Towards Usable Privacy for Social Software

Maryam Najafian Razavi
University of British Columbia
2366 Main Mall, Vancouver, BC
1-604-827-5909
Maryamr@ece.ubc.ca

Lee Iverson
University of British Columbia
2332 Main Mall, Vancouver, BC
1-604-822-3381
Leei@ece.ubc.ca

ABSTRACT
As the use of social software for various personal and professional purposes gets widespread, the issue of providing usable support for managing access to the vast amount of user-created content in such an open environment becomes more of a concern. In a recent work, we proposed a grounded theory of how users manage privacy of their information in a typical social software where information sharing and online collaboration is encouraged and users are producers as well as consumers of information. The grounded theory suggests that users' preferences regarding privacy of their artifacts in such an environment depends on a number of factors, including the current stage in the artifacts 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. In this paper we discuss how the results of the theory can be translated into guidelines that inform design of more usable privacy management mechanisms for social software. We also discuss some of the existing access control models and their insufficiencies in supporting specific privacy requirements in this particular context. Based on our findings, we then propose a privacy control system to provide more usable privacy management for social software.

Categories and Subject Descriptors
H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces – *Theory and model;* H.1.2 [**Models and Principles**]: User/Machine Systems – *Software Psychology.*

General Terms
Human Factors, Theory, Design, Security.

Keywords
Social software, grounded theory, information sharing, information privacy, tagging.

1. INTRODUCTION
Recent advances in the emergence and growth of Web 2.0 applications have made the users producers as well as consumers of information. Social software is a by-product of the Web 2.0 phenomenon: the second generation of internet-based services that include some form of many-to-many publishing (such as social networking, social book marking, weblogs, wikis, and ePortfolios), enhanced organization and categorization of content, and most importantly, encourage generation and distribution of Web content. The Web 2.0 phenomenon is characterized by open communication, decentralization of authority, and freedom to share and re-use [12]. Although sharing information is one of the major motivations behind the use of social software, not everything is to be shared with everyone. While use of social software has moved within the reach of non-technical mainstream, managing selective sharing of published information still requires expertise. Lack of proper access management mechanisms has been identified as one of the major impediments in the wide spread use of social software despite its obvious benefits [22, 3]. Research into access management has generally concentrated on the needs of organizations or distributed systems. However, privacy requirements for user-created content in social software are different from data protection requirements in organizational databases and operating systems: social software is often used for both social activity and engaged work practices and as such, provides users with the opportunity to include a wide variety of artifacts in their environment, from work related documents, to personal opinions expressed in a weblog, to bookmarks and personal collections. Over time, this aggregation of ones' information could present a rich view of his experiences and skills in the form of a searchable life log. This creates a persistent, long-lived online identity for the user, to which he may wish to expose different views to various audiences. The shared artifact and the groups in which it is shared could both be dynamic, and preferences regarding sharing the artifact within a group have to be flexible enough to accommodate frequent changes. Information is not necessarily shared with identifiable, accountable individuals, and sharing might happen in various contexts, for example competitive as well as collaborative. Traditional access control models generally address the problem of enforcing well-defined rules set by central authorities and do not account for the dynamic nature of personal preferences as required by social software. On the other hand, access control models that are proposed in the literature for groupware (e.g. [31, 6]) tend to be rather complex and leave the important question unanswered whether users will be able to cope with this complexity. Thus, there is a clear need for privacy management models that address specific privacy requirements in social software and yet, are easy enough for non-technical users to understand and use. Our research has been motivated by this need: we believe that in order to be usable, privacy mechanisms must reflect users' needs and must be built based on users' mental model of information

privacy. To this end, we recently conducted a grounded theory study of users' information sharing behavior in social software to identify users' privacy needs. The study showed that users' privacy preferences depend on a number of factors, including the current stage in the artifacts 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. In this paper we discuss how the results of the theory can be translated into guidelines that inform design of more usable privacy management mechanisms for social software. We also discuss some of the existing access control models and their insufficiencies in supporting specific privacy requirements in this particular context. Based on our findings, we then propose a privacy control system to provide more usable privacy management for social software.

2. RELATED WORK

In recent years, use of social software has moved from niche phenomenon to mass adoption [10, 22]. 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.

Gross et. al [10] report on a study on patterns of information revelation in online social networks and their privacy implications. Their results are based on actual field data from more than 4000 users of Facebook. They report that 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.

Researchers have also studies users' attitude towards revealing information in several other contexts, including work place, online services, and location-aware mobile services. Olson et. al. [26] take a quantitative approach in conducting an in-depth survey of people's willingness to share a range of everyday information (such as web sites they visit or their health status) with various others, including family members or co-workers. They point out that whether data is anonymized or can be tied directly to people plays a major role in people's willingness to disclose. Other relevant factors reported include general attitude towards privacy (from privacy unconcerned, to privacy pragmatist, to privacy fundamentalist), and personal judgment regarding "appropriateness" (i.e. relevance) of sharing certain information with certain groups.

