A Grounded Theory of Information Sharing Behavior in a Personal Learning Space

Maryam N. 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

This paper presents a grounded theory of information sharing behavior of the users of a personal learning space. A personal learning space is an environment consisted of weblog, ePortfolio, and social networking functionality. It is primarily used within education as a tool to enhance learning, but is also used as a knowledge management tool and to develop communities of practice. Our results identify privacy as a main concern for users of a personal learning space and illustrate challenges users face in ensuring privacy of their information and strategies they employ to achieve the desired level of privacy. We then identify factors that affect users' decisions regarding disclosure of their personal artifacts to various people and groups in a personal learning space. The three main themes as emerged in our study include current stage in 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 a personal learning space. The findings offer some ideas about how to create privacy management mechanisms for personal learning spaces that are based on users' mental model of information privacy. Practical implications of the results are also discussed.

Categories and Subject Descriptors

K.3.1 [**Computers and Education**]: Computer Uses in Education – *Collaborative learning*; H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces – *Theory and models.*

General Terms

Security, Human Factors, Theory.

Keywords

Grounded theory, ePortfolio, weblog, information sharing, information privacy, user modeling.

1. INTRODUCTION

Convenience and ease of access have made the Web a favorable medium for storing personal artifacts. The more personal artifacts are stored in digital form, the easier they are to share. Although often times creating personal artifacts in open online environments happens with the goal of sharing with others, not everything is to be shared with everyone. While use of personal publishing tools such as weblogs, Wikis, and ePortfolios has moved within the reach of non-technical mainstream, managing privacy of the published information still requires expertise. 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 not yet been systematically studied. Privacy concerns for personal artifacts are different from privacy concerns in e-commerce or workplace: Personal artifacts contained in such a system can cover a wide range, from ones' picture, contact information, and interests to his/her social network, scholarly work, and opinions expressed in personal weblog. The shared artifact and the group in which it is shared could both be dynamic, and preferences regarding sharing the artifact within the group have to be flexible enough to accommodate these 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.

One of the areas where personal artifacts are largely shared with a non-uniform group of audiences is the emerging field of personal learning spaces [4, 5, 10, 14]. A Personal learning space is an environment consisted of weblog, ePortfolio, and social networking functionality. It is a new breed of software which allows effective connection of people to each other as well as resources and facilitates the creation of communities. Personal learning spaces are extensively used within education as a tool that is truly learner driven and allows learners to create their own community learning space, and in organizations and commercial businesses to develop communities of practice. They have also been considered as a knowledge management tool which encourage the connection of expertise, resources and people within organizations. At the core of a personal learning space is an ePortfolio, defined as "a Web-based information management system that uses electronic media and services to enable the learners to build and maintain a digital repository of artifacts for

demonstration of competence and reflection on their learning [4]." Combining the artifacts contained in an ePortfolio along with discussion/reflection functionality provided by weblogs provides a rich view of learners' experiences and skills. The ability to manage access to this rich view is an essential part of achieving the ultimate purpose of wide adaptation and usability of the personal learning spaces.

Users are encouraged to include personal life experiences, awards, non-academic activities, and other character/learning revealing artifacts in their personal learning spaces. Also, use of a personal learning space is considered as a continuum, where it will be used during different stages of learning for different purposes. From this point of view, a personal learning space is rather a minable, searchable life-log that contains a compilation of collaboration and in-process work habits of the owner that taken out of context, could misrepresent intended meaning. In such an environment the social implications of having a persistent, tractable online identity must be considered.

