Understanding Community Privacy through Focus Group Studies

Sherley Codio

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
In
Computer Science & Application

Dennis G. Kafura
Manuel A. Perez-Quinones
Andrea L. Kavanaugh

May 1st, 2012
Blacksburg, Virginia

Keywords: Community privacy, Focus group studies

Understanding Community Privacy through Focus Group Studies

Sherley Codio

ABSTRACT

Just as an individual is rightly concerned about the privacy of their personally identifying information, so also is a group of people, a community, concerned about the privacy of sensitive information entrusted to their care. Our research seeks to develop a better understanding of the factors contributing to the sensitivity of community information, of the privacy threats that are recognized by the community, and of the means by which the community attempts to fulfill their privacy responsibilities. We are also interested in seeing how the elements of a community privacy model that we developed are related to the findings from the studies of communities. This thesis presents the results of a series of focus group sessions conducted in corporate settings. Three focus group interviews were conducted using participants from two information technology companies and one research group from the university. Three themes emerged from the analysis of these focus group interviews which are described as privacy awareness, situated disclosures, and confinement of sensitive information. These three themes capture the character and complexity of community oriented privacy, and expose breakdowns in current approaches.

Dedication

To my parents, Francoise and Ruben Codio who are the greatest fans of my life

And to my best friend Cathy Bernard for her support and encouragement

Acknowledgement

Thanks to God for providing me strength and motivation to complete my thesis. I would like to take this opportunity to thank Dr Dennis kafura for his guidance in this project. I would also like to thank my other committee members: Dr Manuel A. Perez-Quinones, Dr Andrea Kavanaugh for all their help in the project, as well as the other members of the research group Dr Denis Gracanin, Peter Radics and Tom Dehart who were helpful in discussing the focus group responses.

I would also like to acknowledge some people and organizations who contribute to my education and success today. Below are a few names out of many:

Dr Patrick Guilbaud

The Higher Education Development (HED) Program

The Google Foundation

The Office of International Research, Education, and Development

My family members for their constant support

My friends who always checked on me to ensure that I was making progress

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Personal information privacy of an individual is a recognized and widely studied issue. Research in this area is multifaceted, including analysis of legal principles [51], studies of sociological factors [6], and substantial work related to the design and operation of information systems. Some research seeks to provide usable means for an individual to influence or control the dissemination of personally identifying information [18] or to be better informed about how the information they disclose might be used [37, 38]. Other research is aimed at preserving the anonymity of individuals when data mining aggregated information [5]. The common theme in all of this research is preserving the privacy of an individual's identity or personally identifying information. While individual privacy is a critical technical and societal issue, there are other forms of privacy of equal importance.

Another, though less studied, form of privacy is how a community of people collectively safeguard sensitive information that they share in the course of their activities. In this thesis, we draw on Wenger's concept of community of practice [16] to define a community as a group of people who regularly interact and collaborate on a set of shared tasks or goals. We also define the term *privacy-aware community* as a group of people known to each other who recognize a mutual obligation to limit the dissemination of sensitive information that they share in the course of their collaborative professional, business, or social activities. In the remainder of the thesis, we use the term *community privacy* or the shorter term *community* to mean the limitation on the dissemination of sensitive information known to the community. Specific examples of communities sharing sensitive information are:

- proprietary technical specifications within a corporate research and development team,

- competitive business data between companies in supply-chain relationships,
- electronic medical records involving care providers and a patient or the patient's extended family, and
- personal communications among family members or close friends.

As seen in these examples, the collective responsibility of each individual to protect shared, sensitive information known to the community is distinct from an individual limiting access to their own personally identifying information. That is, community privacy is distinct from individual privacy.

A community differs from collections of users such as circles [28], virtual organizations [26], or groups (in access control and encryption [49]). These collections of users are defined by the entity owing the shared resources; they do not require the identified users to participate in the collection's definition, agree to be in the collection or, in many cases, even be aware of the collection's existence. For example, Bob can put Alice in one of his circles without Alice's consent or knowledge. While Bob understands the meaning and intent of the circle he has created, Alice has no knowledge of this. Alice is also free to define circles of her own choosing that may include Bob. There is no requirement for the circles defined by Bob and Alice to be related in any way because they are constructs for *individual* privacy protection. Communities, however, involve cooperation by the community members in the formation and operation of the community. Other collections of users are those defined in group key encryption. Here, the members form a passive *audience* that receives information broadcast to them. This is typically the scenario in restricted broadcast media. In contract, individuals in a community actively collaborate in the use and sharing of content. Finally, some recent approaches to individual privacy incorporate information from *crowds* of anonymous (to each other) users. These crowds

may provide information from their own information usage which may benefit others. Community members, however, interact in a deeper and richer manner. They are known to each other and share common goals and resources. Thus, our sense of community differs from owner centric circles, passive audiences, and anonymous crowds. Further discussion of communities is given in the second section of this chapter and in the related work described in CHAPTER 2.

As defined above, community privacy denotes the limitations on the dissemination of the sensitive information shared within the community. As have others, we adopt the view that privacy is a "boundary regulation" process [40, 47] in which privacy decisions are highly context dependent [46]. Because the notion of "context" is difficult to capture [22], we approach privacy controls as a dynamic process of user-driven adjustment under changing circumstances rather than a fixed set of "rules" or "policies" that are automatically enforced. We take it, of course, that information systems supporting community activities will preserve privacy *in addition to* enforcing security. Security in this sense focuses on confidentiality, integrity, and availability and gives preeminence to automatically enforced rules and policies.

The needs of communities must also be viewed through the lens of privacy because this perspective examines information systems from a socio-technical perspective that can account for the human and social factors critical to the trustworthy operation of the information system. Security systems that neglect people as a significant part of the equation "are seldom secure in practice" [8]. Practice is what happens in the moment; it is the activity; it is what is actually done. It is often in the human-centered moment, and not in the computer-centered planning stages, when security policies or mechanisms break down and the safety of sensitive information is compromised [1, 20]. When a breakdown occurs in a social system (as opposed to a computational one) workers do not stop what they are doing. Instead, they create special cases or

methods that allow them to continue by bending formal work policies, e.g. when users write passwords on post-it notes, or shout them across the room, as observed in this work [1, 20]. As a result, there exists a need to study and understand the holistic practice of how security is managed in socio-technical systems [20]. The work presented here not only addresses the need for more practice-based study, but also explores community-centric themes of privacy practice.

Specific threats to community privacy include accidental disclosure violating known privacy expectations [61], lack of awareness of privacy expectations in a given context [40], inability to determine if a given disclosure meets the privacy expectations of the community, inability to use the privacy interfaces to achieve the correct sharing [27, 60], and inability of the system to guarantee the desired privacy. These threats can lead to significant harms to individuals and the community. While we recognize that tremendous damage that can be caused by external attacks (e.g., rootkit attacks, denial of service attacks), we also see that there are significant harms that result from unintended disclosures by well-meaning individuals. For example, one report indicates that the vast majority (80%) of improper disclosures come from inside the organization ([44] cited in [42]). There is also evidence that accidents and errors, rather than malicious intent, account for the majority of information leaks ([29] cited in [42]). Such inadvertent leaks can be costly, as illustrated by the case of an email error that led to a $1 billion loss (example cited in [61]). Our case studies also provided evidence of the occurrence of these types of inadvertent disclosures (see CHAPTER 4).

A community based view of privacy has intuitive and practical motivations. In the real world, privacy is ensured not only by individual means (doors, locks, alarms, fences) but also by community means (neighborhood watch, community policing). In practical terms, community-based control has three advantages over individual control: (1) better utilization of privacy

mechanisms due to sharing of expertise; (2) better awareness of privacy requirements due to sharing of knowledge about privacy threats; and (3) better attention to frequently neglected privacy tasks [33] due to a sense of responsibility to the community.

Achieving community privacy, the focus of this thesis, is a challenge to research because (1) approaching privacy from a community perspective is relatively new, (2) individuals are in multiple and overlapping communities, each having different privacy requirements, and (3) privacy is ill-suited to being captured in fixed security rules or policies because it is a "boundary adjustment" process that is dynamic and context-dependent.

In the thesis we report the findings from the focus group study of community privacy in three organizations: two technology companies, and one research group from the university. The study details are presented in CHAPTER 3. The purpose of these studies was to understand from a privacy perspective relevant community characteristics and practices, and to assess the elements of a community privacy model (see Section 2 below). We recognize, of course, the limited interpretation that can be given to our studies. The limitations are discussed in the conclusion chapter, CHAPTER 6. However, we believe that there are novel aspects of our perspective on privacy and there are interesting outcomes of these studies.

In the remainder of the thesis we present in Section 2 of this chapter the key elements of our a priori privacy model: community, tagging, exceptions, and notifications. In chapter 2 we review work most closely related to our own. In CHAPTER 3 we describe the details of the focus group interview involving participants from one organization. In CHAPTER 4 we present the analysis of the privacy features and practices of the communities within the organizations studied. The analysis of the focus group interviews identified three major themes: privacy-awareness, situation actions, and confinement of sensitive information. The discussion of the analysis in

CHAPTER 5 explains support that was found for four key features of our community privacy model: community privacy, exceptions, tagging, and notifications. The discussion also draws out the similarities between our observations of community privacy, as revealed in the focus group interview, and legal concepts of individual privacy and individual privacy behaviors. Thus, we are encouraged to believe that what we are observing is properly seen as a privacy phenomenon. Brief conclusions are given in CHAPTER 6.

KEY CONCEPTS

In this section we review our a priori community privacy model [32] and describe four elements of this model which were relevant to the design of our interview protocol and the analysis of the focus group responses. Figure 1 is a representation of the community privacy model.

As shown in Figure 1, users have a local *interface* by which they share with each other information *artifacts* through some *channel* of communication. In a concrete system the artifacts may be information resources like email messages, files, or images, the interfaces maybe be software components like email clients or web browsers, and the channels may be communication systems like email servers, file systems, or streaming media. The system mediates the information sharing among users so that the dissemination of information is in line with the users' privacy expectations. These privacy expectations are expressed by the, one or more, tags attached to the artifacts by the users. Each tag names a group of users. In order for a user to receive a given artifact that user must be identified in all of the tags attached to that artifact. The tags, once assigned, are permanent and cannot be removed. The system also makes visible the actions of one user that might be of interest to other users through a *notification*. Finally, the users may agree to permit an *exception*, sharing that is contrary to the dissemination otherwise allowed. Four elements of this model are central to this study: (1) a community is

defined by the users identified by a tag, (2) tags are attached to artifacts, (3) exceptions triggered by the request of a user to seek other users' agreement on sharing of an artifact that would not otherwise be allowed, and (4) notifications sent to a community when an event of interest occurs (e.g., a user attempted an impermissible sharing of an artifact). This model shaped the development of the questionnaire used to guide the individual and focus group interviews.

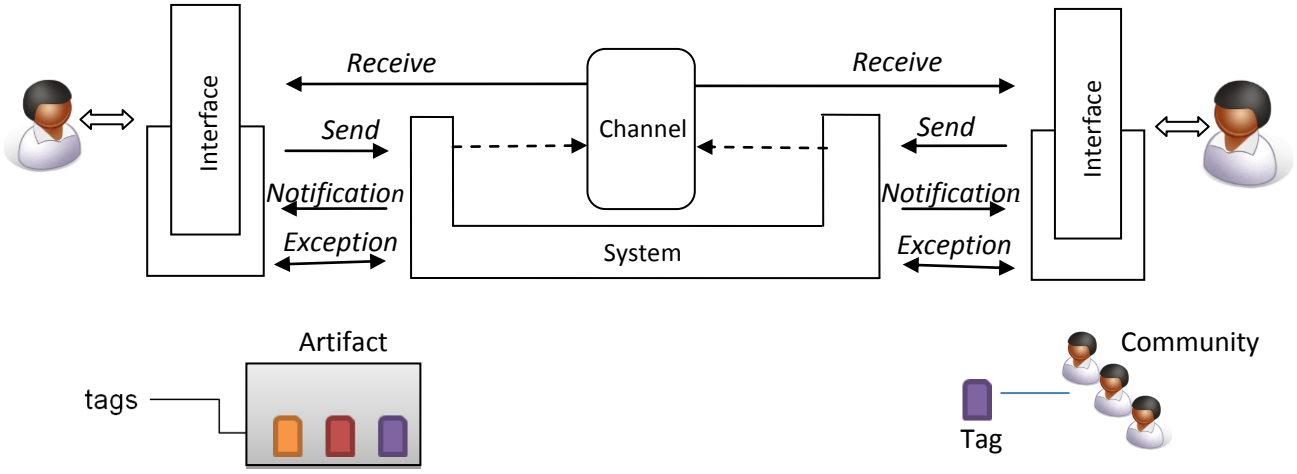

Figure 1: Community Privacy Model

The model and the four key elements are further described and then illustrated by several generic scenarios. These scenarios are given merely to illustrate the elements of the community privacy model and are not related to the situations described in any of the focus group interviews.