In another work, Patil et. al [27] conduct 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 the workplace. They identify which clusters of awareness information are more likely to be shared with whom and in what context (i.e. "team members" received comparable level of awareness sharing with "family" during work hours).

Whalen and Gates [35] 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.

Recent works in the area of knowledge management (KM) have also recognized the need to improve people's ability to control who sees what in their personal information. Erickson [7] 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. In his article on GIM, Group Information Management, he identifies many research questions that need to be explored, including 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.

3. THE STUDY

In the view of the difficulties that HCI researchers have encountered in locating places where the context of privacy can be better understood, we undertook a qualitative study with the aim of identifying privacy needs, concerns, and challenges in social software from users' perspective. This section describes the study and the theory that was derived from it.

3.1 Methodology

The research method that was selected for the study was grounded theory [9, 23]; a primarily inductive investigation process in which the researcher aims to formulate a small-scale, focused theory that is derived from the continuous interplay between analysis and data collection. The purpose of the grounded theory method is building theory, not testing theory; therefore theory concepts are suggested, not proven. The resulting theory is an integrated set of propositions that are grounded in evidence but not traditional quantitative "findings". Therefore, rather than starting with a preconceived theory that needs to be proven, the researcher begins with a general area of study and allows the theory to emerge from the data. The rationale for this approach (as described by Glasser and Strauss [9]) is that the theory that is derived from data is more likely to resemble reality, as opposed to theory that is derived by putting together a series of concepts from solely speculation on how one "thinks" things should work.

3.2 Data Collection

The data that was gathered for this study primarily consisted of semi-structured in-depth interviews with 12 participants who were using a social software system with integrated weblogs, e-portfolio and social networking functionality for over a year. As such, they had a rich experience in using various features of the tools, which was an essential requirement for the emergence of the issue of privacy preferences and selective disclosure of information. The two main reasons that motivated our choice of environment were that it supports creation of ad-hoc groups and communities where privacy issues potentially arise, and that it provides reasonable support for privacy management at a fairly granular level which most other tools simply don't have. Nevertheless, the environment was just used 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 the interviews on the general area of information sharing behavior in the context of social software, rather than limiting the discussion to specific characteristics of the tool.

The gender balance of the selected participants was evenly split; there were 6 females and 6 males. Participants were selected according to their potential for developing new insights using a procedure known as theoretical sampling. 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. Our sampling continued until the study achieved theoretical saturation, the point at which additional data was no longer adding to the concepts and relationships being developed.

3.3 Data Analysis

Our grounded theory was formulated from data using a constant comparative method of analysis with three stages: The first stage of analysis, called open coding, involved breaking the interview transcripts down into discrete incidents (i.e. ideas, events, and actions) which were then closely examined and compared for significant concepts. These concepts were abstractions in the sense that they represented an aggregated account of many participants' story. We used the qualitative analysis software NVivo at this stage to label incidents in the data with code words and to write theoretical notes that captured momentary thoughts.

The second stage of analysis, called theoretical coding, involved taking the concepts that emerged during open coding and reassembling them with propositions about the relationships between those concepts. The relationships, like the concepts, emerged from the data through a process of constant comparison. Neither the concepts nor the relationships were preconceived or forced upon the data.

The third stage of analysis, called selective coding, involved delimiting coding to only those concepts and relationships that related to the core explanatory concept reflecting the main theme of the study. At the end of this stage, we were able to produce a more focused theory with a smaller set of high-level concepts. A more comprehensive explanation of the study and the overall theory is provided in [28].

4. THE GROUNDED THEORY

Two main themes emerged from our grounded theory study: First, we determined that privacy *is* a main concern of users of such systems, and second, we identified factors that affect users' privacy preferences. The next subsections present a more detailed description of each theme.

4.1 Centrality of Privacy as a Concern

The concept map in figure 1 highlights the centrality of privacy as a concern for our subjects. Although the tool was primarily introduced to participants for educational purposes, they were also using it for interacting with each other (social networking), writing personal reflections (weblogging), and showcasing samples of their creative works. This confirms that as with other computer-mediated social technologies (e.g. email [36]), when given a rich environment that provides support for both work related and social activities, user communities will adapt it for more purposes than was initially conceived. Many participants mentioned they see potential benefits in using the tool, such as having all their information in one central place and over the Internet, where they can refer potential audiences to view things rather than having to send them stuff individually. Many also mentioned that it helps them keep track and reflect on their improvement over time and in some cases, get unbiased feedback on their creative artifacts from a community of people who share the same interest.