Even though existing platforms provide some tools for managing selective information disclosure, the act of specifying who should have access to what under what condition could become a burden to the user. The choice between private and public may become onerous, given the volume of data and the diversity of data sources. There is a clear need for a more flexible, comprehensive scale that takes into account the fact that there are commonalities in the way people perceive privacy of their information in this particular context. These commonalities could be taken advantage of in creating privacy management mechanisms that support users in the face of changing needs, without creating so much overhead that they either share everything or share nothing at all. This research has been motivated by this need: We believe current issues with privacy management in personal learning spaces rise from a model of information sharing that is far too simple. To be effective, mechanisms must be built based on models that reflect user experience. This paper presents our first step towards such a goal: understanding users' perception of privacy of their artifacts in a personal learning space, in order to propose a conceptual model of personal information sharing behavior in such an environment.

2. RELATED WORK

Although the momentum is building for the widespread use of personal learning spaces in education, the discipline is still in its infancy. Many issues are yet to be explored, notably security and privacy of information among others. However, researchers have studied people's attitudes towards disclosing personal data in several other contexts, including multimedia, work place, social applications, and location-aware mobile services.

In a series of related works [1, 2, 3], Adams and Sasse summarize the details and results of four qualitative empirical studies to model users' perception of privacy in multimedia environments. 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.

Olson et. al. [16] 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 [17] 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 [19] 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. 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. The issues and concepts raised in this article are so far the closest to the issues explored in this research, although the article focuses more on specifying possible research directions rather than providing solutions.

3. RESEARCH QUESTIONS

Our program of research asks several questions with regard to personal information sharing preferences. We were primarily interested in exploring the following aspects:

- What factors affect privacy of information form a user's perspective in a personal learning space?

- Is privacy management in a personal learning space considered important, and why?

- What are users' challenges in managing selective disclosure?

- What are users' strategies in achieving privacy?

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

- Is it possible to drive a set of default privacy settings for different categories of information in this environment, which users can easily modify later in context?

By trying to find answers to these questions, we were aiming for identifying fundamental concerns with privacy form 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 incorporate those abstractions in a conceptual model of information sharing behavior in a personal learning space.

4. RESEARCH METHOD

The research method that was employed in this study was grounded theory; 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 analysis and data collection. The grounded theory method has been suggested in the literature as the appropriate method for situations where the researcher is trying to reveal user experience or construct a theoretical framework based on reality [8]. The purpose of grounded theory method is building theory, not testing theory. 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 results of this research method are propositions and hypotheses, not findings. Theory concepts are suggested, not proven.

5. DATA COLLECTION

Methodological congruence [15] is an important issue when choosing a qualitative research methodology. It implies that the research method (i.e. grounded theory, ethnography), research strategy (i.e. interviewing, videotaping), and research technique (i.e. 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 have selected semi-structured in-depth interviews for our data gathering strategy, suggested as one of the best fits with the grounded theory methodology [9, 15]. Unlike structured interviews, semi-structured interviews have a flexible and dynamic style of questioning directed toward understanding the significance of experiences from the informants' perspective. This strategy is primarily suitable for situations where the researcher knows enough about the domain to develop questions, but not enough to anticipate answers [15]. 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. All of the interviews in this study were tape-recorded with the informants' permissions, and later transcribed to provide accurate records for analysis. Standard procedures were followed to maintain the confidentiality of the interview data and the anonymity of the informants.

The interviews were structured around a list of topics based on concepts presented in the existing literature, including sharing preferences with regard to the type of information, the person or group with whom the information is being shared, and the purpose behind sharing. The core topics were as follows:

- What kinds of information are shared, and why?

- What categories of objects are perceived to need protection, and what factors shape this perception?

- Are current privacy management mechanisms sufficient, and (if not) what are the problems?

- Has there ever been a change of preference regarding sharing a certain object, and what factors affect this decision?

- Are there differences in the way information is shared with groups and communities of different nature?

