Production in a Networked Culture: Why Youth Heart MySpace. In Annual Meeting of the American Association for the Advancement of Science, St. Louis, MO. February 19, 2006. available online at http://www.danah.org/papers/AAAS2006.html.

Boyle, M. and Greenberg, S., (2005), The language of privacy: Learning from video media space analysis and design. In ACM Transactions on Computre.-Humman Interaction, 12, 2, 2005, pp 328-370.

Brunk, B. D. (2002). Understanding the privacy space. First Monday, 7, 2002. http://firstmonday.org/issues/issue7_10/brunk/

Butler, J.K. (1991). Toward understanding and measuring conditions of trust: evolution of a conditions of trust inventory. Journal of Management, Volume 17, Issue 3, (1991), pp. 643–663

Chiasson, S., van Oorschot, P.C., and Biddle, R. (2006), A Usability Study and Critique of Two Password Managers. In USENIX Security Symposium, Vancouver, Canada. July 2006.

Clark, J., van Oorschot, P. C., and Adams, C. (2007). Usability of anonymous web browsing: an examination of Tor interfaces and deployability. In Proceedings of the 3rd Symposium on Usable Privacy and Security, Pittsburgh, Pennsylvania, July 18 - 20, 2007

Consolvo, S., Smith, I. E., Matthews, T., LaMarca, A., Tabert, J. and Powledge, P. (2005). Location Disclosure to Social Relations: Why, When, & What People Want to Share. In Proc. of CHI '05, Portland, Oregon, pp 81-90.

Computer Research Association Conference on Grand Research Challenges in Information Security and Assurance (Warrenton, VA, Nov. 16–19, 2003). www.cra.org/Activities/grand.challenges/security/.

Corritore, C. L., Kracher, B., Wiedenbeck, S. (2003). On-line trust: concepts, evolving themes, a model, in International Journal of Human-Computer Studies, Volume 58, Issue 6, June 2003, pp. 737-758

Cranor, L. F. (2005), Privacy policies and privacy preferences. In L. Cranor and S. Garfinkel, editors, Security and Usability, pp. 447-472. O'Reilly, 2005.

Cranor, L. F., Guduru, P., and Arjula, M.. (2006). User interfaces for privacy agents. ACM Trans. Comput.-Hum. Interact. 13, 2 (Jun. 2006)

De Paula, R., Ding, X., Dourish, P., Nies, K., Pillet, B., Redmiles, D., Ren, J., Rode, J. and Filho, R. S. (2005). Two Experiences Designing for Effective Security. In Proceedings of Symposium On Usable Privacy and Security (SOUPS), Pittsburgh, PA, pp 25-34.

DeWitt, A. J. and Kuljis, J. (2006). Aligning usability and security: a usability study of Polaris. In Proceedings of the Second Symposium on Usable Privacy and Security, Pittsburgh, Pennsylvania, July 12 - 14, 2006.

Dourish, P., Grinter, R. E., Delgado de la Flor, J. and Joseph, M. (2004). Security in the wild: user strategies for managing security as an everyday, practical problem, in Personal and Ubiquitous Computing 8, pp. 391-401.

Egelman, S., and Kumaraguru, P. (2005). Report on DIMACS Workshop and Working Group Meeting on Usable Privacy and Security Software.. May 3, 2005. Rutgers University, New Burnswick, NJ, available online at http://dimacs.rutgers.edu/Workshops/Tools/dimacsrpt.pdf

Erickson, T. (2006), From PIM to GIM: Personal Information Management in Group Contexts, in Communications of the ACM, 49(1) 2006.

Farrell, S., Lau, T. (2006). Fringe contacts: People-tagging for the Enterprise. In Workshop on Collaborative Web Tagging, WWW 2006.

Farrell, S. Lau, T., Wilcox, E., and Muller, M. (2007a). Socially Augmenting Employee Profiles with People-Tagging, In Proc. In proceedings of UIST 2007, ACM Press (2007).

Farrell, S., Lau, T., Nusser, S. (2007 b). Building Communities with People-Tags. In Proceedings of INTERACT 2007, Springer.

Fuller, S. (2002). Knowledge management foundations, Boston: Butterworth-Heinemann

Glaser, B. G., & Strauss, A. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago, Illinois: Aldine.

Glaser, B. G. (1978). Theoretical Sensitivity. Mill Valley, California: Sociology Press.