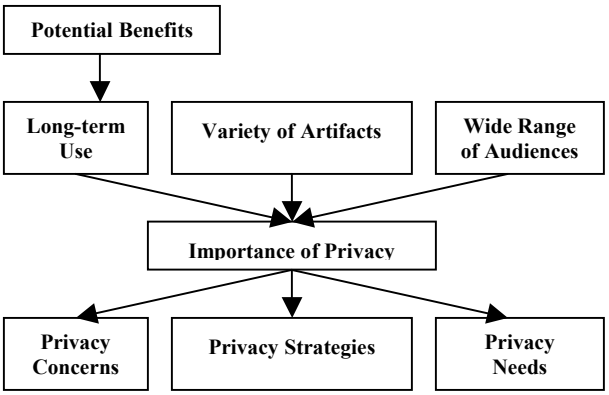

Figure 1. Centrality of privacy as a concern

However, most participants agreed that the tool did not provide sufficient support for privacy control for all their various needs: they had a wide variety of artifacts (ranging from personal profile and reflections to educational material and creative work) targeted to a different groups of audiences (teachers, classmates, friends, various communities) that were not necessarily static. These specific characteristics of the environment plus the tendency for long-term use, gave rise to a need for selective sharing. Two major concerns that were brought up by most users were the concern over loss of control and credit (mostly for creative and educational artifacts), and the concern that their work might be interpreted out of context (mostly for personal opinions and reflections). Because of these concerns, many of the participants employed certain strategies to achieve the desired level of privacy: Some were using other platforms with better privacy management mechanisms for their more private content; others had chosen to stick to one platform, but write their more private content in some sort of a "code language" so that it was meaningless to anyone other than the writer himself; and some had decided not to provide a link to certain material from places where their real identity is known. These strategies pointed to the fact that there are certain privacy needs of the users that the tool fails to support. Almost all participants mentioned that a better privacy management mechanism would improve their experience with the tool.

4.2 Privacy Factors

The second theme that emerged from our grounded theory were the factors that affect privacy of information from users' perspective as shown in figure 2. Our study showed that rather than a binary scale of public vs. private, users' judgments of privacy of resources often reflects a transition from private, to semi-private/restricted share, to public, depending on the state of the information, the receiver, and the context of sharing. The study showed that users have different perceptions of privacy of their artifacts in different stages of the artifact's life cycle. For example, an artifact is often considered private at the time of creation when descriptions, goals, and personal reflections are often included with the artifact. However, during the work-in-progress stage, the artifact may be shared with a restricted

audience to obtain feedback, and it then may be shared with a larger/more public audience once it is completed.

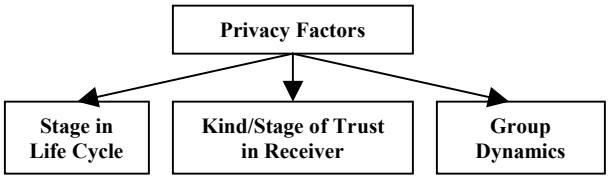

Figure 2. Factors affecting privacy preferences

We also found that users' assessment of the persons or groups who will be the receivers of information plays a strong role in making decisions about information sharing: users tend to share less with people/groups with whom they are in the initial stages of trust, and as their trust moves towards a more mature level over time, they begin to feel more comfortable and share more. The most influential factor in the information sharing attitudes of users however seems to be the dynamics of the groups or communities where the information may be shared. Our study revealed that users often hold back from sharing information in anticipation of loss of control and influence, 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.

5. TOWARDS A CONCEPTUAL MODEL OF INFORMATION PRIVACY

The main objective of a grounded theory study is to improve understanding of a phenomenon and to construct an evidence-based theoretical framework describing the phenomenon. In general, whether it is based on qualitative or quantitative evidence, a theory has both explanatory and predictive force. Whereas a theory may be initially accepted based on its explanatory force (especially if it is about something that is unexplained or insufficiently described), its perceived usefulness is determined by its predictive force. To that end theories often include a model (either formal or informal) that others can test and apply. As such, the model is expected to make predictions that can be evaluated in different situations. In this section, we extend the results of our grounded theory into a conceptual model of information privacy for social software. We first discuss how the privacy factors that were identified by the theory can be translated into requirements that should be supported by social software systems and then discuss existing access control models for their ability to support these requirements.

5.1 Analysis of Findings

The first observation that follows from our study is that users have nuanced ideas about what they want to share with whom and in what context, and they consider it a shortcoming of the tool when their desired level of privacy is not supported. In some cases privacy may even determine their choice of tool or their level of engagement with it. The fact that users try to address lack of desired level of privacy with certain strategies points out to the fact that even though users might adapt their behavior to the tool, that does not mean the tool is good enough for their purpose, which further emphasizes the need for privacy management mechanisms with better support for users' needs.

Considering the three privacy factors that emerged from the second theme, we can see that the theory suggests that in practice, users view the information sharing act as establishing and maintaining a *dynamic* sharing relationship, rather than a single event. Although the information sharing act seems like a simple and straightforward act (user shares something with a group of receivers based on their current relationship), there are various levels of dynamics to this model. Over time the artifacts might change (i.e. research results getting published, patented ideas getting approval, personal opinions reconsidered), the receiver group might change (i.e. competitors joining a group or collaborators leaving), and the relationship between the user and the receiver group might change (i.e. switching to a different project, change of affiliation). In short, all the contributing factors in users' privacy preferences can change over time, and all the three factors that were identified by our study reflect users' needs for support of these kinds of changes: 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: organizing ones' network into various groups is a way of compartmentalizing trust and audience, rather than having to deal with it on an individual basis.