6. LOCATING THE STUDY

Since a grounded theory method looks for emergence of theory from the data, qualitative 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 [15]. Informants chosen for interview 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 experience. For these reasons, we have followed a procedure called *purposeful sampling* as our initial sampling strategy in this study. In this procedure, only participants are selected who know the information required and are willing to reflect on the phenomenon of interest. Our initial set of participants included 9 high school students enrolled in a special program for gifted kids, who were using a personal learning environment for over a year. The environment is called Elgg [8, 18]. Elgg provides the tools necessary to set up personal learning space: It includes tools for online journals (weblogs), an e-portfolio in which pupils can store and showcase their work in a range of formats, and software to support resource tagging and social networking. Students were required to fill in their personal profile, write reflections in their weblogs on the topics covered in the classroom, and join and participate in a special community created for their group. For each of these artifacts (weblog posts, profile items, and personal reflections posted to the community blog), they had the option of regulating access (i.e. make it visible to only oneself, the instructor, a specific community, or everyone in Elgg). Since active use of the environment was part of the their curriculum, these students had in fact 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.

One key feature that distinguishes Elgg from similar tools is its user-centered approach, as opposed to the more traditional presentation-based view that is employed by most other tools [18]. By facilitating the formation of learning communities, Elgg enables creating an environment where knowledge sharing, conversation, and reflection can take place and students are contributors as well as recipients of knowledge. Another key feature of Elgg is that it provides reasonable support for privacy control at a fairly granular level that other tools simply don't have. These features were the main reasons we chose Elgg for the purpose of our study; but even though the study is situated in the context of Elgg, constant effort has been made not to limit the discussion to Elgg. 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 the interviews on the more general area of information sharing behavior in the context of personal learning spaces.

The analysis of the data gathered from our initial interviews with Elgg participants resulted in identifying the *basic social processes* (BSPs), which are the core concepts around which the grounded theory is built. 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. For this process, we interviewed 3 more students who had experience with other similar platforms in addition to Elgg, including other ePortfolios, forums, and weblogs. We also redirected the interview questions in a way to reflect our new goal of verifying the emerging theoretical themes and their relationships. The experience of these 3 participants particularly helped in identifying places where the current privacy mechanism was considered insufficient and users felt the need to switch to other platforms in order to achieve their goal.

The participants' ages ranged from 15 to 17, and the gender balance was evenly split; there were 6 females and 6 males. All participants were quite confident with the Web. All had extensive prior experience with various sorts of open online environments in addition to Elgg.

Our data collection ceased when the indicators pointed to *theoretical saturation*, the point at which we could identify interchangeable examples showing the same phenomenon in different instances, and additional data was no longer adding to the concepts and relationships being developed.

7. DATA ANALYSIS
As with much qualitative research, data collection and data analysis occurred simultaneously in this study. A theory was derived from the data using a constant comparative method of analysis with three stages: open coding, theoretical coding, and selective coding. The next following sub-sections explain each stage in detail.

7.1 Open Coding
This stage of analysis involved going through the interview texts and applying code words to sections that identified pertinent concepts, following Glaser and Strauss' [9] 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.

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. 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.

7.2 Theoretical Coding
The second stage of analysis involved taking the concepts that emerged during open coding and reassembling them with propositions about their relationships. The relationships, like the concepts, emerged from the data through a process of constant comparison. These emerging propositions then formed a

theoretical framework, which served as a guide to further data collection and analysis.

Using the constant comparison method, some codes were subsumed under broader or more abstract categories. At the end of this stage, the core categories and their relationships were shaped.

7.3 Selective Coding
As the theory developed, it evolved around the core categories and their relationship that reflected the main theme of the study, selective information disclosure. The identification of the core categories led to selective coding, the process of delimiting coding to only those concepts and relationships that relate to the core categories, resulting in a more focused theory with a smaller set of higher-level concepts.

8. THE GROUNDED THEORY
Two separate concepts emerged from our analysis: centrality of privacy as a concern, and factors affecting users' privacy preferences. Each concept evolved around a collection of core categories. In the next sections each concept and the related categories are explained in detail.