Glaser, B. G. (1992). Emergence vs. Forcing: Basics of Grounded Theory Analysis. Mill Valley, California: Sociology Press.

Glaser, B. G. (1998). Doing Grounded Theory: Issues and Discussions. Mill Valley, California: Sociology Press.

Goffman, E. (1959), The Presentation of Self in Everyday Life. Garden City, New York, Doubleday Anchor Books.

Govani, T., Pashley, H. (2007), Student Awareness of the Privacy Implications When Using Facebook, Unpublished manuscript retrieved September, 2007 from lorrie.cranor.org

Grabner-Kräuter, S., and Kaluscha. E.A.. (2203). Empirical research in on-line trust: a review and critical assessment , International Journal of Human-Computer Studies, Volume 58, Issue 6, June 2003, pp. 783-812

Gross, R., Acquisti, A., and Heinz, H. J. (2005), Information Revelation and Privacy in Online Social Networks. In Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society, Alexandria, VA, USA, November 07 - 07, 2005.

Hart, M., Johnson, R., and Stent, A. (2007). More Content - Less Control: Access Control in the Web 2.0., in Web 2.0 Security and Privacy Workshop 2007, Oakland, California

Hartson, H. R. (2003). Cognitive, physical, sensory, and functional affordances in interaction design, Behaviour and Information Technology 22(5), 2003, pp. 315-338

Hawkey, K. and Inkpen, K. M. (2007). PrivateBits: managing visual privacy in web browsers. In Proceedings of Graphics interface 2007, Montreal, Canada, May 28 - 30, 2007

Hong, J. I., Ng, J. D., Lederer, S., and Landay, J. A. (2004), Privacy risk models for designing privacy-sensitive ubiquitous computing systems. In Proceedings of the 5th Conference on Designing interactive Systems: Processes, Practices, Methods, and Techniques, Cambridge, MA, USA, August 01 - 04, 2004.

Hsu, M.H., Ju, T. L., Yen, C., and Chang, C. (2007). Knowledge sharing behavior in virtual communities: The relationship between trust, self-efficacy, and outcome expectations, International Journal of Human-Computer Studies, Volume 65, Issue 2, February 2007, pp. 153-169

Iachello, G. and Hong, J. (2007), End-user privacy in human-computer interaction. In Foundations and Trends in Human-Computer Interact. 1, 1, 2007, pp 1-137.

James, R., Kim, W. T., McDonald, A. M., and McGuire, R. (2007). A usability evaluation of a home monitoring system. In Proceedings of the 3rd Symposium on Usable Privacy and Security, Pittsburgh, Pennsylvania, July 18 - 20, 2007

Jarvenpaa, S.L., Knoll, K., and Leidner , D.E. (1998). Is anybody out there? Antecedents of trust in global virtual teams. Journal of Management Information Systems Volume 14, Issue 4, (1998), pp. 29–64

Jensen, C., Potts, C. and Jensen, C. (2005). Privacy practices of Internet users: Self-reports versus observed behavior. International Journal of Human-Computer Studies, Volume 63, Issues 1-2, pp. 203-227

Karat, Clare-Marie N., Karat, J., and Brodie, Carolyn A. (2007), Management of Personal Information Disclosure: The Interdependence of Privacy, Security, and Trust. In

Personal Information Management: Challenges and Opportunities, by William Jones and Jaime Teevan, University of Washington Press, 2007.

Karat, Clare-Marie N., Brodie, Carolyn A., and Karat, J. (2005), Usability Design and Evaluation for Privacy and Security Solutions. In Designing Secure Systems that People Can Use, by Lorrie Cranor and Simson Garfinkle, O'Reilly and Associates, 2005.

Karat, Clare-Marie N., Brodie, Carolyn A., and Karat, J. (2006), Usable Privacy and Security for Personal Information Management. In Communications of the ACM, 49(1) 2006, pp. 56-57.

Karat, Clare-Marie N., Karat, J., and Brodie, Carolyn A. (2005), Why HCI Research in Privacy and Security is Critical Now. International Journal of Human Computer Studies 62(1-2), April 2005, pp 1-5.