5.2 User-Oriented Privacy Controls

The central structure upon which we ground our design model is the description of the kinds of control of privacy that have been shown to be necessary for social software systems for managing and sharing personal information or work products. In short, the theory suggests that users need *artifact control*, *audience control*, *relationship control*, and most importantly, *change control* for all the three of the factors outlined above.

5.2.1 Artifact Control

The principle of *artifact control* reflects the need for control of the privacy of information in terms of both individual artifacts and collections thereof. This is of course mere confirmation of the long-standing model that access rights should be associated with individual objects (e.g. files) and their collections (e.g. folders). But since social software has a different granularity and object creation model, it may be that the way in which these rights are managed to protect privacy and facilitate sharing needs to be different in some essential ways.

Our study suggests that unlike static artifacts for which the set-on-creation access management models may be sufficient, the dynamic nature of the personal artifacts that are generally disposed in social software systems calls for more fluid rights management techniques. For dynamic artifacts, users seem to dynamically match privacy and control to the artifact's degree of completion. We believe we can use this fact to reduce the up-front cost of privacy management by gathering privacy context from the environment. Since users already categorize their information for other purposes, it makes sense then to leverage these categories further by associating default access patterns with different user-defined categories. Of course, categories can be defined in various

contexts and tuned to the application. They could be established globally as a library of workflows that can be used by individuals or groups or built from the ground-up by the users and shared like other artifacts within the social software system. If we hope to provide a global library of such patterns though, it will be necessary to align the models with preexisting mental models in order to guarantee out-of-the-box usability. We suggest that providing a set of such patterns that offer both static and dynamic rights management would help give control to the users without too much overhead: once a pattern is selected for an artifact, the access restriction level of the artifact can be changed by simply selecting which stage of its life cycle the artifact is currently in.

Furthermore, categories with default access patterns can help catering to the needs of both novice and expert users by conforming to the principle of safe staging [37]: users can choose to accept the defaults while they are in the initial stages of interacting with the tool, and once they have moved to a higher level of expertise, they might decide to modify or extend the defaults to better suit their needs.

Finally, if the categories are themselves treated as resources to be shared, discussed and managed then such an evolution may actually happen on a community-by-community basis.

5.2.2 Audience Control

The principle of *audience control* reflects the observation that from a user's point-of-view, the primary concern in assessing the information sharing act is in understanding the audience that will have access to that information. From an access control point of view, this suggests that the most significant access rights to be modeled are those pertaining to the mere visibility of an artifact (e.g. does it even exist at all) and its readability (i.e. ability to access its contents). For user-oriented privacy management, we will use the term "*audience control"* to describe the ability to restrict the visibility and readability of artifacts to certain user-defined groups.

We see some of this control currently being expressed in certain social software systems, notably Del.icio.us [15], Orkut [20], and Facebook [16]. 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 Ma.gnolia [18] and Bluedot [14]) 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. In Facebook and Orkut, are services that are largely concerned with identity construction and maintenance [2]. The greater risk of exposing identity attributes to a worldwide audience has thus resulted in the deployment of a great deal of audience control for one's personal profile information. In essence, one can choose which of a variety of different categories of "friend" and "colleague" will be allowed to see any particular piece of identity attribute (e.g. phone number, address, AIM id) or posted content.

Audience control is clearly most directly related to the group and trust factors described above. In essence, the choice of audience for a particular artifact or personal attribute is primarily expressed in terms of a *group* of others who one *trusts* with that particular piece of information. While there are many models of trust in the literature, we do not depend on any one in particular. It is important to note, however, that 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 will revisit this issue when we discuss change control below.

5.2.3 Relationship Control

The principle of *relationship control* reflects the finding that many of our information sharing needs can be 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. At first blush, this seems simple and obvious, but in terms of rights management it strongly implies that the potential audience for some artifacts or attributes is likely defined in the user's own terms, and not in terms of any organizational "roles" or groups. In other words, each and every user needs the ability to define "groups" of friends or collaborators in their own terms and then to be able to use this model of their relationships with others as the basis for audience control (at minimum).

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). 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 ego-centric *groups* of users, so both of these factors are essential. Less obvious, perhaps, is the way in which this interacts with the *privacy life cycle* of artifacts. In essence, it is very likely that the best match for the assignment of audience and other rights (e.g. modification rights) to an artifact through its life cycle are via these relationship groupings, and not via traditional "group" or "role" assignments. Whereas it certainly makes sense for an organization to align access rights to organizational roles, it makes little sense for a user to align privacy rights with those organizational roles. Given that an egocentric relationship model naturally aligns with patterns of trust and information sharing for personal information, it is essential that audience and other access control rights be assignable based on these user-controlled relationship models.

5.2.4 Change Control

The principle of *change control* is something of a cross-cutting concern within the other control patterns. In essence, with social software systems one must never forget that the artifacts, audience and relationships used to define privacy and sharing patterns are dynamic. In essence, our privacy and user interaction models must reflect an assumption 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 and trust patterns within those relationships will change, and that the whatever access rights are consequences of these models must change whenever they do.