8.1 Centrality of Privacy as a Concern
The concept map in figure 1 draws together the categories of importance that gave rise to the concept of centrality of privacy as a concern. As the diagram indicates, while many participants are considering long-term use of their personal learning space because of its many potential benefits, certain characteristics of the environment combined with the long-term use give rise to the importance of the issue of information privacy. Some of the participants were already employing strategies to forego the shortcomings of the tool in this matter and overcome privacy challenges. A more descriptive account of each element in the diagram plus table summaries of how each element emerged from participants' descriptions is provided below.
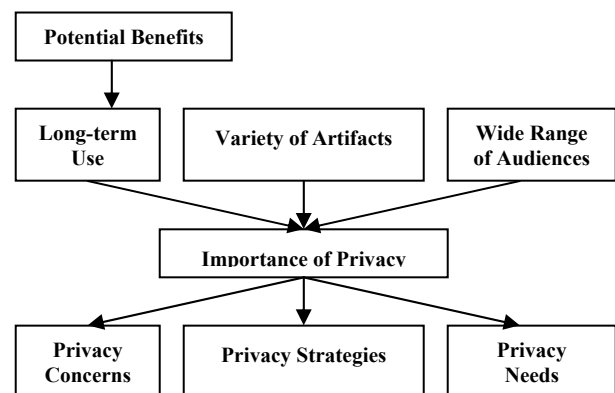

Figure 1. Centrality of privacy as a concern

8.1.1 Benefits of a Personal Learning Space
Our results show that most of the informants envision using their personal learning space as a continuous process. Many stated that they would be using their space after the end of their program

(where they are *required* to use it). The reason for this was the many benefits of using such an environment in various dimensions, including 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; The opportunity to keep track and reflect on ones' improvement over the time; The advantage of getting anonymous, unbiased feedback on their creative artifacts form a community of people who share the same interest; and ease of presentation to potential future employers/university administrators where they would apply for education. Table 1 summarizes some of the statements that present the benefits of long-term use of a personal learning space as viewed by the participants. To ensure anonymity of the participants, they have been referred to by numbers: in all tables, the first column refers to the participant number, while the second column refers to the comment made by the participant regarding the specified issue.

Table 1. Benefits of a personal learning space

	Comment
2	By keeping my ePortfolio up to date, teachers and other evaluators can see how I am developing.
1	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.
6	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. It's better on the Internet.
12	It's a good way to keep track. Even if you don't change it a lot, at least it's there if you ever need it.
11	[It's good to share personal artifacts because] you get more feedback. Because if you show it only to your teacher, your teacher can only say so much. But your friends might have more ideas; they might say you could fix this up or fix that up. The teacher is probably busier, too. Your friends might have more time. The public might give you ideas, too, if you let them see it.

As the table shows, even though Elgg is primarily used by the students for the purpose of enhanced learning (#1, 2, & 11), they are using it for more than that: It is also used as a knowledge management tool (#6 & 12) and for showcasing their work (#1). This shows that when users are provided with such a rich environment that provides support for both work related and social activities, they used it for a lot more than initially conceived. Support for privacy may not have been initially

considered in the tool for all these areas. That's when privacy concerns emerge.

8.1.2 Privacy Concerns

The artifacts contained in a personal learning space have a wide variety, ranging from personal profile and reflections to educational material and creative stuff. Each category may be targeted to a different group of audiences that are not necessarily static. These specific characteristics of the environment plus the tendency for long-term use, gives rise to the concern for selective sharing. Some of the concerns expressed by the participants were affecting people's attitude by disclosing certain stuff, fear that some stuff might be interpreted out of context, fear of losing control especially over their creative material, and sometimes just the simple feeling of "awkwardness" at the thought of certain stuff being exposed to public. Table 2 summarizes some of the statements that reflect these concerns.

Table 2. Privacy concerns

	Comment
8	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.
2	Most of them [blog entries] aren't public. Because when I first started everything was public, and I found that some strangers leave random 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.
12	Even your educational information sometimes needs to be kept private. For example, I think 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 your educational and social information from certain people who really don't know a lot about you, you are treated more like an equal.