Kumar, R. (2008) One-In-Five Employers Use Social Networking Sites To Screen Applicants, available online at http://www.webguild.org/2008/09/one-in-five-employers-use-social-networking-sites-to-screen-applicants.php, September 11, 2008

Lansdale, M. (1988). The psychology of personal information management. Applied Ergonomics, Vol.19, No 1, pp.55–66

Lau, T., Etzioni, O., and Weld, D. S. (1999), Privacy interfaces for information management. Communications of the ACM, 42, 10, 1999, pp 88-94.

Lederer, S., Mankoff, J. and Dey, A. K. (2003), Towards a Deconstruction of the Privacy Space. In UBICOMP 2003 Workshop on Ubicomp Communities: Privacy as Boundary Negotiation, 2003.

Lederer, S., Mankoff, J., and Dey, A. K. (2003), Who wants to know what when? privacy preference determinants in ubiquitous computing. In CHI '03 Extended Abstracts on Human Factors in Computing Systems, Ft. Lauderdale, Florida, USA, April 05 - 10, 2003.

Lederer, S., Hong, I., Dey, K., and Landay, A. (2004), Personal privacy through understanding and action: five pitfalls for designers. Personal and Ubiquitous Computing, 8, 6, 2004, pp 440-454.

Locke, K. (2001). Grounded Theory in Management Research, Sage publications, London

Lutters, W., Ackerman, M.S., Zhou, X. (2007), Group Information Management, In In Personal Information Management: Challenges and Opportunities, by William Jones and Jaime Teevan, University of Washington Press, 2007.

Mayer, R.C., Davis, J.H., and Schoorman, F.D. (1995). An integrative model of organizational trust. Academy of Management Review Volume 20, Issue 3 (1995), pp. 709–734

McGrenere, J., and Ho, W. (2000). Affordances: Clarifying and Evolving a Concept. In Proceedings of Graphics Interface 2000, May 15-17, 2000, Montreal, Quebec, Canada. pp.179-186

Nielsen, J. (1992). The Usability Engineering Life Cycle. Computer Vol. 25, No. 3, March 1992), pp. 12-22

Nielsen, J., Clemmensen, T., Yssing, C. (2002). Getting access to what goes on in people's heads? - Reflections on the think-aloud technique. NordiCHI, October 2002.

Nielsen, J. (2004), http://www.useit.com/alertbox/20040301.html

Nielsen, J. (2001), Usability metrics, Alertbox January 2001

Norman, Donald A. (1988): The Design of Everyday Things. New York, Doubleday

Norman, Donald A. (1999): Affordances, Conventions, and Design. In Interactions, Volume 6, Issue 3, pp. 38-41

O'Hear S. (2006). Elgg: Social Network Software for Education, ReadWriteWeb, available online at http://www.readwriteweb.com/archives/elgg.php

Olson, J. S., Grudin, J., and Horvitz, E. (2005), A study of preferences for sharing and privacy. In CHI '05 Extended Abstracts on Human Factors in Computing Systems, Portland, OR, USA, April 02 - 07, 2005.

Orlikowski, W. J. (1993). CASE Tools as Organizational Change: Investigating Incremental and Radical Changes in Systems Development, MIS Quarterly, Volume 17, September 1993, pp. 309-340

Pace, S. (2003). A grounded theory of the flow experiences of Web users International Journal of Human-Computer Studies, Volume 60, Issue 3, March 2004, pp. 327-363

Palen, L. and Dourish, P. (2003). Unpacking "Privacy" for a Networked World. In Proceedings of CHI '03, Ft. Lauderdale, FL, pp 129-136.

Pandit, N. (1998) Towards a Grounded Theory of Corporate Turnaround: A Case Study Approach, doctoral thesis, University of Manchester, UK.

Patil, S. and Kobsa, A. (2005), Uncovering privacy attitudes and practices in instant messaging. In Proceedings of the 2005 international ACM SIGGROUP Conference on Supporting Group Work, Sanibel Island, Florida, USA, November 06 - 09, 2005.

Patil, S. and Lai, J. (2005), Who gets to know what when: configuring privacy permissions in an awareness application. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Portland, Oregon, USA, April 02 - 07, 2005.

Phillips, D.J. (2004), Privacy policy and PETS: The influence of policy regimes on the development and social implications of privacy enhancing technologies. New Media Society, 6, 6, 2004, pp 691–706.

Preibusch, S., Hoser, B., Gürses, S., Berendt, B. (2007), Ubiquitous social networks' opportunities and challenges for privacy-aware user modelling, in Proceedings of the Workshop on Knowledge Discovery for Ubiquitoues User Modeling, 2007.