Community

We characterize a "community" by three attributes: collaboration, cooperation, and collective actions. Collaboration means that a group of people are affiliated to achieve some purpose based on a task-oriented goals or on a shared concern. Cooperation means that members of the group

are well disposed toward the group's privacy goals. Thus, a deliberate violation of the group's privacy (i.e., an insider threat) is not considered. Collective actions mean that the members of the group share a sense of responsibility for protecting the privacy of the group's information. Thus, members of the community can be expected to employ privacy controls so long as these controls are sufficiently usable.

Tagging

Community members annotate, or tag, sensitive information with metadata to define and enforce community privacy. The metadata, a *community tag*, is a named structure that identifies the members of the community and is used to annotate artifacts (information, devices, places) that are within the community's privacy boundary. Artifacts not so labeled are outside the community's privacy boundary.

The purpose of the community tag is to regulate the visibility and dissemination of tagged artifacts so that the privacy intentions of the community are realized. The tag is permanent in that it cannot be removed from the artifact. The tag is propagated so that information derived from a tagged item carries the tag of that item. Thus, derived information stays within the privacy boundary defined by its source(s). For example, an email with a community tag can only be replied, sent or forwarded to some subset of members of the community associated with the email's tag.

The community identified by the tag defines the privacy boundary for the tagged information; that is, the tag defines the set of individuals who are allowed to change that boundary (by changing the composition of the community or approving exceptions). In this way, the community is invested with authority by being empowered to define and regulate its privacy

boundary. Depending on the nature of the community, all members of the community may be involved in regulating the boundary that defines the privacy of the community's information.

Exceptions

The community is also empowered to make *exceptions*, that is, agree to permit a disclosure that is contrary to the existing privacy boundary. The exception does not change the privacy boundary. Rather the exception allows an extraordinary action to be taken when the community is so disposed. Exceptions are seen as necessary to account for situations where unusual situations arise that compel the extraordinary action to be taken. We will see in our focus group analysis cases of competing privacy priorities that must be resolved and compelling community interests that push sensitive information outside of its normal channels.

Notifications

Through *notifications* the actions of community members are made visible, allowing all members of the community to be aware of specified actions taken by others in the community. The community tag specifies the conditions under which notifications are sent. Notifications give additional substance to the notion of community because it allows the community to be aware of potential threats due to attempts by its members to make inappropriate disclosures. It also provides a way for the community to be aware of activities within the boundary that may have special significance. For example, the community working on a document may be interested in knowing that there is an intense interaction between two of their members on a key part of the report.

The following scenarios illustrate the ideas of *community*, *tags*, *notifications*, and *exceptions*.

Scenario 1: Community tags. Three members of a larger planning group want to share drafts of an outline among themselves. The information shared among the planning group is tagged with the community tag *Planning*. The three members form a sub-group by creating a new community tag, *Outline*, and use this tag to keep their sub-group activity private. Information objects tagged with *Planning* are visible to the planning group, but information objects tagged with either *Outline* or with both *Outline* and *Planning* are visible only to the sub-group. The planning group needs some specialized input from a person outside their group. A new community tag, *Consultant*, is created that includes the outside person and all members of the planning group. Information objects tagged with *Consultant* are visible to all.

Scenario 2: Notifications. An intern joins the planning group and is added to the *Planning* community tag. The intern does not correctly understand the privacy restrictions associated with the planning group's activities and tries repeatedly (but unsuccessfully) to share the group's information with other interns. A notification is sent by the system to members of the planning group alerting the intern's mentor to better inform the intern of the privacy restrictions.

Scenario 3: Exceptions. A member of the planning group receives a call from the company's CEO asking for a status report. Rather than adding the CEO to the *Planning* community tag for this single request the planning group agrees to make an exception and allow the current draft to be sent to the CEO.

As illustrated by these scenarios, our privacy model mirrors and supports the collective actions of individuals who are entrusted with preserving the privacy of sensitive information. For example, the community tags *Planning* and *Outline* mirror the structure of the planning group and its sub-group. Significantly, our model allows communities to define and regulate the privacy boundary by changing the community membership. When a new member joins or a

current member leaves the planning group, the *Planning* community tag can be appropriately changed. Exceptions provide a principled way to deal with unanticipated, context-dependent situations. For example, the planning group was able to choose to divulge its draft to the CEO as an exception. Notifications provide visibility and accountability because the acts of disclosing community information cannot be done surreptitiously but are instead made observable to the group. This privacy model was part of our research agenda when we undertook the studies reported in this thesis. In part, the studies were intended to assess the degree to which the model elements would be useful in the practical situations described by the studies' participants.

CHAPTER 2

RELATED WORK

In this chapter we present a review of related prior work. This review focuses on four topics: privacy, privacy awareness, usability of privacy, and boundary regulation.

Privacy

Privacy related to the design of information sharing technology is an important area of research. It has been argued, for example, that privacy concerns should be taken into consideration during the design process of technologies [47]. Cavoukian has been advising against treating privacy as an afterthought for over two decades; she proposed the concept of "privacy by design" [14].

However, as concerned as designers and researchers seem to be about privacy there is no common definition of the term, with different people associating different meanings to the term privacy [13]. For example, communication privacy management treats privacy as "the feeling that one has the right to own private information, either personally or collectively; consequently, boundaries mark ownership lines for individuals [48]." The sociologist Altman views privacy as the boundary regulation between private information and disclosed information [6]. Caine considers that different pieces of information have different level of "privateness" which ranges from very private to not really private.

In this thesis we take the view that privacy is a "boundary regulation" process [40, 47] in which privacy decisions are highly context dependent [46]. More specifically, community privacy is defined as the limitations on the dissemination of the sensitive information shared within the community. These "limitations" are seen as changing in response to contextual circumstances and by a social process within the community.

Privacy Awareness

Previous work has focused on preserving the privacy of an individual's identity or personally identifying information. Examples of such work, [45, 17] extending access control models for individual privacy, [18] providing privacy-enhancing technologies for web-services, or geo-location [30]. While preserving individual privacy, which is greatly affected by technology in use today, equally important is privacy consideration where a group of people share sensitive information.

Community privacy protection differs from individual privacy protection because it explores the complexity and capability that comes from members of a community collaborating to achieve their privacy goals. Also, this previous work focused only on personally identifying information while the proposed work focuses on the more general class of sensitive information. That is, medical information about a patient is sensitive information which a health-care providing community (doctors, nurses, technicians, etc.) wants to keep private among themselves but which is not personally-identifying information about any member of the community.

The sense of community among a group of people distinguish this work from related work which treats the group as a crowd – a collection of anonymous individuals that serve as a source of experiential information but are otherwise disengaged from each other. This other work can be characterized as *community-assisted* because the community is a source of data by which users can assess risks to their privacy. Unlike communities in our sense, in these approaches there is little or no direct or extended collaboration among the users. Nascent work on community-assisted privacy protection has been done in four areas: social networking, email communication, web services, and geographic location.

Social networking. Recently, social networking services have added circles (Google+) and groups (Facebook) that are defined by an individual to control that individual's posted information. A community, however, is defined by agreement among the group members to control the community's shared information. This distinction is fundamental. Also, circles and groups lack community mechanisms for exceptions and notifications. Group consensus for privacy settings on photographs in a social networking site has been studied [53, 54]; and collaborative strategies to cope with online disclosure on social networks have been proposed by [40]. These creative approaches use game theory to devise a fair voting protocol through which a group of individuals can agree on the privacy settings for photographs. The annotation attached to the photograph can also be used subsequently to recommend privacy settings for similarly annotated photographs. This work is similar to ours due to the concern with community privacy control, the use of annotations (tags), and the provision of a consensus mechanism.

The approach in our privacy model is distinct from this work in several significant ways: (1) we include other community mechanisms (e.g., notification and exceptions), (2) the sense of groups is different, and (3) tags in our model are the means of relating the group to the controlled item while in social networks the tags are used only to suggest similarity between items; identical tagging need not imply identical privacy control.

Email communication. One study [61] examined a large corpus of corporate email. Through machine learning techniques it was possible to identify interacting groups of users. This work improved on early email analysis by showing how to reduce the rate of false positives (i.e., misidentifying legal communicators as illegal ones) by a cascading approach which combines group analysis and individual communication patterns. Aside from the common concern with email there are many differences between this work and our own. The study of email patterns

attempts to infer what users in our model state directly (the group of concerned users) and is analytically-oriented as opposed to enforcement-oriented. For example, it is not clear how to use the analytical approach as an effective enforcement technique in cases of first use (where there is no history to analyze), or consensus among users on joining additional members to a group.

Important work has been done on group email privacy. Exemplars include systems using mail lists [10] or attribute-based encryption [9]. Our approach differs from this work in several respects. First, we include ubiquity where email is only one of several communication mechanisms that must be included. Second, our notion of community does not require a moderator as in [10] nor require its members to be within a single organization as in [9]. Also, these approaches do not include important community mechanisms such as exceptions and notifications.

Web services. A collaborative privacy management approach has been added to PRIME [37, 38]. In this approach the P3P technologies allow users to report their experience with the privacy provided by web services. These reports are stored in a shared repository so that users can compare their privacy preferences not only to the policies stated by the web service but to the actual experience reported in the shared repository. The differences between this work and our work include: (1) in PRIME the members of the community only share their experience but do not actively collaborate in any deep sense as in our model, (2) the information sharing model is different since in the case of PRIME the information is disclosed by the user to the web service while in our approach the users are sharing information with each other, and (3) PRIME enables an awareness of privacy threats while we enable collaborative enforcement of and negotiation about privacy boundaries.

Geographic location. In Spybuster [36] users can tag geographic locations which have surveillance cameras. These geo-tags are made available to other users who, on a radar-like display on their handheld device, can be aware of the surveillance near their current location. In Aegis [43], a game theoretic technique is used to reveal a user's current geographic location only to that user's nearby friends. Aegis is collaborative because it enables the disclosure of geographic location among a group of voluntarily participating users. Lastly, the "alibi phone" [22] presents artificial context information on behalf of a user wishing to conceal their current location. This system is collaborative because it relies on the artificial context information being provided by other users who have access to this context. While sharing a common concern for community, our model differs from these works in the type of information being protected, and the degree of interaction among community members. The cited systems promote geo-awareness while our approach enables collaborative enforcement of and negotiation about the privacy of shared information.

Usability

Prior work has explored several different aspects of the usable security and privacy problem. For example, there has been a large body of research and development of technologies that support educating, informing, and understanding behaviors of the user [50]. The work of Ackerman [2] has suggested the use of privacy labels so that users can be notified of possible data capture, use, and reuse as they interact with web-based technologies. Another example is the work of Cranor et al. called Privacy Bird [19]. Privacy Bird reads the P3P (Platform for Privacy Preferences Project) [19] information in a website and then displays a visual indicator of how closely the privacy standards for that website match personal settings (e.g., green indicator means a good match). Karat, Karat and Bridie have designed SPARCLE, a policy authoring and transformation

tools to “enable organizations to create and transform natural languages policies into machine readable code for real time enforcement decision. [35]” Caine explored individuals’ behaviors to design systems for privacy. Her work consisted of understanding the nature of privacy concerns, how they influence behavior in order to provide technology solutions to these concerns [13]. Our work examines a different aspect of usable privacy; we examine current practices of how other people manage the privacy of their clients as a method to inform future technology design.