8.1.3 Strategies for Achieving Privacy

Our results show that many of the participants already employ certain strategies to achieve their desired level of privacy when the tool does not provide it. 8 out of 12 participants reported using other platforms with better privacy management mechanisms for their more private content. In more extreme cases, some had decided to forego deploying certain stuff in an online environment because of lack of acceptable privacy levels, although when asked if a better privacy management scheme would change their attitude, in 5 out of 6 cases the answer was yes. Other strategies included writing private content in some sort of a "code language" so that it is meaningless to anyone other than the writer himself, and having the contents somewhere (i.e. a web page, weblog, or an ePortfolio), but not providing a link to it from places where their real identity is known. Table 3 summarizes statements regarding strategies for achieving privacy.

Table 3. Strategies in achieving privacy

	Comment
9	[my private ePortfolio] 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.
7	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.
12	[for private stuff] I wouldn't write them down anywhere because the easiest secret to keep is the one that is never told.
10	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. So there is some privacy, not everyone sees everything, which would be nice. Because my ePortfolio is personal so I like to keep some privacy.

Table 4. Privacy needs and challenges

	Comment
8	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.
9	The problem I have with that [current privacy mechanism] 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.
12	When I share [some form of creative stuff] it is usually because I want some feedback on it. But there is a problem with making it public because then you need to keep it really censored: sometimes you don't want to give too much information about yourself like where you live or something... but then it's not possible to make some parts private and some parts public.
5	... I just took that off. I would have made it private if I could, but it's hard to make just one word private.

8.1.4 Privacy Needs

Current privacy management mechanism in Elgg is restricting access through creation of access groups. After joining the system, a user can use tags to search through system *resources* (weblog posts, repository items, user profiles, and existing communities) to find others with similar interests and add them to his network of friends. However, access to the person's information will be limited to only what he has made available to public. By creating an access group it will be added to one's access levels for any resource created. The user can then categorize his network into different access groups with different access privileges based on his sharing preferences.

Although all the participants reported using current privacy management mechanism to maintain their preferred level of privacy, they were not always getting the results they were looking for. 9 out of 12 participants mentioned that a better privacy management tool would improve their experience with the system. Among the needs expressed were the need for hiding the fact that something is not being shared, need for the ability to control privacy at a more granular level, and the need for a more comprehensive grouping rather than public/friends/private. Even though the tool provides the functionality to manage access rights on a case-by-case basis, 6 out of 12 participants mentioned that they find it too labor-intensive to apply on a regular basis. Table 4 summarizes some of the participants' comments regarding the issue of privacy needs and challenges.
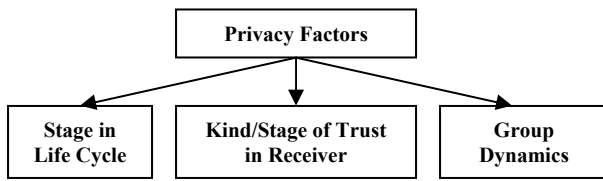
8.2 Factors Affecting Privacy Preferences

With recognizing privacy as a main concern for users of a personal learning space, we approached the problem of unraveling how they perceive privacy of the resources they dispose in such an environment and what factors affect their perception. We arranged our questions along three main lines: the resources users feel they need to protect, people/groups with whom the resource is being shared and to whom access privileges are granted, and the perceived usage of the shared resource. Our starting point in designing the questions was the model of privacy factors and issues for multimedia environments provided by Adams and Sasses [3]. In their model, Adams and Sasse break the privacy factors into three main categories: information sensitivity, information receiver, and information usage. Although their model is targeted to the multimedia environments, we found it general enough to be used for deriving our initial set of interview questions, keeping in mind that each category will have a different definition in our specific domain.