Ridings, C.M., Gefen, D., Arinze, B. (2002). Some antecedents and effects of trust in virtual communities, The Journal of Strategic Information Systems, Volume 11, Issues 3-4, December 2002, pp. 271-295

Rowse, D. (2006). Blog stalkers – personal safety for bloggers, available online at http://www.problogger.net/archives/2006/02/07/blog-stalkers-personal-safety-for-bloggers, February 2006

Sandhu, Ravi S., Coyne, Edward J., Feinstein, Hal L., & Youman, Charles E. (1996). Role-based access control models. In Computer. Volume 29, Number 2, 1996, pp 38-47

Saltzer, J., and Schroeder, M. (1975), The protection of information in computer systems. Proceedings of the IEEE, 1975. 63(9), pp 1278-1308.

Scott, J. (1998). Organizational Knowledge and The Intranet, Decision Support Systems Journal, Volume 23, pp. 3-17

Simonetti, E. (2004). I was fired for blogging, available online at http://news.com.com/I+was+fired+for+blogging/2010-1030 3-5490836.html, December 2004

Spiekermann, S., Grossklags, J. and Berendt, B. (2001). E-privacy in 2nd Generation E-Commerce: Privacy Preferences versus actual Behavior. In proceedings of Electronic Commerce 2001, Tampa, Florida, USA, pp. 38-47

Strauss, A., & Corbin, J. (1990). Basics of Qualitative Research, Sage Publications

Strauss, A., & Corbin, J. (1998). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Second Edition: Sage Publications

Strauss, A. (1987). Qualitative Analysis for social scientists, Cambridge University Press

Stutzman, F. (2006), An Evaluation of Identity-Sharing Behavior in Social Network Communities. In Proceedings of iDMAa and IMS Code Conference, 2006.

Tallon, P. P. and Scannell, R. (2007). Information life cycle management. In Communications of the ACM, 50, 11 (Nov. 2007), pp. 65-69

Thomas, R., Sandhu, R., (1997). Task-based authorization controls (TBAC): Models for active and enterprise oriented authorization management. In Database Security XI: Status and Prospects, North-Holland.

Tosh, D., and Werdmuller, B. (2004), ePortfolios and Weblogs: One Vision for ePortfolio Development, ePortfolio research and development community (ERADC), available online at http://www.eradc.org/papers/ePortfolio_Weblog.pdf.

Tosh, D. (2005). ePortfolio Research and Development Community (ERADC), available online at http://www.eradc.org/blog/

The Vancouver School Board Web Site, www.vsb.bc.ca/vsbprograms/kto12/apfe/VSBUBCTransition.htm

Westin. A. F. (1967), Privacy and Freedom, New York, NY: Atheneum, 1967.

Westin. A. F. (1991), Harris-Equifax Consumer Privacy Survey 1991. Atlanta: Equifax Inc.

Westin, A. F. (2003). Social and Political Dimensions of Privacy, Journal of Social Issues, 59(2), pp 431-453.

Whalen, T., Gates, C. (2005), Private Lives: User attitudes towards personal information on the web, poster, in SOUPS 2005

Wharton, C., Rieman, J., Lewis, C., and Polson, P., The cognitive walkthrough method: A practitioner's guide. In J. Nielsen and R. L. Mack, editors, Usability Inspection Methods, pp. 84-89. Wiley & Sons, 1994.

Whittaker, S., and Sidner, C. (1996). Email overload: exploring personal information management of email, in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground, Vancouver, British Columbia, Canada, April 13 - 18, 1996

Whitten, A., and Tygar, J.D. (1999), Why Johnny Can't Encrypt: A Usability Case Study of PGP 5.0, in Proceedings of the 8th USENIX Security Symposium, August 1999.

Whitten, A., and Tygar, J.D. (1998), Usability of Security: A Case Study. Technical Report CMU-CS-98-155, Carnegie Mellon University School of Computer Science, December 1998.

Yee, K.-P. (2002), User Interaction Design for Secure Systems, University of California Berkeley Tech report, May 2002, Tech Report CSD-02-1184, available online at http://www.sims.berkeley.edu/~ping/sid/

Appendix A: OpnTag's Field Trials

Since its first release in June 2006, OpnTag has been adopted by over 100 users. In addition to individual usage, various groups have been using OpnTag for educational or organizational purposes. Here we present two experiences of deploying OpnTag in real world situations. We discuss the scenarios, the feedback, and the changes we made to the design as a response.