Bellotti and Sellen [11] present a design framework for understanding privacy and security needs through the aspects of user feedback and control when creating ubiquitous technologies. Similarly, the work of Flechais et al. [25] demonstrated the difference between social and technical security measures. By their definitions, human-based security measures are progressive and adaptive yet are unreliable due to the effects of emotions and variable circumstances. Technical security, on the other hand, works well on repetitive tasks, but is less capable of being flexible in unanticipated situations. These studies emphasize the importance of understanding both the social and technical context of the user to provide adaptive security needs. In our studies we explore how security is affected by both technical and social conditions. We do this to provide a holistic conception of security in socio-technical systems that accounts for trust, privacy, and negotiation with respect to social and technical security mechanisms.

Boundary Regulation

Privacy as an interpersonal boundary process by which people control their interactions with others was first expressed by Altman [6] in a social context, and was later interpreted by others [47] in a technical context. In this view privacy is seen as an individual's dynamic adjustment the boundary separating disclosed information (information outside the boundary) and private information (information inside the boundary). The process is dynamic in that the boundary is regulated with changes in context. For example, if an individual assumes they don't have enough privacy, they may choose to reduce their interactions with the outside world; similarly, if they have more privacy than desired they may also choose to increase their interactions with the outside world [13]. These mechanisms for the regulation of privacy are described by Altman as the interpersonal boundary regulation.

Stanton [56] defined a framework called information boundary theory. His framework requires individuals to follow rules for "boundary opening" and "boundary closure" for the uses of information technology. "Boundary opening and closure are dynamic, psychological processes of regulation by which people attempt to control the flow of "intimate" information [56]." Petrinio [48] builds his on Altman's work and defines boundaries as permeable or impregnable and may be linked with other privacy boundaries. The lines of ownership are not always clear. According to Petrinio, boundary management may not always work when there is invasion from outside sources [48].

The concept of boundary regulation was an early guide for our work as we considered the nuanced and context-dependent situations reported in the interviews. However, while this concept served as useful inspiration we were unclear about how to apply it in many cases.

CHAPTER 3

METHOD

This chapter presents: a description of how the survey was created and the focus group interviews conducted, and (2) a description of the qualitative analysis method, called framework analysis. Together, these two elements formed the basis for the generation and analysis of the data whose major themes are presented in the next chapter.

Interview Process

For this study, we developed an interview protocol that was approved by Virginia Tech's Institutional Review Board (IRB). Materials used in these studies included a questionnaire, a focus group script and two audio recording devices. The questionnaire was developed to probe the four elements of our community privacy model and consisted of five main categories of questions. The complete questionnaire can be found in the appendix. The five top-level categories of questions are:

- Can you tell us about the group?
- Can you tell us about how the group works and makes decisions?
- Can you tell us about how the group keeps information private?
- Can you tell us how the group shares private information with someone outside the group?

We designed the focus group script to encourage participation in the discussion of the five main questions as well as the more specific probes.

To recruit participants, we contacted several local companies and asked them to participate in our study. We targeted industry participants because they deal with sensitive information either related to the company directly or to clients. We conducted two individual interviews before the

first focus group session. One of the individual interviewees also participated in the first focus group discussion. The individual interviews helped us to develop the probe questions for the focus group sessions to be conducted. The individual interviews lasted each an hour. We conducted three focus group sessions for this study; all three focus group sessions were during business days. Each of the focus group interviews were conducted on site in conference rooms made available by the focus group's organization. In total there were sixteen participants.

The following paragraphs describe the general structure of each focus group session, each of the focus group's organizations, and the process used to generate the record of the focus group session.

Each focus group session began by thanking the participants for being available, and then we explained the purpose of the study. We informed participants that there were no right or wrong answers for any given question, thus they could freely and openly express their opinions and experiences. It was also noted that while the study was concerned with sharing of sensitive information, the participants were not being asked to disclose any sensitive information. They were also given the opportunity to contact the research team afterwards if they believed they had inadvertently revealed anything that should be obfuscated in the transcript of the interview. No such requests were made of the research team. We informed the participants that an audio recording would be made of the interview, the place where the audio files would be securely stored, as well as the precautions taken to keep their identities confidential. We asked the participants to read and sign the consent form before the discussion started. It was not explicitly stated that participants could leave at any time, but participants were allowed to leave at any time.

Once the consent forms were signed, the moderator began to ask questions from the questionnaire to open the discussion. The moderator moved on to the next question once the discussion following a probe ran its course. Sometimes participants would bring up topics relevant to other probes that had not yet been discussed or would refer back to probes previously discussed. The moderator would follow up on these topics and skip them when they appeared later in the questionnaire or would come back later to ask specific questions about the topics for more details.

Group Characteristics

Participants for the first focus group session came from an information technology company providing email, application support, and cloud hosting services to its customers. The company is geographically distributed in multiple locations across the United States with a total of approximately 4000 employees. As a cloud computing company, they are very concerned about security issues. For that reason they employ several mechanisms to ensure a high degree of security such as the use of a virtual private network to access their internal network remotely, frequent mandatory changes of passwords, and encrypted file systems for some data. The seven participants were from different divisions within the company; some of them were from human resources, others from software engineering or product development. Several of the participants also played management or supervisory roles. The focus group session lasted one hour and twenty-five minutes.

Participants for the second focus group came from a research group here at Virginia Tech. It is a group that works closely with numerous offices on campus. The group members deal on a daily basis with students and faculty data for their analysis and reports. The group normally has ten or

eleven people with different backgrounds, including: industrial system engineering, political science, statistics, mathematics, and criminal justice. Five of their team members joined our focus group discussion. The session lasted an hour and twenty six minutes.

The third focus group session consisted of participants from a privately held technology company that provides asset performance management software solutions. The company is geographically distributed with offices in many locations in the United States and around the world. A total of five employees from the company participated in the focus sessions. One of them remotely participated in the discussion via speakerphone. These participants came from different divisions within the company; some came from software engineering, other from the legal department or usability engineering. Several of the participants also played management or supervisory roles. The session originally started with four participants, the fifth participant joined the group during the discussion. After fifty seven minutes, one of the first four participants left the discussion due to another appointment. The entire session lasted an hour and twenty six minutes.

Constraints and Goals

We chose to conduct the focus group discussions with these three because they made themselves available to us. We contacted several companies from an industrial affiliated group, companies collocated in a research park, and groups in our institution, but these three were the ones agreeing to participate in the study. We deliberately targeted structured organizations because we thought such organizations would be able to acknowledge the issues we are trying to study. We decided to interview a set of similar participants for the possibility of drawing stronger

conclusions. We know that the choices we made for the focus group studies, voluntary or imposed on us, have impacted the limitations of our conclusions.

Recording Process

We recorded the focus group sessions and the individual interviews on two different devices for reliability and accuracy. The first device is an Olympus VN-6200 and the second device is the TagPad app on an iPad 2. The TagPad is an app that allows the iPad to record interviews with an animated interview schedule. The audio files from the iPad were transferred via Dropbox to the computer where they were stored and analyzed; the audio files from the Olympus VN-6200 were transferred to the computer using an USB cable. For the purposes of the analysis reported here, we partially transcribed the individual interviews and fully transcribed the focus group session. We assigned each participant an anonymous identifier of the form*GxPy where X* represents the focus session (first, second or third) and *y* represents the participant semi-random number from 1 to n based on the number of people present at the focus group session. For example, the first participant from the first focus group would be identified as G1P1.

Analysis Process

Framework analysis [39] is the method that guided our analysis process. Framework Analysis is a flexible method that can be adapted to research with specific questions, a limited time frame, and a priori issues [55]. This form of analysis allows for the inclusion of a priori as well as emergent concepts, for example in coding. We were motivated to use Framework Analysis for two reasons. First, Framework Analysis is explicitly designed to assess a priori beliefs or models. This matched our situation where there was an existing privacy model whose elements were to

be tested and refined through the analysis of the focus group interviews. Second, although Framework Analysis was originally developed for policy analysis, we believed that the elements of the a priori privacy model could also be analyzed using this method. In an intuitive sense, the privacy model was akin to a "policy" to be followed to achieve the desired degree of privacy. The focus group testing then amounted to an assessment of whether this "policy" conformed to the requirements exhibited in the focus groups.

The steps in the Framework Analysis are shown in Figure 2 and explain in the following.

We followed for out of the five steps proposed by framework analysis. We analyzed the data for each focus group separately using four steps of the framework analysis process. The five steps process consists of:

1. Familiarization
2. Identifying a thematic framework
3. Indexing
4. Charting
5. Mapping and Interpretation

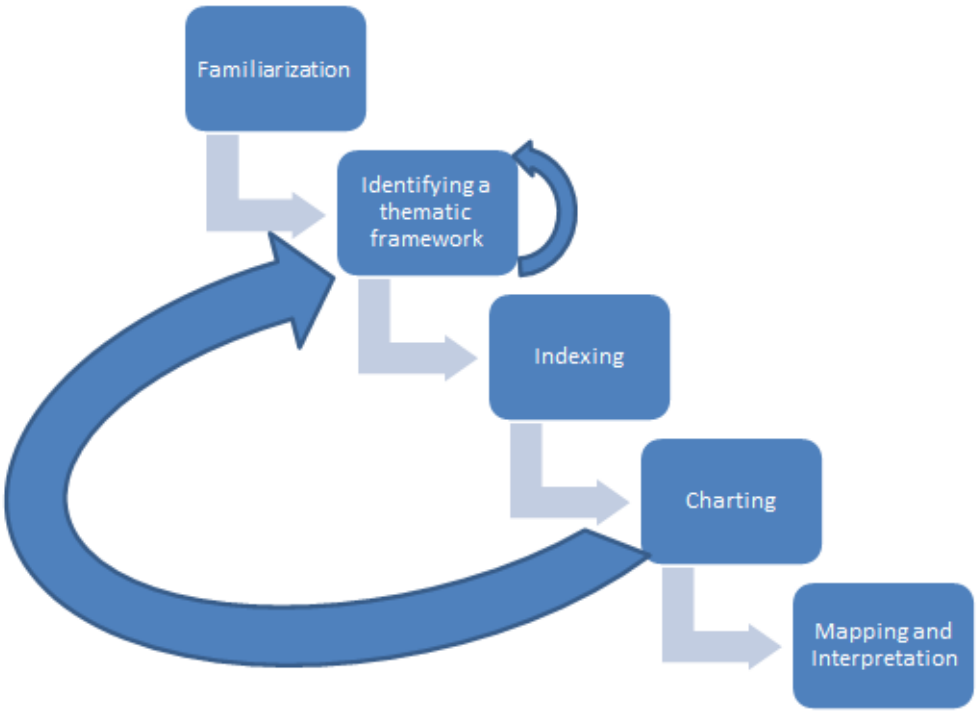

Figure 2: Modified Version of Framework Analysis

In step one, the familiarization step, two members of the research team listened to the audio recordings and consulted the transcripts and the notes to gain an overview of the data collected [55]. Through discussion they sought to determine the meaning and significance of each statement made in the interview.

For step two, identifying a thematic framework, Three multi-hour sessions were needed to review the first focus group interview and to create approximately twenty categories into which the statements could be placed (see Figures 3 and 4). The two researchers spent additional time reviewing the transcripts and statements between the face-to-face meetings. In some cases a given comment was placed in more than one category. Some comments were peripheral or lacked sufficient significance to be included. For example, descriptions of the company's

security mechanisms (e.g., use of virtual private networks) were not directly germane to our privacy interests. Approximately ninety statements were organized into these initial categories. For the second and third focus group data, since we have already gone over the process, we used the first focus group categories as a base and seek out for the emergence of new categories.