The concept map in figure 2 draws together the categories of importance that emerged from our analysis of factors affecting privacy of information in a personal learning space. As the diagram indicates, we identified one substantial factor in each dimension: our results indicated that sensitivity of a shared resource is largely perceived by the current stage in the life cycle of resource. 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 descriptive account of each element in the diagram plus table summaries of how each element emerged from participants' descriptions is provided below.

Figure 2. Factors affecting privacy preferences

8.2.1 Information Sensitivity: The Effect of Life Cycle

Our results show that users' judgment of privacy of a resource is rarely based on a binary scale of public/private. Rather than a simplistic classification of artifacts into categories with different degrees of privacy, analyzing participants' reports show that most of them apply a more sophisticated and flexible scale of information privacy that reflects the transition from private, to semi-private/restricted share, to public, with regard to the state of the information, the receiver, and the context of sharing. Table 5 summarizes some of the comments that refer to the change in privacy preferences.

Table 5. The effect of life cycle

	Comment
8	Sometimes, I would put it 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.
3	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.
1	I use it [transition from private to public] for 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.
12	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.
11	For example, for [a particular course] there are some questions due today on [some topic], so we have to put them on our ePortfolio. So I had them on my private weblog, but I will make them public tonight so that I can tell [the instructor] the site and she can look at it.

As the comments show, one particular criterion that continually appeared in the accounts of users was changing sharing preferences based on content life cycle. Content life cycle refers to the path that a particular piece of content takes from its source or creation to final publication on and then onto a storage archive for future reference and possible revision. Our results show the existence of a similar pattern for information privacy, which we call as "privacy life cycle", the privacy requirements of an artifact during various stages of its life cycle. We found out that for many users, their decision about information disclosure is affected by the current stage of the privacy life cycle. We identified three main stages of privacy life cycle:

- *Creation/Authoring Phase*: An artifact is often considered private at the time of creation. Descriptions, goals, personal reflections, and personal assessments may be included with the artifact in this phase.

- *Review/Work in Progress Phase*: At this phase, the artifact is considered semi-private. Restricted access may be granted to peers, support groups, and collaborators to read and possibly provide feedback. Comments and annotations are possible, but with restricted read and write access to a closed circle.

- *Distribution/Publishing Phase*: at this phase, the artifact is shared beyond the group who created it, although this does not necessarily mean making it public. It could be useful to revoke write access from all parties at this point to ensure originality, especially for creative stuff.

As users' comments show, this factor mostly affects artifacts of a dynamic nature (i.e. assignments, group projects, creative works). Among various components of a personal learning space, weblog posts and file repository are the two main places where artifacts of this sort are mostly shared. This suggests that these components might benefit from a set of access patterns that can be assigned to resources at the time of creation that will allow privacy life cycles to automatically follow the artifacts' production life cycles.

8.2.2 Information Receiver: The Effect of Trust

Our results indicate that users' assessment of the persons or groups who will be the receivers of information plays a strong role in deciding about information sharing. The most influential factor seems to be trust. We found out that users' sense of trust (either in another person or in a group) moves along a progressive path (i.e. from less trust to more) rather than following a binary trusted/not trusted pattern. This is inline with findings in existing trust literature in HCI and e-commerce [12, 13]. 8 out of 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, causing them to share more freely. This is very much in line with the way face-to-face trust is shaped in the real world and between real people, pointing to the fact that the notion of groups in social software must reflect the way human groups behave and work, rather than mathematical models. Table 6 summarizes some of the comments that refer to the developmental nature of trust.

Table 6. The effect of trust

	Comment
11	Right now I am on a forum and I remember in the beginning I was really careful about saying [my age] and [where I live]; and that was about all I said, because that's really general information and nobody can do anything with it. But I really avoided things like exposing 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.
6	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.
10	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.