This principle then strongly suggests that a model that assigns access rights based on these factors at the time of an artifact's creation or modification will be inadequate. In essence, the access rights must track changes in whatever models are used to fulfill the above principles dynamically and visibly. This may be implemented in many different ways, but it essentially demands that either the access control regime be based on the privacy model directly or that whatever rules connect the privacy model to the access control regime be dynamic and incremental, reacting to whatever changes are made to the social parameters that define the sharing model.

This may, of course, require some rule-based system to maintain this connection (e.g. [4]), but it is likely to require significant interaction with the social software's notification system as well. For example, if we follow the user interface "principle of least surprise" [13], then when a user (A) adds some other user (B) to a relationship category, both A and B should be notified in some way of the consequences of this change (e.g. user B now has access to a new collection of information). For the initiating user (A) such a notification (or at least ability to explore the consequences of this action before it is taken) can be critical to making the decision in the first place. For user B, the granting (or restricting) of rights to a body of information is an important piece of data to be able to assess their own relationship model.

5.3 Candidate Access Control Model

We now examine some of the existing access control models and discuss their suitability to be applied to social software. For each model, we use the principles and philosophy behind the model as the basis for our discussion on its ability to support the user-defined privacy controls as discussed in the previous section.

5.3.1 RBAC

We start with RBAC [30], as one of the widely accepted best practices for managing access permissions in the literature. RBAC was originally designed for controlling access to services and resources within organizations. The main characteristic of the RBAC model that makes it a suitable candidate for use within organizations is the ability to assign enterprise-specific access permissions to organizational roles rather than individuals. As such, the success of RBAC model depends on clear assignment of roles to users, and access rights to roles, by the system administrator (thus no user-oriented artifact control or audience control is supported). The effectiveness of the model is based on the underlying assumption that there are pre-defined roles and that the role/permission association changes less frequently than user/role association (thus assuming no user-oriented relationship control and change control). While this would be a valid assumption for the organizations and commercial applications world, it is not necessarily true for social software: users of social software do not conform to an underlying organizational structure and personal information is not always shared with identifiable, accountable individual. Assigning appropriate roles to these users thus becomes an irrational and ad-hoc exercise.

Although our study suggests use of user-controlled group definition as a way of enabling users to specify their trusted audience, using roles for that matter as pertained to organizational structures would be counter-intuitive: in order to provide support for user-oriented controls, role definition and assignment need to be performed by non-technical users, as opposed to a system administrators with deep technical knowledge. Considering the dynamic nature of user-created resources, audiences, relationships, and privacy references in social software as shown by our study, this calls for frequent changes in user/role assignment that needs to be handled frequently by the user, which would be too labor-intensive and counter to RBAC philosophy. We conclude then, that RBAC is not a suitable candidate for privacy management in social software.

5.3.2 TrustBAC

Over the years, researchers have proposed various extensions to the original RBAC model to tailor it to the specific needs of certain applications. One of these extended models is TrustBAC [5]. It adds the notion of trust levels to the original model. Users are assigned to trust levels instead of roles based on a number of factors like user credentials, user behavior history, user recommendation etc. Trust levels are then assigned to roles, which are then assigned to permissions as in the conventional RBAC. TrustBAC is proposed for open and decentralized multi-centric systems where the user population is dynamic and the identity of all users are not known in advance, such as service providers over the Internet.

In social software, however, access regulations to a large part depend on users' privacy preferences and attitude, rather than the receiver's credentials. Trust in social software mostly resembles the way face-to-face trust is shaped in the real world and between real people, which is based on implicit group norms and cultures rather than individuals' credentials. The notion of credential-based audience control as provided by TrustBAC with the addition of the notion of trust levels does not contribute to the support of any of the user-oriented privacy controls. Thus, like RBAC, TrustBAC is not a suitable candidate for privacy management in social software.

5.3.3 RelBAC

RelBAC [1] adds another level of abstraction to the original RBAC model by using the Resource Access Decision facility (RAD [29]) to include the notion of dynamic relationships between arbitrary entities in access decisions. The model is primarily targeted towards healthcare system, although the authors claim that it is general enough to be applied to any domain that requires relationships in access decisions. Relationships are explicitly defined using UML association or dynamic attributes. A combination of roles and relationships is then used to determine whether a permission should be granted or denied.

Like the original RBAC model, RelBAC is suitable in domains where there are clearly defined roles and relationships, for example when roles and relationships are dictated by requirements placed on access to information by the governmental or organizational rules. Considering the notion of relationships in addition to roles in making access decisions provides a more fine-grained right management compared to the original RBAC. The model supports relationship control (and audience control through it), but not in the user-oriented form as neither roles nor relationships are defined or controlled by the end users. Since the model is not geared towards dynamic information, the notion of user-oriented artifact control does not apply. Change control is indirectly supported through the assumption that relationships are short-lived and thus managed through other component of the system (e.g. registration component), rather than the central authority.