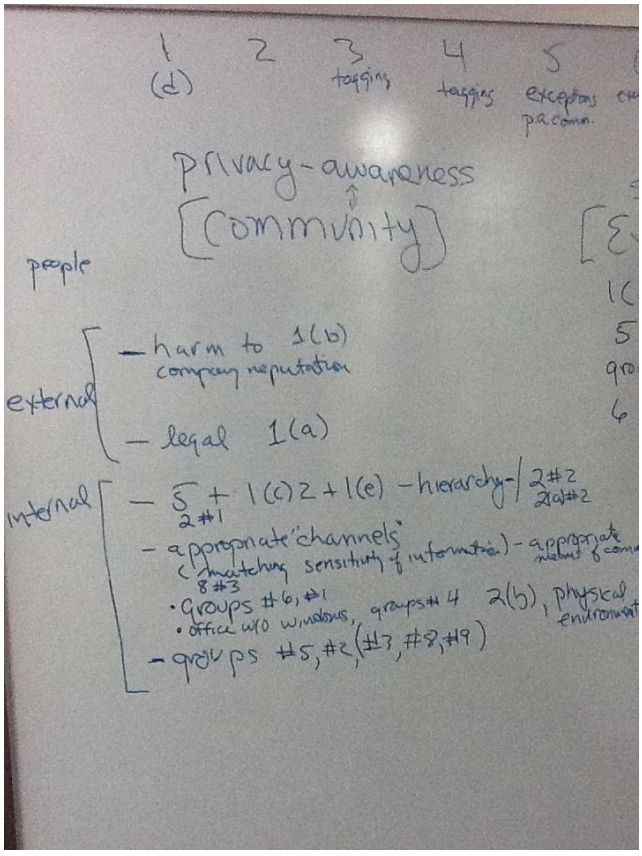

Figure 3: Emerging Themes from Analysis

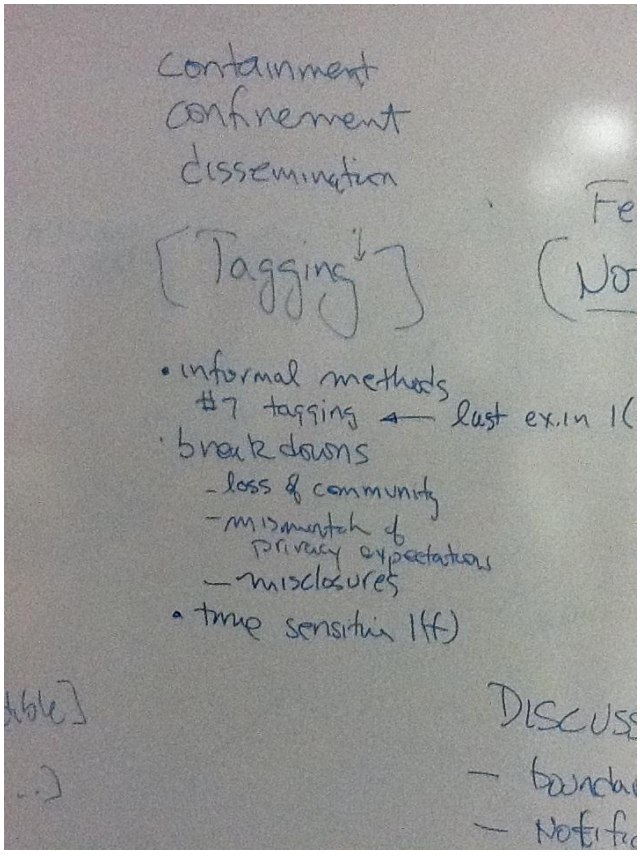

Figure 4: Emerging Themes from Analysis

In the third step, Indexing, the same two members of the research team worked together to abstract the initial categories, rearranging the categories and moving statements across categories to where they best fit. As progress was made in the analysis, for the first focus group analysis six categories emerged: factors of sensitivity, boundary regulation, group structure, metaphors, types

of privacy behaviors, and types of unintentional disclosure. For the second and third focus group, more categories emerged. This step consisted of several meetings each lasting one hour or more. Later on, three other members from the research team helped refine the categories and discuss whether a specific statement of group of statements belonged to a given category. These meetings suggested some further possibility of abstraction and some alternative views on the meaning of the categories.

In the fourth step, Charting, the initial two members of the research team focused their attention on the dominant themes that emerged from the earlier analysis, setting aside some issues that seemed less relevant and further combining and reorganizing the categorization.

In the fifth step, Mapping and Interpretation, the dominant themes and categories were related to the elements of the privacy model and to other known privacy concepts. The relative strength of support for each category and sub-category was also determined.

The results of steps 1-4 are given in Chapter 4 (Analysis) while the results of step 5 (Mapping and Interpretation) are given in Chapter 5 (Discussion).

CHAPTER 4

ANALYSIS

In this section we present the qualitative analysis of the focus group discussion we conducted. Framework analysis is the process that we followed to analyze the data. Three general categories emerge from the data analysis and are described below. The three categories are: privacy awareness, situated disclosures, and confinement of sensitive information

Privacy Awareness

As defined in the introduction section *privacy-aware community* is a group of people known to each other who recognize an obligation to limit the dissemination of sensitive information which they share. Thus, we were interested in knowing: (1) whether we could identify in the participants' responses descriptions of behavior that evidenced the existence of a privacy-aware community, and (2) what circumstances and factors affected the behavior of individuals within the community.

Table 1 summarizes the factors related to privacy awareness that were found in our analysis and are discussed in this section.

We found description of situations where the participants described experiences where they acted to control the dissemination of sensitive information to parties *external* to the organization as well as controlling dissemination *internal* to the organization. We consider these two types of situations, external vs. internal, separately.

Table 1

Privacy Awareness Control over the dissemination of sensitive information	
External (factors)	Legal considerations • General legal factors • Imposed legal factors • Voluntary legal factors
	Reputation considerations
	Misrepresentation of information
	Agreement on information releasing
	Competitive advantage
Internal (mechanisms)	Explicit communities
	Sub-communities
	Tacit communities
	Appropriate channels
	Seclusion

External Factors

Institutions are often facing outside threats either from competitors, government, customers, or other institutions, and have to act consequently to deal with these threats. Different situations were described where participants acted to limit the disclosure of information to external parties.

These situations are regrouped under two sub-categories: legal considerations and reputation considerations.

Legal Considerations:

Different legal concerns were called out in the focus group discussions, they are: *General legal factors* that are not necessarily related to any specific law or contract - *Imposed legal factors* when the community has certain laws to which they have comply. - *Voluntary legal factors* that are contractual obligations which the community negotiates.

General legal factors

First, participants' behavior was conditioned by actual or potential consideration of legal consequences. As a publically traded company some information was embargoed by insider trading regulations. As one participant from the first focus group session explained:

"But we also receive a lot of information that is as a publicly traded company that would say: this is for internal use and it's flagged as such."[G1P1]

Another participant from the same session similarly noted that dissemination was limited for similar reasons:

"*If we receive any sort of preliminary finance data that hasn't been available, made available to the market yet." [G1P2]*

Beyond adhering to specific legal requirements, more general concerns for legal implications were expressed by participants from the same session. One response was:

"*if it is other sensitive information you may not want the recipient, even if it is the right recipient, passing that along, that is more for legal reasons or things like that. There is always that*

concern too as you are talking about information; this is a piece of information that could either be subpoenaed." [G1P1]

While another said:

"*I know there are certain types of conversations that I will only have over the phone because of the nature of the subject. That way if there ever a lawsuit or anything like that all of that information could be subpoenaed." [G1P4]*

Imposed legal factors

Participants described how some of their actions regarding outside parties were dictated by the legal framework they are bound by. A participant from the third focus group sessions gave several comments related to the legal framework they are bound by. The participant said her research group declines to give information on certain groups of students especially if these groups are relatively small in order to protect them and their identities:

"We do try to minimize the impact of survey on our students, particularly students in underrepresented groups. We are often asked to over sample [the data]. If you are a Native American and there are 15 surveys that are coming through there is a good chance you are going to fall in the sample group a lot more than someone who is not a Native American." [G2P2]

The participant also said:

"And that's not safe if somebody figure you out. There is one Native American in this particular major, this student record or data field that we created belongs to that student." [G2P2]

They also try to protect the faculty members by not imposing too much on their time.

"Our faculty members [are doing too many things], we also have a certain role to play there in just protecting people being hit with three or four surveys a day." [G2P2]

To make sure that new and existing employees follow the legal framework, in this case the FERPA (Family Educational Rights and Privacy Act) "which is a Federal law that protects the privacy of student education records" [54], the same participant stated:

"We ask them to read the FERPA information throughout the registrar website which is somewhat [complete] and covers most what we do. We are also asking the registrar to come and speak to us about the new interpretations and new ounces of FERPA. At the staff meeting, we just hired a new person that's always a good time to bring her back as FERPA officer to not only inform new people but to remind those of us who've been around for a while what we should be doing."[G2P2]

Participants of the third focus groups did not explicitly mention any legal framework, but they talked about customer contracts and the legal concerns associated to them. One participant said:

"*So in that sense, one of the big things we try to do is try to protect our customers, [be] cause they provide us with information and data to help us developing software for them. So we try to be very careful in terms of how we use that information. Who we share that information with, we make sure that it's not posted openly and shared it with others." [G3P2]*

The unit also blocked the disclosure of information when the information would refer to a too small group. For example, when posting grade distribution on the school website, if the class had less than ten students she said:

"If it is class thaught with 9 or fewer students the data is blocked. If you have 10 or more students in the class we will show what the distribution of the grades were on the website. [G2P2]

Certain communities establish procedures to follow with legal implications when they receive external request. One participant described the procedure his former office followed in this way:

"*When I was in [that other office] that was our standard if it was an external request. Our standard procedure was: tell [people in the public relations], they will phone it to the right or department and assess it, [subpoenability], as well as the information being released." [G2P5]*

Voluntary legal factors

Because humans can be considered as ambulant sensitive information holders, companies sometimes ask employees to sign legal agreement on intellectual property for non-disclosure. One of the participants for the second focus group said:

"The other aspect is really people, people have knowledge, we are a software industry, so people are secrets because they know everything. So those are broadly I think what we would be concerned about. That's why we have employee agreements and other things as well that they cannot share certain things." [G3P4]

Another participant said:

"[…] it's really a risk out there that the information may seep out. An employee may leave the company and take that information and spread it. So what we're looking at here is to have the necessary discipline, the necessary precaution if somebody does that; legal precaution." [G3P1]

While describing his job title, another participant from the third focus group session expressed his concerns about the non-disclosure agreements he signs on behalf of the company hoping that other employees will comply with them.

I mainly deal with customers and vendors directly. What I do for the company is that I do all the legal work. So a lot of non-disclosure provisions that I engage with the customers, with vendors and suppliers. The tricky part, there's obviously that I'm signing that [The company] as a whole group keeps confidential what we do, which I put my signature on something that I can only hope that everybody complies with." [G3P1]

These responses indicated an awareness of the explicit or implicit need to limit the kinds of information that should be disclosed outside of the organizations.

Reputation Considerations:

The second major category of external factors related to privacy awareness is reputation considerations. Participants described self-imposed limitations on information dissemination due to concerns for the standing or reputation of the organization. The reputation concerns expressed relate to preserving a sense of competence, trust, or credibility with either specific external entities or a more generalized external public.

One participant from the first focus group described the concern for ensuring that disclosure be deferred until accuracy could be ensured:

"*we had that a couple years ago when some rather large entity was affected and to make sure that the wrong information didn't get to that customer through social media, just employee being fanatical and helping customers; trying to be transparent that's what we try to do but be accurate. So let's strive for accuracy and not release of information too soon." [G1P5]*

The early release of incomplete or misleading information was deemed potentially harmful to the organization. One form of harm was identified by another participant from the same session as:

"*The end result is to prevent customer turnover. We want a tailored message going out to our customers cause they are going to take what we send then put it out on some other, like you said, form of social media." [G1P2]*

Another form of harm was expressed as a potential loss:

"for maintaining trust, If we change our story then customers will trust us less, like 'they don't know what they are doing'."[G1P3]

This concern for reputation was also expressed differently in one of the individual interviews this way:

"*There are some codes in there that's probably embarrassing. I think that's probably true, any software we've done a lot of work to clean up a lot of those things but there are still some stinky codes in there. I don't know that it would necessarily reflect badly on the company, I think that it's definitely something that I would be, there's something in there that I didn't write but I would be embarrassed for people to see my name attached to it." [G1P8]*

In order to maintain their credibility participants from the third focus group explained how they try to control the information they give out to other entities before the entities make it public:

"Undergraduate admission [raised] the question with the race, ethnicity, categorization came from federal government, [and] decided they were going to prioritize them in a different way. So we sort of like to help little bit to control over that information before it goes out so that they can't say this is what[our team]says." [G2P1]

Before they respond to any request regarding releasing information, one participant from the second focus group said:

"Sometimes we get to be very clear when we respond "yes" we can provide that information or No we can't provide that information, by going back to them and ask them what are going to do with this data? What type of analysis are you going to do? Is this just for publication or for your department, to improve operations within your department? Or is this something to improve operations within your department that you want to publish later? Just like the registrar office investigates with me what is the purpose of us requesting this data." [G2P2]

And they usually will go over a set of questions like to make sure they clearly understand the purpose of the request of this particular information.