In a comprehensive study of online trust, Corritore et. al [6] specify four different dimensions for trust between humans and computers: generality, kinds, degrees, and stages. In our study, though, what appeared in accounts of users as relevant factors that affect their information sharing attitude in groups and communities, was only 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 (#6).

8.2.3 Information Usage: The Effect of Group Dynamics

Our results show that users' willingness to share something they have vested interest in also depends on their perception of how it will be used. The two main concepts that emerged in this category as the main issues in holding back from sharing were the risk of loss of control and influence, and the risk of not getting credit for their works. One drawback of current social software systems in general is that although they provide facilities for creating and participating in groups and communities, there is no indication of what the information sharing manners are in a particular group or community. So while in real life, implicit group cultures (with regard to membership, visibility, etc.) 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. Table 7 summarizes some of the users' comments regarding the effect of group dynamics on their information sharing behavior.

Table 7. The effect of group dynamics

	Comment
4	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.
3	[What I share in a particular community] would really depend on the subject. Like, in [a particular community they have in their program] I know the students [who are members of the community], 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.
5	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. But usually the discussion-based ones are pretty open, and that's OK.
12	[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.
11	[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.

We believe a clear notion of group characteristics (i.e. size, public/private visibility, open/controlled membership) can highlight the potential trade-offs between risks and benefits of information sharing in a particular group or community. It could give users' a high level overview of the effects of their sharing decisions so that they can see how potentially a particular resource could be used in a particular community and tune their sharing decisions accordingly.

9. DISCUSSION

The theory described in this paper provides a relatively comprehensive answer to the research questions. Overall, 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 of a personal learning space: 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 information 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 group dynamics on information sharing attitude) may be considered new insight into information sharing behavior in learning communities. Even for the concepts that have been studied before, their role in the particular context of learning communities 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' experiences and opinions rather than deduced from the literature. As such, they give valuable insights into the processes entailed in information sharing in learning communities, and they provide a framework to direct further research.

Even though existing literature had pointed to some of the results found in this study, the design of privacy management tools for social software has not yet been informed by these results. As such, our next step will be to build a prototype based on the conceptual framework suggested in this study and to perform usability studies to verify whether a privacy management tool that is built on users' mental model can in fact improve users' experience.

Other possible directions for future work are complementing this study with results from the study of a possibly wider and more diverse group. Since personal learning spaces are primarily targeted to educational communities, such as high schools and higher education, a possible direction would be to approach other potential knowledge practitioners such as instructors and administrators 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 personal learning spaces are also a fundamental part of any knowledge-oriented work practice, we plan to extend this study into work environments where the cooperative-competitive balances are likely to be different.

10. IMPLICATIONS FOR DESIGN

Our results also provide some guidelines for the design of privacy support mechanism for educational software: current mechanisms might seem good enough only because users make do with what they have. They usually come up with creative ways to "gel" what the tool offers with what they want to achieve at a "task" level. The following is a summary of some of the direct implications of our findings on the design of privacy support mechanism for personal learning spaces.

- Our results show that as in other contexts, there are commonalities in the way users arrange their information

with regard to sharing in a personal learning space: there is a category of artifacts, such as personal reflections and character revealing content, which most users prefer to keep private or semi-private; there is another category of artifacts, such as in-progress projects and samples of creative work, to which most users apply a privacy life cycle, meaning they allow different recipients (i.e. teachers, peers) to perform different actions (i.e. read, comment) on the artifact based on the current stage in the artifacts' life cycle; and finally there is the category of general information, which users consistently feel safe to leave public. In essence, there are 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 with a limited, trusted audience. These seem to be well served by the current "set on creation" model. The other category however, seems to match privacy and control to a document's state. More fluid techniques for specifying these states and matching access-control settings to status seem necessary to manage this privacy life cycle.