5.3.4 TBAC

The TBAC model is another extension of RBAC that introduces domains with task-based contextual information. Access control in this model is granted in steps that are related to the progress of tasks. Each step is associated with a protection state containing a set of permissions. The contents of this set change based on the task. This is similar to the concept of privacy life cycle as identified in our study, although permissions change based on various stages of tasks, not artifacts. TBAC is an active model that allows for dynamic management of permissions as tasks progress to completion. The model also supports validity period and expiration for the access rights.

The notion of artifact control is somehow supported because artifacts are assigned different permissions at different stages, although again this is not managed by the end users. Audience control and relationship control are handled through role assignment as in original RBAC and are not user-oriented. User-oriented change control is not supported.

5.3.5 BCSW

Sikkel [32, 33] presents a general authorization model with an emphasis on conceptual simplicity and ease of use. The model is provided in two forms: The basic form and the extended form. The basic form extends the canonical ACL model with the notion of groups that are used for both assigning roles (user groups) and permissions (access groups). The extended form adds support for delegation, negative rights (exclusion), conditional authorization, and explicit role switching. The model is modular in the sense that the extensions that are not needed in a particular application can be discarded, thus avoiding unnecessary complexity. Context (time, location, etc.) is also supported by the notion of conditional access rights applied to groups.

Use of *user groups* as the basic audience categorization mechanism (based on which roles and other kinds of categories can be modeled) seems to provide the level of flexibility in group definition as required in social software: Since user groups are collections of people without the attributes and operations for various types of roles, they enable group definition based on factors other than organizational roles. Compared to the notion of roles used in RBAC and other models that extend it, the notion of user groups in the BSCW model seems to be a better match for satisfying the audience control requirement in our conceptual model. Artifact control can be supported through the use of access groups, by assigning an artifact to different access groups through various stages of its life cycle. Also, the notion of conditional authorization in the extended model could be used as the basis for adding support for relationship control. Support for change control, however, depends on the actual implementation of the model and usability of the user interface that accompanies it.

As we can see, each of the discussed access control models at best supports some of the requirements of social software (either directly or indirectly). We now move to a short description of a privacy control system based on the user-oriented privacy controls that we are incorporating into an open-source social networking and information management system and our plans to test this against other approaches.

6. Tagging people: a new model for relationship control

One of the most significant challenges in developing a system for audience and relationship control, and thus for supporting user-oriented privacy control and information sharing, is the subtle and nuanced way in which our patterns of trust change over time and the ways in which this interacts with our transactional approach to information sharing and exchange. In this respect, our study confirmed existing theories of knowledge sharing that compare the exchange of information between people with the exchange of money in economic systems [8]. While the analogy is not perfect, this observation highlights the contextual nature of the choice to share knowledge and the degree to which these choices depend on an assessment of the personal benefit to be gained from sharing weighed against the risk of sharing that specific information with that particular audience.

At minimum, a solution to this problem must involve an ability to associate collections of information-bearing *artifacts* with groups of people (the *audience*) defined largely in ego-centric terms (the *relationships*). The insight that leads to a potential solution is that the organization of relationships can be treated in the same way that we organize information itself, and that the model of personal information organization called *folksonomy* or *tagging* has exactly the characteristics necessary to facilitate relationship management for information sharing.

Simply put, 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 the social bookmarking system del.icio.us [15] and the photo exchange system Flickr [17], it has been adapted to 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 selfish purposes (what Vander Wal has termed *broad* folksonomy).

In relating information management to relationship management, we highlight a number of features of the tagging model:

1. Many tags (and thus categories) can be associated with each artifact;
2. The choice and control of tags is entirely in the control of the individual user;
3. The act of tagging is simple, intuitive and well-adapted to granular information collections (e.g. web bookmarks); and
4. The collections created by coincidental tagging (i.e. all artifacts tagged with the same words or the same set of words) form natural categories.

For these reasons, we propose to model relationships for information sharing by tagging people, represented by their profile pages in a social information sharing network.

Opntag [19] is a web-based, open source system for note taking and bookmarking we have developed to experiment with personal information management and exchange in sensitive environments (e.g. within corporations). The fundamental unit of information storage in opntag is the "memo", a tagged textual annotation that may optionally refer to any URL-addressable object. Fundamental to its implementation is an ability to restrict the visibility of these memos to one or more groups (including the "private" group consisting only of oneself). To this point, we have used a fairly traditional model of user groups, based on the hierarchical BSCW model (e.g. a particular memo and its associated tags may only be visible to the "opntag developers" group).

One of the experimental focuses of opntag has been to exploit the opportunities presented by tagging or creating memos that refer to other objects in the system. For example, a memo that refers to another existing memo is considered to be a "reply" to that memo and becomes automatically threaded into the conversation that the first memo is part of. Memos on collections become associated with those collections and we are investigating the consequences of other tags applied to collections (e.g. in one experimental extension such tags are viewed as "implication markers" that signal semantic implication and automatic tagging).