*"There are sometimes you just don't feel that somebody is presenting the entire picture to you and being in the position we are in. we have to be a little bit cynical." [G2P2]*Because certain releases could affect the characterization of the university community as mentioned by one participant:

"It is sensitive in a way that it could affect, as I said before, the characterization for the institution. So it is not individually sensitive but it is sensitive to [the university]" [G2P1]

Sometimes the information that goes out is not sensitive per se, but it conveys an inappropriate message about the community. One participant from the third focus group shared an experience where he shared information with a customer about design ideas the development group was not committed to yet.

"I think one related example is a call when I was working on designs for a particular product and it's something for a particular customer then shared it with [the customer]. I wanted them to review the design and part of that design was something where, I think, the Project Manager came back and said 'well we didn't actually want to communicate to them that we were thinking

about this type of functionality or this type of thing.' [….] that was more like well we don't want to really tell them this because we haven't committed to it that sort of thing." [G3P2]

On the related note, another participant said:

"[…] when we actually expose something to the customer we want to do it in the context of making them short and understandable that this is something that is actually going to release or not. We don't like giving expectations that are not going to be met. So that's one of the things that we sort of deal with a lot when we are actually exposing our R&D effort to the customer to make sure that we are not making any promises or expectations that we are not entirely sure about."[G3P3]

Every company cares about maintaining their competitive advantage. One way of doing so is by protecting their intellectual property. For that reason, intellectual property becomes a type of sensitive information that companies are concerned about. One participant explicit said that they avoid disclosing this information in order "*to maintain competitive advantage in terms of developing software." [G3P2]*

Talking about how they prepare request for proposal, another participant said:

"we always have to worry about how much we're saying in those [RFP] because some of our RFP's go through complication. So we have to worry about protecting our own information as well besides customer's information. How much IP gets out there. So our own IP protection, which is a big thing for us."[G3P4]

Misinterpretation/Misrepresentation of Information

Although this category did not emerge in all of the focus group sessions, multiple cases about misrepresentation or misinterpretation of information were mentioned in the third focus group discussion. They care about misrepresentation/misinformation because it affects the image they project to the outside world. Some credibility issues expressed in the *reputation considerations* section are also related to misrepresentation of information. While discussing about releasing date to outside party, one participant mentioned:

"I know we have given data which has been misinterpreted or analyzed incorrectly." [G2P2]

Another participant gave a concrete a concrete example of misrepresentation:

"In [that other] office we did some analysis in a college; [when we went to] another meeting [what we saw] "was our chart but not our numbers." [G2P4]

The same participant along with another one explained the strategies they use to prevent misrepresentation/misinterpretation of information:

"I have had issues that's when pdf came about, that became a great tool where you can create an excel spreadsheet within an analysis and certain numbers and send it to someone. and there are people who will change it. So if you have a concern about that you can send them a pdf where it's hard they can't change it, they can't reinterpret it." [G2P4]

"We had something recently from [an office] where we were not sure where they were going with it and so we did the analysis before we sent it to them so that we would know whether we agree with what they would ultimately find. It serves a way of checking if there is misrepresentation of data on down the road." [G2P1]

Agreement on information releasing

Participants brought up several cases where they had some sort of internal discussions to understand what information is appropriate to disclose to outside parties. These internal discussions are not related to irregular situations but are rather considered as normal practices. Before signing agreement with outside parties, one participant said:

"I mean when you sign an agreement, we are to be legally bound by it. From that point we have, any information that goes out, to make sure that our proper people review and approve it." [G3P4]

In some cases, the discussion occurs to make sure they are following the organization procedure properly, as the following participant said:

"If somebody contacts me directly, I know for instance I am working with [another faculty member], and [the associate director] gave me a task in the past of helping her with ten-year track process, following group of individuals through 10-year track process. She sends me a note saying she likes some more information of these folks, maybe email address something like that. She may contact me directly, generally she will cc [the associate director] on it as well. [When] she sends me a request, again I go right to [the associate director]. I do [this] for 2 reasons: 1- to make sure it's alright to send this information. 2- to make sure my method matches his method." [G2P3]

On the same note of safety check another participant stated:

"[…] from my role I have literally daily question, can we do this? can we do that? So I'll look at the contract, I'll study the contract. I'll research it , saying in this case we can and know what

exactly you' re trying to release and make management decision and even sometimes leadership decision, what we can and what we cannot do." [G3P1]

With the hierarchical structure that exists in these communities, it is sometimes implicitly required to verify with people in the upper chain authority whether or not some information can be disclosed. On that matter, one participant stated:

"In my position anytime I get a request from anybody, a faculty member it has to go to through […] the associate director; he is who I work with primarily. Most of the people we report to, we already have established relationship [….], but I would get request from faculty members every now and then. If I get a request from any faculty member for any information it goes right to [the associate director]." [G2P3]

Another one said:

"Sometimes we are contacted by external organization and if it is the US department of education we comply, a lot of time if it is private organization we won't. And that they will ask for a list of contact information or faculty members so that they can survey them. And that type of thing we will send, we will ask human resources [….], do you think we should participate in this." [G2P2]

The same participant also said:

"One of the things, we almost always asked for permission to release, and we've released it every time but we just feel compelled to again ask 'are we still okay with doing this?'. There are a number of vendors out there who will file a freedom information act with us to get grades that professors are assigning for each course they teach so that they can market a product to prospective students showing the type of grades you can expect to get if you take professor x at institution y." [G2P2]

The tension that continually exists between what to disclose to a customer and what not to disclose because of competitive advantage issue, requires internal discussion before the release of such information.

"Sometimes they'd [customers] ask what's the maximum number of users do you support. We'd say that we support this technology, that technology. […] we have to be honest and at the same time we don't want competitors [to] use that information against us."[G3P4]

Internal Mechanisms

The second major category of privacy awareness concerned control over the flow of sensitive information within the organization itself. Five themes related to internal disclosures were identified in the interviews: (1) the formation of *explicit communities*, (2) the formation of *sub-communities*, (3) the formation of *tacit communities* in relation to the organization's reporting structure, (4) the choice of *appropriate channels* for communicating information of varying sensitivities, (5) and the use of *seclusion* for communicating without noise from the outside.

Explicit Communities

The formation of explicit internal communities was seen in how the tools for internal communication were conditioned to limit the disclosure of information. The use of group controls on an internal social networking tool was described as:

"…you can create a group and manage membership. So you can control to some degree that sees what within that tool."[G1P2]

Similarly, another participant noted that:

"There are two types of groups: groups that anybody can join and there are locked groups." [G1P1]

The evident purpose of these groups was to control the dissemination of information. Another motivation for the creation of groups was to facilitate focus. As one participant related:

"I would say generally even the private groups, it's not really meant for private conversation it's just meant for people that are part of one group to be able to have a conversation without noise from the outside, to kinda keep it more focused." [G1P3]

Explicit groups were also formed based on an individual's job responsibilities. One person noted how:

"*Groups are established too by your job title. In [other location] there is HR team and there are subsets of the HR team based on your job title. You don't get added to a certain group unless your title changes to fit the profile of that group." [G1P2]*

Another one said:

"We have the team system which is basically a source control slash work ticket line system and that's where everyone knows which work ticket [….] is assigned to them by their name, or they receive an email to it. […] And then we have email list. […] so someone knows that's the email list they belong to, so there exchanges that happens based on the work they are doing on the email list. And we also have a development share point [], so everyone knows exactly which [one] you can go to and what sort of the work going on within the team." [G3P3]

More cases of groups that were created based on job's responsibilities are described below. One participant *"Actually there is another usability intern, but us as a group together, we're on multiple different projects simultaneously so that's why we're kind of always talking with these different folks and different groups of people." [G3P2]*

"They usually are assigned and they are assigned different team, groups. For developers there are work tickets, they have all the tools, they get assigned for that so they know they have to work on this particular thing." [G3P4]

"when we are distributed and everything we encourage people [to] put a lot of stuff in email so that they can go back to it and all that. If a meeting like this happens and someone was thinking about a ticket that he was supposed to work on and misses 2 minutes of conversation, you know it's gone; right. But if you have the same thing recorded on email when he has to use it he can go back and say alright let me look at that information and at least it's there. So that's the reason why there's a lot of efficiency with email group. Everything we have emailed are cross functional team, we have email groups or when they are multi-sittings, we have email groups." [G3P4]

Being a member of certain explicit groups grants access to all information related to the group with no exception. Two participants mentioned:

"For us in [the research group] we are pretty much, once you are in [the research group] if you are a full time employee in [the research group] you pretty get access to everything that everybody else in IR has access too." [G2P2]

"And we try very hard not to be cryptic, if you have access to our server you going to know what's in the folder that says degree awarded, time to degree or graduation. We deliberately try not to be cryptic so that we can have an office environment where everybody can see what everybody else is doing and if you need access to somebody's program, you can." [G2P2]

"But if I know that [person] works On Time to Degree project and somebody asks me for something and that [person] is out on disability, I can do a string search on the server for time to

degree and that would bring up on of her programs, that's in terms of seeing what other people, we have access to pretty much all of that." [G2P1]

Some groups are formed on the fly to satisfy the organization's need as one participant stated it. Members of these groups can come from different other groups of the organization and form *"a lot of overlapping groups." [G3P4]*

"I think there's a lot of function in our company that are unique, [.....].You see a lot of times when I get ready to do a contract, obviously I have a salesperson, I have a support person, I have client services person but all need to be brought into this whole request for proposal so there's a team formed by function of process, a function of who is the best available person for that. [....]To help find what the customer's needs are in order to make proposal; if that makes sense. So the teams are all dynamic. So it's not necessarily a group that communicates by themselves. But it's a cross functional group that did a lot of what is going on." [G3P1]

In certain situations groups are not indefinitely created. The following example is about lifetime of ad hoc group created to work on a specific project.

"Something in RFP could be a couple weeks or months but it could be an implementation, strategic implementation going on. So we form group cross disciplines" [G3P4]

Sub-Communities

Communities are often divided into sub-communities when members of such communities are working on different projects or different part of a same project. Communications that are relevant to some members in the group are not necessarily relevant to all members. To avoid any communication overload or distractions, they create sub-communities as one participant mentioned it in the following quote: *"If there are multiple persons directing a project they could*

choose to create, let's say there are five of those persons, it's rare but there could be an alias just for those developers, they are more free to share that information as developers. In 80%, 90% of the cases privacy is not involved it's a matter of if that information is relevant to that person or not." [G3P4]

A participant explained how an HR group is sub-divided based on titles"

"I am on a list in [that other city] it's only people with the same title, [this other person] is on a list in [that other city] with the same title that I am not on." [G1P2]

These two examples show that sub-communities exist to facilitate the flow of communication as well as keeping it more focus.

Tacit Communities

The example below suggests a relationship between the channels of communication and the structure of the organization, limiting the flows of certain information to paths defined by the managerial structure. One participant described the confidentiality of information this way:

"*I feel like confidentiality goes up but not always down. I would say there is very little information that I wouldn't talk about with my manager but there are a lot of things I would not talk about with people working with me." [G1P3]*

Similar views were expressed by participants who related this situation:

"The person I manage, he deals with two different contracts that otherwise I would have no information about. But as his manger it is my job to advise him and to help him. So I [have access] to the information that he has so that I can mentor him appropriately. It's information that in any other context I wouldn't need to have information about but I do because I need to help him." [G1P1]

"Because she is my manager and she's been in this role a whole lot longer, I don't feel like there any information that I wouldn't or couldn't share or ask."[G1P4]

"I think it's okay to share some pieces of information because it's my job to help and provide some discretion on what's shared outside of HR. [G1P2]

A related example concerned salary information:

"*There are privacy issues as well, we don't want other people to know about salaries and other incident we have to go to HR for. That's something between a manager and an employee. That should not be widened*." *[G1P7]*

This next example is related to someone who doesn't belong to any specific group, yet they interact with all existing and established groups. He described his case as followed:

"As an organization, I'm under development and I'm a little unique in that when I first started with [the company] they didn't have a separate user experience or usability person. So I'm kind of anomaly, I guess, I report to development, But I work with all these other groups to be able to do my job effectively. So I'm in regular contact with all these different groups, especially, I try to be in contact as much as possible with the product managers who deal most directly with customers in terms of defining requirements. I try to work with them to get contact with customers, end users so I can gather data from them and also sometimes recruit them to do user tests; that sort of things." [G3P2]

In all of these cases, implicit communities are formed around the organization's reporting and managerial lines. That is, the authority relationships created a tacit community within which sensitive information could be shared. Interestingly, we will see cases later where, for certain information (about employee health conditions), one tacit community is replaced by a separate

overriding tacit community. We will also see tacit communities formed around peer relationships rather than authority relationships.