1. CrowdTrust

CrowdTrust[13] is a small start-up focused on creating collective intelligence solutions and active in the development of OpnTag. The company has 8 members, including designers, developers, marketers, and CEO. The CrowdTrust team has been using OpnTag for information management and sharing within the organization for over two years. Separate groups have been created to serve different information sharing purposes: the "CrowdTrust" group is the main group that all the corporate staff are a member of. Issues relevant to all team members such as meeting plans and agenda, meeting minutes, competing companies, similar products, and potential customers are shared between staff by creating memos either in the CrowdTrust space, i.e. in situations where any CrowdTrust member is expected to contribute; or in member's personal space visible to CrowdTrust, so that other CrowdTrust members can also see it. There are also two other groups, each with a selected subset of corporate staff as members: "CrowdTrust Help", used by developers for communicating help materials on company's product to the customers; and "CrowdTrust Board", used by company board members for discussing management issues.

The CrowdTrust experience has helped us clarify which features of the application users appreciate the most and which parts of the interface are confusing to them. Since some of CrowdTrust members were also OpnTag's developers, this pilot group was not a candidate for a formal usability study of OpnTag. However, the level of users' participation and contribution shows that for most of the group members, engaging with the tool has become an essential daily activity and valuable resource; which in turn suggests that from

[13] http://crowdtrust.com/

users' perspective, engaging with both the privacy control and tag-based organization is probably simple and not a large barrier to usability.

Although cases of confusion were few, most centered around the implementation of the notion of audience control in OpnTag: We realized that there is need for better management of the space between *audience control*, which bounds the audience for any particular item or conversation, and *audience notification*, which makes the audience specifically aware of certain activities. As a result of this feedback, we extended the notification model to allow more specific control by both information producers and consumers of the streams of notification information managed by OpnTag.

2. ETEC522

In the fall of 2007, OpnTag was used as the main course information and interaction system for ETEC 522, an online course on educational technologies offered by the University of British Columbia. Students used it for both their own information management within the course and for conversation and sharing resources with the rest of the class. Throughout this process, we had no negative feedback with respect to the privacy or information management aspects of the system; however, the major criticisms from use in this context were centered on the management of conversation and awareness using the tool and model, which confirmed the findings from the CrowdTrust pilot experience.

At the start of the course, when students asked for the memos for the group "ETEC 522", the system would select those memos "owned" by the group. It was clear that this was inadequate, in the sense that the students expected that specifying the audience of a memo for "ETEC 522" would also have the effect of it being seen in the group "space". After this, we revisited the selection of memos considered to be part of a "space" (for an individual or group) and realized that there are various ways of both claiming a memo for oneself and providing it to a group. Currently, when visiting an individual's space the memo set includes all memos created, modified or tagged by that user. When visiting a group's space the memo set includes memos owned by the group, tagged in the group and memos made explicitly visible to the group. Moreover, OpnTag's message system notifies all members of a group

when any of these memos are created or modified. In this way, membership in the group now allows one to both contribute to and monitor the group in a variety of ways.

Appendix B: Certificates of Approval from UBC Research Ethics Board

The University of British Columbia
Office of Research Services and Administration
Behavioural Research Ethics Board

Certificate of Approval

PRINCIPAL INVESTIGATOR	DEPARTMENT	NUMBER
Iverson, L.	Electrical and Computer Eng	**B05-0414**

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT

UBC Campus ,

CO-INVESTIGATORS:

Najafian Razavi, Maryam, Electrical and Computer Eng

SPONSORING AGENCIES

TITLE :

Modeling Information Sharing Behaviour in Open Online Environments

APPROVAL DATE	TERM (YEARS)	DOCUMENTS INCLUDED IN THIS APPROVAL:
JUN 2 1 2005	1	June 10, 2005, Consent forms / Questionnaires

CERTIFICATION:

The protocol describing the above-named project has been reviewed by the Committee and the experimental procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval of the Behavioural Research Ethics Board by one of the following:
Dr. James Frankish, Chair,
Dr. Cay Holbrook, Associate Chair,
Dr. Susan Rowley, Associate Chair