Within this environment then we have started to experiment with the tagging of your own (identity tagging) and others profile pages as a way of "categorizing" friends and collaborators. This might be useful simply as a way of signaling our assessment of others (e.g. I might tag a seller on eBay as "unreliable") or as a way of signaling a relationship (e.g. I will tag my graduate students as "student" and "grad student"). When viewing that profile page then, I may be able to see the person's own tags for himself (e.g. a self-assessment of identity), my tags (signals of our relationship) and other's tags (third-party opinions). This is all, of course controlled by opntag's visibility management facility, so I will only see those tags that the taggers have allowed me to see, and thus it is reasonably safe to "opinion tag" others, but this is highly volatile and private information, so likely to be lightly shared.

As we have highlighted above though, associating the visibility of these tags with invitation-only or open membership groups (e.g. online communities) is probably not sufficient in most cases, since we often make such sharing decisions based on relationships more personal than shared membership in a community. For example, I may want my "friends" (i.e. those others I have tagged with "friend") 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. The obvious solution to this then is to treat each of these "tagged categories" of other users as a "relationship group" which is usable as a visibility category. Thus, the act of "tagging a person" (via their profile page) is equivalent to asserting their membership in a group whose membership is entirely under my control.

Currently, this implementation is incomplete and scientifically untested, but we can assert that it fulfills all of the control criteria outlined above. Sharing control within opntag is already done on the basis of *artifact*-specific privacy control, since each memo in the system and its associated tags is visible only to its specific *audience*. The visibility management model in opntag is also clearly a user-driven *audience* control approach, with the audience for an item defined as the set of users the object is visible to[1]. The people tagging establishes the egocentric *relationship* control our study suggested and tying that to the visibility model allows one to exploit these relationships for audience management. The one aspect of the problem not directly addressed by this solution is *change control*, although the visibility of a memo or tag a user has created is always modifiable. What is needed is a way to match the changes in audience to identifiable stages in a privacy life cycle model, still to be developed.

The most salient comparison with this approach is the one exemplified by Facebook. A contrast to the bottom-up, user-defined vocabularies is the traditional application- or community-defined taxonomy. In Facebook, relationships are classified with a set of standard assertions represented by the dialogue in Figure 3: Facebook friend categories. We suggest 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 my friends or networks (Facebook's groups).

Figure 3: Facebook friend categories

We are on track to complete the "tagging people" implementation in opntag and release it to a wider audience than the lab within a few weeks (on the hosted opntag.net site). Once we do so, we will conduct a survey and controlled tests comparing opntag's approach to relationship-based information sharing with that implemented in Facebook.

7. DISCUSSION

The representations of the data that emerged from our grounded theory analysis provide a set of propositions for understanding privacy requirements in social software. Our most important finding was that users have a fundamental assumption that when they put something in the tool, they should have control over it. Our data confirm the intuition that users can be reluctant to share personal information when the consequences of doing so are unclear, or when they are unable to control the transactional aspects of knowledge sharing activities. A counterintuitive consequence of this may be that some users are more ready to share personal information in a space that affords virtually no privacy control (e.g. blogs or Myspace pages) than one which offers them an inadequate 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 in the social system they used. When they found it inadequate, they

[1] In opntag, visibility implies readability, so there is no "I can see that it exists but cant read it" issue.

often chose to not place information into the system because of the inflexibility of the tools or the lack of ability to model consequences of their actions.

This points to the importance of perceived affordances of privacy management mechanism for social software (as for any other user-oriented tool). As defined in the HCI field, perceived affordance is "action possibilities which are readily perceivable by an actor [11, 25]". Simply put, the concept emphasizes that suggested interactions with a tool must be in accordance with the ability of the actors to perceive those interactions. Perceived affordance has been identified as a major contributor to enhancing usability of a design [21, 24]. Because our privacy management mechanism is based on users' mental model of information privacy, we believe it provides better perceived affordance, thus improving the overall usability.

Even though the required privacy controls that were identified by our results were mere confirmation of the factors known by existing access control models, the fact that in social software these controls need to be in the hands of the users calls for new approaches in design of privacy management mechanisms in this context. We believe the insufficiencies of current mechanisms are the results of a significant gap between the perceived affordances of the underlying model and user requirements. We expect our findings to contribute to reducing that gap.

8. CONCLUSIONS

Although the use of social software for a variety of purposes has moved from leading edge to mainstream over the past few years, it is still in the early-adopter phase. Among issues that need further investigation are the issues of privacy and access management in these environments. We believe that the ability to understand and control information sharing in a natural, fluid manner is essential to the acceptance of these tools by a broad set of users, and yet, none of the existing access control models in the literature address the specific privacy needs of social software.

This research summarizes the results of our investigation into privacy issues as they pertain to the specific context of social software. We used the results of a grounded theory study of information sharing behavior to propose guidelines for the design of privacy control mechanisms. We discussed current access control models and explained why they are not sufficient for specific needs of social software, and then presented our proposed solution for a privacy management mechanism for social software that we believe can address those insufficiencies.

An 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' experiences and opinions rather than deduced from the literature. As such, they give valuable insights into the processes entailed in information sharing in social software, and they provide a framework to direct further research.