Appropriate Channels

Privacy-awareness was also evident in how the perceived sensitivity of information influenced the choice of communication channels used to convey the information. Examples of factors that influence the choice of communication channels are: trust, information leakage, and control. We have already seen one example of this above (concern for creating a record subject to legal subpoena). One participant described a different situation this way:

"For example if we are talking to a manager about a hiring decision for a candidate, we normally don't put reasons we are or are not hiring someone in an email, we will have that verbally." [G1P2]

Differences among electronic communication mechanisms were perceptible to the participants. One individual stated that:

"*I think if I had a question I would not use [internal instant messaging system] to communicate it. It is more fun and informal. I would never put something confidential on there." [G1P2]*

While another similarly observed:

"We trust email more than we trust [the internal instant messaging system]." [G1P5]

The sense of appropriateness was also reflected in channels of communication outside of the company's control. One individual explained that:

"Who knows what [an outside providers] privacy policy in terms of use that the different email providers might have. So they will have access to information that we never intended them to have access to." [G1P3]

Similarly, even useful channels of communication have limited utility because of privacy concerns. In this case, the same individual related how:

"*Multivideo chat is great example where, there is no commercial solution; [what] we have that is good enough for to have 10 people videochatting, but google hangout does this perfectly and it's free. [But] … I wouldn't say very much confidential information, probably private information." [G1P3]*

In both of these cases, the lack of privacy guarantees limited the use of these communication tools. Another illustration of privacy-awareness in choosing a communication channel concerns the use of physical space and its counterpart in videoconferencing. One participant in the human resources group told how:

"*If we are having a performance discussion or we are ending someone's employment, I am not going to go to a room that has a bunch of windows." [G1P2]*

The concern in this case was that the reaction of the individual being addressed in the meeting could be visible to others, thus leaking information about that individual's employment or performance status. The human resource office itself was especially viewed as a setting for limiting information disclosure. This was succinctly captured in the words of a participant:

"In HR we talk about the rectangle of trust … what happened there it stays in that room. It doesn't leave that room." [G1P4]

This use of physical barriers to condition information dissemination also has its counterpart in virtual communications. When describing the use of videoconferencing, the ability to ensure the same physical limitations used in face-to-face meetings against unintended disclosure was described this way:

"… if I have to share something I think is an inflammatory HR thing I want to be able to see people I am talking to. Not be talking to that person on the phone and potentially be on speakerphone with other people listening. I don't want to communicate with someone when they are just sitting on their desk, I want them to be in a conference with the doors close where I can see them." [G1P2]

It is intriguing that the videoconferencing channel is deemed appropriate because it allows for the emulation of the precautions taken in the physical world. A similar example was reported by another participant where physical limitations were imposed:

"When it gets to a point where actual sensitive information would have to be shared in order to determine whether we can or not we usually have a meeting such as in a conference room where it's private with a overhead to show data, to show what we have, what we can. It's usually not shared through email, the actual data." [G3P1]

For distributed groups, conference call is used as means of communication. One participant described the safeguard put in place to prevent disclosure of information:

"Our leadership team is distributed so we do use conference call and stuff on network and these are private network. And we try to have passwords for conference call on such sensitive topics." [G2P4]

The means of communication depend on the sensitivity of the information; one participant listed the conditions where his team would use IM's:

"some companies don't allow public IM's to be used they give just a local IM to share those things, very sensitive information. In our cases, it's more "hey what this variable do", small communication happening between developers and Q/A. Even if it falls into some competitor's hands, we know that it's very difficult for them to figure out what's going on. But that's a call,

depending on the nature of the project, what kinds of agreement you have signed. We may choose not to use IM or email. Because it's so easy to breach it out." [G3P4]

While the proper means are used to transfer the data, there can be some concerns about security of the data transferred. One participant expressed the concerns in this way:

"And because we are creating text file sometimes to send to [State Oversight Board] SCHEV in an encrypted environment, we find that we can't do that. So you temporarily have to write a text file to a less secure non encrypted area it is still on the server, it's not encrypted anymore in order to submit that through secure FTP to SCHEV. But for that brief period of time, it may only be 10 minutes or so it's not encrypted. And for people like me, [you just think that somebody will try to hack in]." [G2P2]

Seclusion

During the focus group discussions, there were cases where privacy was taken in the sense of withdrawing from the whole group to have some space apart. One participant expressed it this way:

"I would say generally even the private groups, it's not really meant for private conversation it's just meant for people that are part of one group to be able to have a conversation without noise from the outside, to kinda keep it more focus. I can't think of any rare cases where specific confidential information is discussed within these groups. We know these people, it is only this group that is really involved in this, so only this group should be having the need to be having this conversation." [G1P3]

In summary, as shown in Table 1 above, the participants described privacy-awareness related to disclosure both outside of the organization (external disclosures) and also within the organization

(internal disclosures). External disclosures were affected by legal and reputational factors. Privacy factors related to internal disclosures were evidenced in the creation of explicit groups for containing information, the recognition of tacit groups organized around the company's lines of authority, and the choice of appropriate channels of communication.

Situated Disclosures

The second principal theme to emerge from the analysis of the focus group interview concerns how exceptional situations influenced the way in which individuals dealt with sensitive information. We use the term situated disclosures because it relates to the view of Suchman on situated action. Situated action refers to the ability "to adapt to circumstances, deal with contingencies, and act at the right time in seizing favorable opportunities [7]." In several instances participants mentioned situations where they had to act outside of the norms for various reasons. In this section we present both external and internal factors that require situated disclosures.

Table 2 summarizes the factors related to situated disclosures that were found in our analysis and are discussed in this section.

External Factors

Organization Needs:

Some cases were described that involved unusual circumstances where information was disclosed to outside parties. One case involved the use of an external consultant:

"One interesting example is, there is a system that requires a lot of domain knowledge that very few people here might have so we use consultants to help us when the systems are at faults and in one case there was an urgent need to use someone, we needed to give them data, customer data for them to help fix it. So all the customer data was scrubbed, all identifying information was taking out and replaced with some identifiers that the consultant could not use and then we

got the information back, we would take those identifiers and replace them with customers' data. So we scrub the data before transiting in, we securely ship them on an encrypted hard drive." [G1P3]

In this situation the tension between protecting the sensitive information (the customer data) and the need for external expertise was resolved by two actions. First, the data was sanitized before release (and re-specialized afterwards). Second, security measures were used to ensure the confidentiality of the information in transit.

Table 2

Situated Disclosures Disclosures in exceptional circumstances	
External (factors)	Organization needs
	Internal agreement needed
Internal (mechanisms)	Privacy vs. organizational needs
	Degrees of privacy
	Tacit communities outside reporting lines

Internal Agreement needed:

Cases involving the disclosure of sensitive information to outside parties were described as follows:

"We also get request from time to time asking for release of information, release of data to perhaps to another institution of the state or even to a vendor of some sort who is trying to create a database of faculty members or something like that. In those cases when it is specifically refers to faculty data […] we ask the [person in charge] for faculty affair […] what do you think? Should we share this information?" [G2P2]

" I have request the other day for some birthdates, somebody was putting together who has students enrolled in a class which is going to study abroad in order to purchase overseas airline tickets for everybody, for overseas travel now you have to supply TSA with birthdates. Is that a FERPA issue, can we release that information? Can the professor designing the course release that information without getting sign off from the students? That's a question that never comes up before and immediately I ask the registrar about that. She gave me the answer No, we need to get the students permission." [G2P2]

In all those cases, discussion and agreement within the organization was needed to form a judgment of what information to release and what information to withhold. The salient features of these last examples are: (1) privacy management is highly context-dependent and involves reacting to unusual situations and circumstances, (2) interaction and agreement among internal parties is part of the process that leads to disclosure to outside parties.

Internal Mechanisms:

Privacy versus organizational needs

Other cases involved only people within the organization itself. One such situation was described by a participant who operated outside of the hierarchical authority structure and against the "need to know" guideline. The participant related that:

"Just today I was asking some [one for] advice about salary information about some we are going to hire. That's exactly something where he doesn't need to know but I just want his professional advice on that information. We both understand the sensitivity of that information. There is a trust relationship that is really making that okay." [G1P7]

In this situation, the behavior indicated that a concern for a proper salary negotiation superseded the normal restrictions on knowledge about salary information. Another situation also appeared to involve resolving the tension between conflicting privacy goals. The participant told how:

"*Normally I would not share with our IT team when someone left why they left, they just get a ticket and our system says at midnight tonight turn this off. But there is at least one occasion where I had to go to their office to say hey I need you to just stop what you are doing, turn off the system access now. And that is in some way sharing the information that this person didn't leave on their own, but you have to do that to protect the customer data." [G1P2]*

The same participant also said:

"When someone leaves the company, we can, we don't always do that, the manager would say whether they want this or not. We would forward that person's emails to their

manager. So if there are open items, anything that was in their inbox goes to their manager's inbox, so their manager can help close out open items and we turn off the mail box at a certain point. But that wouldn't of course happen in any other way." [G1P2]

In these cases, some degree of privacy related to the terminated employee was judged to be less important than the need to protect the privacy of the customer data. Another participant mentioned the need to sometimes share information within different groups for business reason:

"We encourage them to share if it is relevant to share. In most cases it's not a privacy issue; it's more knowledge issue, dependency issue, technical issue. Something that's happening here could happen here as well, so by the way please, this is the thing." [G3P4]

In other cases the welfare of a member of the organization may call for the violation of privacy. After describing how medical information of an employee is held in confidence, even from the employee's manager, a manager said that:

"*If there were a medical emergency happening and you know the conditions. There might be a time you have to say something to protect their health."[G1P1]* Accessing the right information to accomplish a certain task can, sometimes, be more critical than respecting the restrictions imposed on the information. One participant explained how their team do not usually access financial aids data but can be required to do so to execute certain tasks.

"We have some financial aids data. Usually not because we go out and acquire it ourselves from the database but because of something we do, we need data from the financial aid office which we store on our server." [G2P1]

These situations all involved circumstances, people, and information internal to the organization. Interestingly, in these cases we see (1) information sharing that creates a tacit community outside of the authority structure, (2) an implicit recognition of a hierarchy of privacy concerns, and (3) tension between privacy and other community goals.

Confinement of Sensitive Information

A third major theme to emerge from the focus group analysis concerned how the members of the various communities in the organization sought to control the dissemination of information. It has already been observed that technical mechanisms (group controls and listservs in Section privacy awareness) were used to form sub-communities for this purpose. In addition, as summarized in Table 3, we found (1) informal methods being used to signal the degree of sensitivity and condition the subsequent handling of information, (2) a number of breakdowns of the technical and informal methods, and (3) that time affects the sensitivity of some information.

Informal Methods

Signaling:

Signaling in this case refers to the use of certain means of delivering sensitive information to communicate the degree of sensitivity of the information. This may

involve use of metadata or procedural steps. In a number of different contexts, individuals created informal means to signal the sensitivity of information. One participant described this procedure as:

"*I created an encrypted PDF, sent them the PDF, and then call them to give the passphrase to open [decrypt the document], and the passphrase was something like "I read for my eyes only". So as they type [the passphrase] it was obvious that they are not supposed to share it. If they did share it, it's on them not on me." [G1P5]*

For important meetings with distributed team, the use of password on conference calls indicates that the topics of the meetings are sensitive. One participant described the use of password in this way:

"…our leadership team is distributed so we do use conference call and stuff on network and these are private network. And we try to have passwords for conference call on such sensitive topics." [G3P4]

In the first case, the passphrase both conveyed the high level of sensitivity of the information but also conveyed the sender's intent that this information was not to be further shared. Alternatively, this can be viewed as a means for transferring the responsibility for properly handling the sensitive information from the sender to the receiver.