This Certificate of Approval is valid for the above term provided there is no change in the experimental procedures

The University of British Columbia
Office of Research Services
Behavioural Research Ethics Board
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

CERTIFICATE OF APPROVAL - FULL BOARD

PRINCIPAL INVESTIGATOR:	INSTITUTION / DEPARTMENT:	UBC BREB NUMBER:
Lee Iverson	UBC/Applied Science/Electrical and Computer Engineering	H07-03188

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:

Institution	Site
UBC	Vancouver (excludes UBC Hospital)

Other locations where the research will be conducted:
The study can be done at any place that is convenient for the subject, including subject's home or office and reseracher's lab.

CO-INVESTIGATOR(S):

N/A

SPONSORING AGENCIES:

N/A

PROJECT TITLE:

A usability study of the privacy management mechanism in a social-personal information management (SPIM) system

REB MEETING DATE:	CERTIFICATE EXPIRY DATE:
March 13, 2008	March 13, 2009

DOCUMENTS INCLUDED IN THIS APPROVAL: **DATE APPROVED:** April 23, 2008

Document Name	Version	Date
Protocol:		
OT-ResearchProposal	N/A	February 12, 2008
Consent Forms:		
OT-ConsentForm	N/A	February 12, 2008
Advertisements:		
OT-RecruitmentFlyer	N/A	February 12, 2008
Questionnaire, Questionnaire Cover Letter, Tests:		
OT-InitialSurveyQuestions	N/A	February 12, 2008
OT-InterViewQuestions	N/A	February 12, 2008
OT-Tasks	N/A	February 12, 2008

The application for ethical review and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval is issued on behalf of the Behavioural Research Ethics Board and signed electronically by one of the following:

Dr. M. Judith Lynam, Chair
Dr. Ken Craig, Chair
Dr. Jim Rupert, Associate Chair
Dr. Laurie Ford, Associate Chair
Dr. Daniel Salhani, Associate Chair
Dr. Anita Ho, Associate Chair

Appendix C: Previously Published Work

Much of the research presented in this thesis has been published in part in peer reviewed conference proceedings and journals. Here is a list of previous publications related to the research covered in this dissertation.

M. N. Razavi, L. Iverson (2009). Improving Personal Privacy in Social Systems with People-tagging, In Proceedings of the ACM GROUP '09 on Supporting Group Work, Sanibel Island, Florida, USA, May 10 –13, 2009

M. N. Razavi, L. Iverson (2008). Supporting Selective Information Sharing with People-tagging, In Proceedings of the ACM CHI '08 Extended Abstracts on Human Factors in Computing Systems, Florence, Italy, April 5 - 10, 2008

L. Iverson, M. N. Razavi, V. Mirzaee (2008). Personal and Social Information Management with OpnTag, In Proceedings of ICEIS '08 International Conference on Enterprise Information Systems, Barcelona, Spain, June 12 – 16, 2008

V. Mirzaee, M. N. Razavi, L. Iverson (2008). Improving Personal and Social Information Management with Advanced Tagging, In proceedings of the CAIS/ACSI '08 conference of Canadian Association for Information Science, Vancouver, Canada, June 5- 7, 2008

M. N. Razavi, L. Iverson (2007). Designing for Privacy in Personal Learning Spaces, In New Review of Hypermedia and Multimedia, Special Issue on Studying the Users of Digital Education Technologies: Theories, Methods, and Analytical Approaches, Vol. 13, No. 2, December 2007, pp: 163-185

M. N. Razavi, L. Iverson (2007). A Framework for Privacy Support in Group Information Management Systems, In Proceedings of the ACM GROUP '07 Doctoral Consortium on Supporting Group Work, Sanibel Island, Florida, USA, November 4 - 7, 2007

M. N. Razavi, L. Iverson (2006). Design Guidelines for an Information Privacy Management System for Personal Learning Spaces, In Proceedings of the eLearn '06 Conference *on E-Learning in Corporate, Government, Healthcare, and Higher Education,* Honolulu, Hawaii, USA, October 4 - 7, 2006

M. N. Razavi, L. Iverson (2006). A Grounded Theory of Information Sharing Behavior in a Personal Learning Space, In Proceedings of the ACM CSCW '06 Conference on Computer Supported Cooperative Work, Banff, Alberta, Canada, November 4 - 8, 2006