- We also discovered some problems with the current "group" and "community" concept in Elgg (and other tools): There seemed to be a lack of clarity on the definition of a group vs. a community, especially as to how they differ and what purpose each one should be used for. Also, we found that many perceive the community concept to be inflexible. Some of the common problems mentioned were no way to invite people to join one's community, no way to prevent unwanted people from joining, and no way to ban members or disable their membership. We believe personal learning spaces 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 fixed/controlled membership.

11. CONCLUSION

This research is an attempt towards modeling users' perception of privacy of information in a community of knowledge. It is a first step towards the goal of creating a conceptual framework for building privacy management tools that address the particular needs of personal learning spaces. Although the use of social software in education has moved from leading edge to mainstream over the past few years, it is still in the early-adopter phase. Instructors and students alike need to be shown a way to incorporate these technologies into their daily routines if we want the technology to be taken up more significantly. We believe privacy management is a key point in making this goal materialize: by providing support for selective information disclosure that maps to users' mental models of information privacy, the learning community will become a safe place for sharing ideas and reflections, enabling effective collaboration, and enhancing the learning process. By ensuring information privacy, the learning community is envisioned as a habitat, providing the users with control over sharing of information of different degrees of sensitivity with a small or large group, and the opportunity for easy transition of information sharing process into productive, collaborative work.

Although this research is situated in the context of an educational environment, the benefits of this work are not limited to the group under study. The combination of functionalities provided by such a rich environment makes it a suitable tool for knowledge management in communities or business organizations as well. Examples of uses of Elgg outside of the classroom context include the GUSSE project [11], where it is used to establish a community of interest in urban sustainability over the world, and France Telecom R&D [8], where it is used to browse among people, files, blogs, etc. in search of competencies within the company's social network. As such, the group under study is only a small representative of a more general population that can benefit from a safe place for sharing problems, experiences, and resources with others. Ensuring privacy of information in this kind of environments significantly contributes to creating a trusted network for sharing both the process and the product of knowledge transfer, benefiting all potential knowledge practitioners.

12. REFERENCES

[1] Adams, A., Users' perception of privacy in multimedia communication, Extended Abstract, *in Proceedings of CHI 1999*, Pittsburgh

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

[3] Adams, A., Sasse, M. A., Privacy in multimedia communications: protecting users not just data, *in Proceedings of IMH HCI*, 2001, pages 49 – 64

[4] Batson, T., Electronic Portfolio Boom: What's it all about? Syllabus 2002, available online at http://www.syllabus.com/article.asp?id=6984

[5] Cohn, E. R., Hibbitts, B. J., Beyond the Electronic Portfolio: A Lifetime Personal Web Space*, 2004, available online at http://www.educause.edu/apps/eq/eqm04/eqm0441.asp.

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

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

[8] The Elgg project home page: http://elgg.net/

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

[10] Gathercoal, P., Love, D., Bryde, B., McKean, G., On Implementing Web-Based Electronic Portfolios, *in Educause Quarterly 25(2),* 2002, pages 29-37, available online at http://www.educause.edu/ir/library/pdf/eqm0224.pdf

[11] The GUSSE project home page: http://gusse.org/

[12] Karvonen, K., Creating Trust, in *Proceedings of the Nordsec'99, the Fourth Nordic Workshop on Secure IT systems, November 1999*, Kista, Sweden

[13] Lewicki, R.J., Bunker, B., Developing and maintaining trust in work relationships, *in Trust in Organizations: Frontiers of Theory and Research*, Sage, Newbury Park, CA, 1996, pp. 114–139

[14] Love, D., McKean, G., Gathercoal, P., Portfolios to Webfolios and Beyond: Levels of Maturation, *in Educause Quarterly*, Volume 27 Number 2, May 2004, available online at http://www.educause.edu/pub/eq/eqm04/eqm0423.asp?print=yes

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

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

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

[18] Tosh, D., Werdmuller, B., ePortfolios and weblogs: one vision for ePortfolio development, March 2004, available online at http://www.eradc.org/papers/ePortfolio_Weblog.pdf

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