Table 3

Confinement of Sensitive Information Means to control dissemination	
Informal methods	Signaling
	Appropriation
	Feedback from peers
Breakdowns	Misclosures • misidentification • inappropriate channels/roles • misaggregation
	Loss of Community
	Mismatch of privacy expectations
Time	Loss of sensitivity over time

Appropriation:

Another type of informal method described by different participants was appropriation one example was described as follow

"…sometime we'll use a prefix in the subject if it says confidential that's a signal that means do not forward this message, it's not appropriate, it's only for your eyes. We use

words like "urgent", "confidential", "do not forward" things that tell the person reading the message hey this is only for you." [G1P2]

In other words, the email subject line was appropriated to convey metadata related to the privacy of the email and to hint at how the information might be further shared. This same participant elaborated on other similar email practices:

"..you can tell this by the body of the message or the recipients list; so we have a group people in Blacksburg, people that we consider our Blacksburg leadership team. If it is our standard leadership recipient list and in the email it says this is a leadership team topic we're going to talk about it, then that basically signals it stays within that group, it should not be sent out to anybody else so that you can tell who the message is intended for by the recipients list and sometimes just by what's in the body of the message." [G1P2]

Here the privacy metadata is embedded implicitly in the understanding of the recipient list or explicitly in the body of the email. Another participant also described this practice by saying:

"Keep this within this group" that is a phrase I hear a lot." [G1P7]

Breakdowns

We next examine breakdowns, situations where the normal practices, such as those just described, fail to achieve the desired goals. The breakdowns reported by the study participants fell into three classes: (1) misclosures -- mistaken disclosures [12], (2) loss of community over time, and (3) mismatching privacy expectations. These breakdowns result in actual or potential loss of privacy by allowing disclosures that were unintended.

Misclosures

Misclosures, unintended dissemination of information, occur for a variety of reasons.

- *Misidentification*

One cause is misidentification, where simple mistakes cause unintended recipients to be confused with or included among the intended recipients. One participant noted misclosures occurring because two individuals in the same group have the same name:

"One of the guys I work close with his name is [the same as mine]. So every now and then someone will ask me to do something that's not really my role but his role but they don't really understand the difference between the two of us." [G1P3]

This same participant reported other forms of simple mistakes:

"*But I've seen in the past accidental typos in To lines meaning emails get sent to the wrong people or to larger group than intended." [G1P3]*

In other cases the cognitive mistakes underlying the misclosure are more subtle. This person said:

"When talking to a possible candidate. We copied someone who we were talking about inviting back to be rehired, that person's name should have been in the content of the email but not in the To [field]." [G1P2]

In this case the sender made the cognitive error of failing to distinguish between the use of an identity in the body of the email versus its use in the recipients list.

Misspelling error can also cause misclosure. In the example below someone described the case where unintended information was sent to someone due to misspelled email address:

"I have an example (not in this office). When I was in the [other office] office we doubled our faculty salary and raises in a way that we'd send information to departments about because the deans with provost and finance people and they 've always made decisions about who get what sort of thing. This is when we first started going to the paperless and the person [who was in charge said] we will send those [instead of the notebook] to the different colleges and departments, we will do it electronically. And one of those names in the email was spelled incorrectly, and all this information about the discussion about who get what and why went to another person and luckily they were able to get the IT folks to grab that email before it went to the wrong person." [G2P4]

- *Inappropriate Channel/ role*

A second cause of misclosures is when the correct identity is used but the communication uses an inappropriate channel or addresses the recipient in the wrong role. The use of an inappropriate channel was described by one person this way:

"*Within our team the thing that keeps happening is email will go to a different source. Like someone have his Gmail and his company email and if they get into the thread, emails will somehow start going to his Gmail account. … So you are taking email out of the context of the company email system into another email system. So there is great privacy concern there." [G1P1]*

In this situation an otherwise proper communication spills over into channels outside of the company control, raising concerns about how this information might be treated. A case where the correct person was addressed in the wrong role was related by another participant as follows:

"*I actually reached out to somebody not with the company but still had a [company] account. I thought he was still working here. I started talking to him about an operational issue that was going on. Then he said, you should talk to [someone else] about this because I am not here anymore." [G1P7]*

The misclosure in this case occurred because of lack of awareness of the status of the individual being addressed.

- *Misaggregation*

A third type of misclosure arises when information with higher and lower degrees of sensitivity is disclosed as if it contained only a lower level of sensitivity. Two examples of this were cited by two participants from two focus group sessions:

"Someone sends a spreadsheet to the company, one of the rows which you can't even see on the main view but if you scroll over you will see it. It contains private information about individuals. That's an example of a scenario where something went out accidentally to people that has something private but shouldn't been." [G1P3]

"The person in [that other office] office who sends a spreadsheet and underneath of the spreadsheet there were social security numbers. They posted it on the web; our office was not involved at all. That sort of things we are extremely sensitive about and careful with." [G2P1]

The misclosure in these incident resulted from lack of awareness of the full information being disclosed. Thus, the sender inadvertently exposed more sensitive information than intended.

In summary, various forms of misclosures were reported. The causes of these misclosures resulted from the confusion of identities, unintended leaks to people in the wrong status

or using less trustworthy channels, and failure to realize the confluence of information having different levels of sensitivity.

Information with different degrees of sensitivity is sometimes intertwined and there is the need for a cleansing effort before it goes out to a customer.

"[…] we have work ticket systems and customers want to know the issues. And a lot of work ticket systems are retained within aspect it's an internal communication so there a lot of written on email, there's could be privacy information, discussion and if you accidentally expose that work ticket; customers also want to look at things and saying" I want to know what are the issues with this product" and when you are releasing it. So there could be things that could be exposed, you may need to go through a cleansing effort of all the work ticket in your system before you expose it to the customer." [G3P4]

Loss of Community:

The second major form of breakdowns involved the loss of community. That is, it became difficult to know over time the true extent of the community. A member of the company noted that:

"*Email chains will start to build up, people add three people they think might need to be part of the conversation, but in fact they don't need to be. And then those three people are on there for the rest of the conversation even though we determined ten emails back they don't need to part of this anymore. I think another thing that can happen is, you send somebody an email and they get forwarded, it just grows and grows." [G1P7]*

A participant in another group reflected on this type of breakdown:

"*I can't recall specific instance but I can imagine how that can happen. Like we might have something where an email goes out to a group of people, internal people and at some point someone external gets cc and may be you don't realize."[G3P2]*

The circumstance described here arises because the sender of the email has no effective control on the growth of the community, those who become aware of the information contained in the email. The effect of forwarding the email and including other recipients serves to expand the community without control.

A related phenomenon occurs when the definition of the community remains the same but the status of the individuals in the community changes. An example of this was:

"*Old [company] mailist … [where] … sensitive conversation you see going to a subject @list.[company].com is really going outside of the company and sometimes to employees that are no longer employees." [G1P5]*

In this scenario the actual membership of the community is not properly reflected by the mechanism used to name the community.

Mismatch of Privacy Expectation:

The third major form of breakdowns involved communication between people with different privacy expectations. A situation of this form came about in this way:

"For example if I say hey this person is not pulling their way on this project I know some people who have no problems forwarding that email directly to the person who is not pulling their way because they are not trying to avoid conflicts, they don't care about conflict. They want the words to get out there. I am more of a person who likes to avoid conflict and would approach it in a different way. So this is an example of how people

view the same email in very different ways about who the intended audience could be." [G1P3]

A similar situation

"*well, all that happens with email, that's the danger of email. When you are, like P1 said he's communicating to this person the context is you're talking to this person and later to this and you make the assumption that the person knows this and this. And this other person makes a judgment call that he forwards to somebody else but P1 didn't want this other person to know." [G3P4]*

"*I mean it's not that you don't want that person to know. If you were talking to that other person, you would give them a lot more background and then talk, right. Because you know that person would be upset or is from others group or something like that. It could be internal, political stuff or something like that." [G3P4]*

In effect, differences in the privacy values of the two parties led to differences in their expectations of the appropriate community. This breakdown is possible because the privacy expectations are not articulated, but remain below the surface.

In summary, three general forms of breakdowns were seen. These breakdowns related to misclosure, loss of community, and mismatching privacy expectations. These breakdowns point to areas where systems can provide better support in an effort to minimize or eliminate the causes of the breakdowns.

Time Related

Loss of sensitivity Over Time

The third element related to the theme of confinement of sensitive information concerns how the sensitivity of information might change over time. Several members of one the focus group gave example where information would be sensitive only for some period of time. One person indicated that:

"*I think every now and then too I will say "keep this between me and you until x" [be]cause a lot of time sensitive information is sensitive for a certain amount of time." [G1P3]*

A second noted that:

"The only other phrase I've seen is "this isn't public yet". Just something is going to be announced soon and they want a few people to know ahead of time." [G1P7]

The same individual said:

"It's not just financial information it can be a new product launch that's does happen through email; that is sent to everybody and that says "don't talk about this yet, but this is coming." [G1P7]

The significance of these examples is that mechanisms to help members of a community deal with sensitive information must have some means to allow the sensitivity of some information to expire after a period of time.

Feedback from peers

Feedback from peers which can relate to the notification element in our privacy model was called out by one participant in each focus group session. This sends a warning to members of the community in cases where misclosures are about to occur.

" Aside from the [upper] chain of the file through which this information will go, I don't think it would get passed, we never release it on our own, I don't think it would get passed the second or third person that [wouldn't] be asked about it." [G2P1]

"I think the PM came back and said well we didn't actually want to communicate to them that we were thinking about this type of functionality or this type of thing." [G3P2]

"I actually reached out to somebody not with the company but still had a [company] account. I thought he was still working here. I started talking to him about an operational issue that was going on. Then he said, you should talk to [someone else] about this because I am not here anymore." [G1P7]

In all three cases mentioned above, the notification was sent after the fact which could heavily affect the community. In our privacy model is designed to send notification before a misclosure happens.

CHAPTER 5

DISCUSSION

The focus group study was designed to help us understand how communities deal with sensitive information for which they have responsibility. We undertook the study with some notions of what elements might be part of a community privacy model such as: community tags, exception mechanisms and notification. The discussion section is organized as follow. First, we reflect on the findings from the analysis section. Second, we relate the analysis of the focus group sessions to our privacy model elements. Third, we relate our work to individual privacy based on the findings and confirm that community privacy does exist.

Reflections on Findings

To reflect on the findings, we create three tables in which we map the comments and statements made by the participants of all three focus group sessions to the different themes that emerge from the analysis. This mapping visually shows how strongly the different themes were supported by the participants. Table 4 shows the number of comments on factors related to privacy awareness. Participants across all three focus group sessions made a total of 37 comments about external factors and a total of 36 comments on internal factors (mechanisms) related to privacy awareness. These results show clearly that privacy awareness is strongly supported by these communities. There was a total of sixteen participants in the three focus group sessions. There is at least one comment from each of them in this category.

Table 4

Privacy Awareness Comments from the Focus Groups				
		G1	**G2**	**G3**
External (factors) 37 comments	Legal considerations	p1, p2, p1, p4	p2, p2, p2, p2, p5, p2	p2, p4, p1, p1
	Reputation considerations	p5, p2, p3, p8	p1, p2, p2, p1	p2, p3
	Misrepresentation/misinterpretation		p2, p4, p4, p1	
	Agreement on information to be released		p3, p4, p3, p2, p2	p4,p1
	Competitive advantage			p2, p4
Internal (mechanism) 36 comments	Explicit communities	p2, p2, p3, p2	p2, p2, p1	p3, p2, p4, p4, p4, p1, p4
	Sub - communities	p2		p4
	Tacit communities	p3, p1, p4, p2, p7		p2
	Appropriate channel	p2, p2, p5, p3, p3, p2, p4, p2	p4, p2	p1, p4
	Seclusion	p3		

This great support from the participants was somewhat expected because of the nature of the communities we interviewed. It would be interesting to see if this category would be

so strongly supported if the communities were ad hoc groups or less structured groups. While this category is well supported overall by the participants, the sub-categories show various level of support. Participants from all three focus group sessions acknowledge *Legal considerations and reputation considerations* as factors that influence their behavior for the dissemination of information to external parties. The *Misrepresentation/Misinterpretation* factor was only mentioned by participants of one focus group. However, we can easily believe that this is still an important concern for communities dealing with a big dataset. The *Agreement on information to be released* factor was mentioned by every participant from the second focus group and one participant from the third focus group, while nobody from the first focus group raised this as a concern. Only two participants from the third focus group considered *competitive advantage* a factor related to privacy awareness.