It is yet to be determined whether our proposed solution has been successful in improving users' experience with the privacy management mechanism. Clarifying where our solution stands compared to existing solutions (through performing usability studies) is part of our continuing research agenda.

9. REFERENCES

[1] Barkley, J., Beznosov, K., and Uppal, J. 1999. Supporting relationships in access control using role based access control. In *Proceedings of the Fourth ACM Workshop on Role-Based Access Control* (Fairfax, Virginia, United States, October 28 - 29, 1999).

[2] Boyd, D. 2006. Identity Production in a Networked Culture: Why Youth Heart MySpace., In *American Association for the Advancement of Science*, St. Louis, MO. February 19.

[3] Burrow, A. L. (2004). Negotiating access within Wiki: a system to construct and maintain a taxonomy of access rules, In *Proceedings of the fifteenth ACM conference on Hypertext and hypermedia*, Santa Cruz, CA, USA, pp 77 - 86.

[4] Cao, X. and Iverson, L. 2006. Intentional access management: making access control usable for end-users. In *Proceedings of the Second Symposium on Usable Privacy and Security* (Pittsburgh, Pennsylvania, July 12 - 14, 2006)

[5] Chakraborty, S. and Ray, I. 2006. TrustBAC: integrating trust relationships into the RBAC model for access control in open systems. In *Proceedings of the Eleventh ACM Symposium on Access Control Models and Technologies* (Lake Tahoe, California, USA, June 07 - 09, 2006).

[6] Coulouris, G. and Dollimore, J. (1994), A security model for cooperative work, *Technical Report 674, Dept. of Computer Science*, Queen Mary and Westfield College, University of London, 1994.

[7] Erickson, T. From PIM to GIM: Personal Information Management in Group Contexts, *in Communications of the ACM*, January 2006.

[8] Fuller, S. (2002). *Knowledge management foundations.* Boston: Butterworth-Heinemann.

[9] Glaser, B., Strauss, A., *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Chicago, 1967

[10] Gross, R., Acquisti, A., and Heinz, H.J. III, Information revelation and privacy in online social networks, In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, p. 71– 80

[11] http://en.wikipedia.org/wiki/Affordance

[12] http://en.wikipedia.org/wiki/Web_2.0

[13] http://en.wikipedia.org/wiki/Principle_of_least_astonishment

[14] http://bluedot.us/

[15] http://del.icio.us

[16] www.**facebook**.com/

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

[18] http://ma.gnolia.com/

[19] http://opntag.net

[20] http://www.orkut.com

[21] McGrenere, Joanna, Ho, Wayne (2000): Affordances: Clarifying and Evolving a Concept. In *Proceedings of Graphics Interface 2000*, May 15-17, 2000, Montreal, Quebec, Canada. p.179-186.

[22] Millen, D. R., Feinberg, J., and Kerr, B. 2006. Dogear: Social bookmarking in the enterprise. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada, April 22 - 27, 2006).

[23] Morse, J. M., Richards, L., *README FIRST for a User's Guide to Qualitative Methods*, Sage publications, 2002

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

[25] Norman, Donald A. (1999): *Affordances, Conventions, and Design*. In Interactions, 6 (3) p. 38-41

[26] Olson, J.S., Grudin, J., and Horvitz, E., A study of preferences for sharing and privacy, In *Proceedings of CHI 2005,* Portland, Oregon

[27] Patil, S, Lai, J. Who gets to know what, when: Configuring privacy permissions in an awareness application, In *Proceedings of CHI 2005,* Portland, Oregon

[28] Razavi, M. N. and Iverson, L. 2006. A grounded theory of information sharing behavior in a personal learning space. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work* (Banff, Alberta, Canada, November 04 - 08, 2006)

[29] Resource Access Decision (RAD), *Object Management Group Healthcare Domain Task Force*, Revised submission, OMG TC Document corbamed/99-04-04, April 26, 1999.

[30] Sandhu, Ravi S., Coyne, Edward J., Feinstein, Hal L., & Youman, Charles E. (1996), Role-Based Access Control Models. Computer, Volume 29, Number 2, February 1996, 38-47.

[31] Shen, H. and Dewan, P. 1992. Access control for collaborative environments. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work* (Toronto, Ontario, Canada, November 01 - 04, 1992).

[32] Sikkel, K. (1997a), A Group-Based Authorization Model for Cooperative Systems. European *Conference on Computer Supported Cooperative Work* (*ECSCW'97*), Lancaster, UK, 345–360.

[33] Sikkel, K. (1997b), A Group-Based Authorization Model for Computer-Supported Cooperative Work. *Arbeitspapiere der GMD 1055*, GMD, Sankt Augustin, Germany.

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

[35] Whalen, T., Gates, C., Private Lives: User attitudes towards personal information on the web, poster, *in SOUPS 2000*

[36] 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).

[37] Whitten, A., and Tygar, J.D., Safe staging for computer security. In *HCI and Security Systems Workshop*, CHI 2003, Ft. Lauderdale, Florida, April 2003.