Fourteen out of the thirty six comments in the internal factors related to privacy awareness were about explicit communities. These comments were distributed across all three focus groups. Participants strongly supported the existence of explicit internal communities that help control the dissemination of information. It is also shown in the table that privacy awareness dictates the channels used to transfer sensitive information. Twelve comments about *appropriate channels* were made by participants in all three focus groups discussions. They recognize that choosing appropriate channels of communication can reduce the risk of information disclosure. The factor *Tacit communities* is not as well supported as the other two sub-categories in the internal factor (mechanism) category, but has a good number of comments from the first organization.

Two sub-categories, *sub-communities* and *seclusion*, did not receive many comments from the participants.

Table 5

Situated Disclosures Comments from the Focus Groups				
		G1	**G2**	**G3**
External (factors) 3 comments	Organization needs	p1		
	Internal agreement needed		p2, p2	
Internal (mechanism) 6 comments	Privacy vs. organizational needs	p7, p2, p2, p1	p1	p4

Table 5 shows the number of comments made by participants related to actions taken to either disclose or withhold information in exceptional circumstances. Unlike the first category which has a total of 73 comments, this category only has 9 comments. Participants in all three focus group sessions commented on situations where the community has to decide between privacy and organizational needs before disclosing certain sensitive information. Even though there are a small number of comments for this category, it is still important to understand that privacy management is highly context dependent and take that into consideration in early stage of the design of community privacy systems.

Table 6

Confinement of Sensitive Information Comments from the Focus Groups				
		G1	**G2**	**G3**
Informal methods 8 comments	Signaling	p5		p4
	Appropriation	p2, p2, p7		
	Feedback from peers	p7	p1	p2
Breakdowns 15 comments	Misclosures	p3, p3, p2, p1, p7, p3	p4, p1	p4
	Loss of community	p7, p5		p2
	Mismatch of privacy expectations	p3		p4, p4
Time 3 comments	Loss of sensitivity over time	p3, p7, p7		

Table 6 shows the number of comments made on *informal methods* used to signal the degree of sensitivity, *breakdowns* of the technical and informal methods, and how *time* affects the sensitivity of some information. The *breakdown* sub-category was the most supported out of all three sub-categories. Fifteen comments on breakdowns were reported by participants of all three focus group sessions, with 9 comments on misclosures alone. This shows that breakdowns constitute a significant issue in these communities where privacy is very important. Having the means to prevent such breakdowns as we propose in our privacy model could be very useful to the communities.

Relationship of findings to our model

The focus group analysis is related to our model elements in the following ways: First, the focus group discussions demonstrated a keen sense of privacy awareness. This awareness was reflected in the existence of various communities and sub-communities dealing with a variety of sensitive information. In some cases the community was the entire organization itself. This occurred when the organization was subject to legal requirements but also in cases where there existed a distinction between the organization and others outside of the organization. In other cases tacit communities were grounded in the organizational structure, peers with the same responsibilities or non-peers related by a supervisory relationship. The privacy-awareness also influenced what communication channels were used. The choice of channels was done to insure the integrity of the information being conveyed within the community. The term "channel" in this sense meant not only a technical means (e.g., email) but also included physical elements (e.g., a windowless room). Privacy awareness also encouraged members of the communities or sub-communities to sometimes use seclusion in order to have a more uninterrupted conversation. These examples give us reason to focus on the community as an important element in privacy protection. The use of *community tags* in our privacy model is aligned with these findings because the tag is a visible structure that allows the community to define itself and articulate its privacy protections. We also note later in this section how community tags can prevent or minimize privacy breakdowns (i.e., misclosures).

Second, the analysis under the theme of situated disclosures showed that context plays an important role in privacy management. This is not surprising but it emphasizes how limited and brittle policy-based mechanisms are likely to be. Thus, mechanisms that

allow the community to deal fluidly and flexibly with exceptional or unusual circumstances are important features of a community privacy model. In many cases, the reported situations involved some degree of discussion or agreement among the members of the community on the form or content of the information ultimately disclosed. The exception mechanism in our community privacy model is a start at addressing this requirement. However, we did not anticipate in our model that notion of time sensitive information (i.e., information that is sensitive only for a specific period). This aspect can be easily accommodated.

Third, the efforts to contain sensitive information showed how existing structures (subject lines, listserv group names) were appropriated to convey metadata about the sensitivity of information and/or suggest appropriate ways in which this information might be disclosed. The use of community tags is a more principled way to provide this metadata. Because a tag can be attached to any information or device, it is possible to create a uniform and ubiquitous way of defining privacy boundaries.

Fourth, a number of breakdowns were seen in how the dissemination of sensitive information was controlled. These "misclosures" can, in many cases, be eliminated by our privacy model. For example, the loss of community cannot occur because the change in community membership must be agreed to by the members (or by the appropriate authority), the differences in privacy expectations can be avoided because the community tag conveys explicitly the sender's notion of the appropriate community with which to share this information. In addition, some of the forms of misidentification can also be eliminated by more explicit use of community tags.

Fifth, we saw little evidence of notifications as an element in community privacy. It is not apparent whether this is due to notifications being unnecessary or simply that such feedback is not part of the current toolset. Therefore, we believe that notification should continue to be studied because the lack of feedback via notifications leaves the community ill-informed about the attempted misbehaviors of its members.

Relationship of findings to individual privacy

Beyond the relationship to our privacy model, it was also interesting to observe certain parallels between individual privacy and community privacy. We note two of these parallels here.

First, legal scholar Daniel Solove notes [51] that the privacy harms articulated by Justices Warren and Brandeis [58] are "dignitary harms" that, like harms recognized in defamation law, "lowered people in the esteem of others." We found it interesting that one of the participants said that some poorly written code should not be disseminated outside the company because it would reflect badly on the quality of the software produced by the organization, and him in particular (See [G1P8] comment in the Reputation Considerations Section). Certainly an organization's reputation is reflected in trademark protection and branding. But through the lens of privacy we can also view privacy violations as bringing about dignitary harms, in this case to the community rather than to the individual.

Second, Solove also observes that an individual's right to solitude as a privacy right is recognized because "a space apart from others has enabled people to develop artistic, political, and religious ideas that have had lasting influence and value when later

introduced into the public sphere." The generation of these ideas is seen as being imperiled by disruption of the privacy which is necessary for their development. It is interesting to note that a similar note was sounded by one of the participants. This participant described how internal communication took place about an outage incident. The internal communication led to a better understanding of the full circumstances of the outage and a clearer report that was shared with the customer (See [G1P5] comment in the Reputation Considerations Section). The privacy of these internal communications is analogous to the private deliberations of an individual.

Third, we can see in retrospect that the types of privacy behaviors previously seen in individuals [12] can also be seen in the communities described by the focus group. The study of younger and older adult's privacy behaviors identified three categories of behaviors: avoidance, modification, and alleviatory. Avoidance, taking steps that circumvent situations where privacy is at risk, can be seen in the use of a "room without windows" to conduct a sensitive human resources action (See [G1P2] comment in the Appropriate Channel Section).While the specific actors were individuals, the behavior was grounded in the community. Modification, taking steps to reveal as little as possible, is similar to the example cited in the focus group of "scrubbing data" being given to an external consultant (See [G1P3] comment in the Organization Needs Section). This scrubbing was done to minimize the exposure of sensitive customer data. Again, the modification behavior, though enacted by individuals, makes the most sense against the backdrop of the larger community. Finally, alleviatory behavior – preventing the spread of information or reducing its consequences – has some parallels in the focus group study. For example, the action of stopping of removing an email from the email queue

before it goes out to the wrong recipient (See [G2P4] comment in the Misidentification Section).

The three focus group interviews we conducted show that communities are as concerned about the privacy of sensitive information entrusted to their care as individuals are concerned about their personally identifying information. The same factors that determine the level of sensitivity of information for individuals are also seen in community settings. Instead of just one person actively to acting to control the dissemination of the sensitive information, collectively the individuals take responsibility for the information shared within their communities. Also, just people manage privacy through a personal boundary regulation [6, 13], in the analysis of the focus group data we see the communities striving to regulate their boundaries with other communities. The findings from the analysis of the focus groups confirm that community privacy is as significant and nuanced as individual privacy. We also observe communities struggling with inadequate technology support to achieve their privacy goals. Through the development of our privacy model and additional study of real communities we hope to provide better understanding and support for community privacy.

CHAPTER 6

CONCLUSION

The purpose of the focus group studies conducted for this thesis was to develop a better understanding of the factors contributing to the sensitivity of community information, of the privacy threats that are recognized by the community, and of the means by which the community attempts to fulfill their privacy responsibilities. We were also interested in seeing how the elements of a community privacy model which we developed are related to the findings from the studies of actual communities.

We conducted a series of focus group study of community privacy in three organizations: two technology companies, and one research group from a university. We reported the findings of the focus group studies in the analysis chapter and reflected on the findings in the discussion chapter. We recognize, of course, the limited interpretation that can be given to our studies. However, we believe that there are novel aspects of our perspective on privacy and there are interesting outcomes of these studies.

We approached the focus group studies with the belief that community privacy was important but less studied than individual privacy. Furthermore, we had tentative ideas regarding the major elements of a community privacy model. In these studies we believe that all of the community privacy model elements have been given a degree of affirmation.

First, the notion of *community privacy* itself was evidenced by the behaviors reported in the focus group and categorized above under *privacy awareness*. Individuals, acting as agents within various communities within the company, sought to limit the disclosure of information sensitive to their members, to their customers, or to the organization itself. The sensitivity resulted from a variety of factors ranging from legal mandates to simple concern for the dignity of their co-workers. The complexity of these communities and the interactions between business and privacy factors were illustrated in several situations.

Second, the analysis of *situated behavior* gave further substance to the idea of *community privacy*. In some cases, extraordinary circumstances motivated the disclosure of sensitive information outside of the usual community. Cases of extra-community dissemination were seen either when information was disclosed outside the company or to another individual within the company. We believe that the idea of an *exception* mechanism is a useful basis for beginning to develop prototype privacy systems for communities.

Third, we observed in the analysis of the *confinement of sensitive information* support for the idea of an explicit structure to expresses privacy boundaries, our notion of *community tags*. In several cases we observed the practice of using barriers provided by communication mechanisms (e.g., email groups). In other cases we saw how aspects of the communication mechanism or the content were appropriated to hold metadata that gave privacy hints to the receiver. However, the variety of breakdowns reported showed that these ad hoc or informal approaches were insufficient. We noted in the discussion

chapter that many of these misclosures can be eliminated or minimized by the use of community tags.

Finally, we believe that the phenomenon being studied is one of privacy. The parallels that were observed with individual privacy principles and the correlation with individual privacy behaviors are highly suggestive of this point of view.

Limitations

We note a number of limitations in the focus group studies conducted. These limitations can be summarized as size, similarity and protocol. First, we had a small number of participants; we were only able to conduct the focus group studies with sixteen participants from three organizations. Second, all sixteen participants came from groups that are representative of a certain type of communities, highly structured with lines of authority. These groups are professional organizations as supposed to social and volunteer organizations. All three groups are US companies in the same geographic region with the same business culture. Third, we recognize the limitation of our protocol. Our questionnaire was designed to probe issues we thought that were important according to our priori point of view on community privacy. We believe that the way we designed our questionnaire could cause us to miss out on other critical issues. For example, our questionnaire did not probe issue related to time. Fortunately some of the participants brought it up during the focus group discussion. As future work, we will need to study less structured and ad hoc communities to understand the factors contributing to the sensitivity of community information and compare the findings. We may also ask different questions.

Technology Implication

Technology can help solve many problems raised in the focus group discussions. We believe that our privacy model can be applied to several technologies to limit the dissemination of information. For example an email client leveraging the tagging mechanism of our community privacy model can prevent or minimize privacy breakdowns (i.e., misclosures). The email client would trigger a notification to the community represented by a tag when email actions are taken by an individual member of the tag. These email actions include forwarding tagged emails to others outside of the tag or adding more recipients to an email thread than original recipients. Our community privacy model can also be applied to a streaming media application to limit a conference call to only members of a specific tag. The application would use the community privacy model infrastructure to screen participants to ensure that they are members of the actual tag. While we acknowledge the application of our community privacy model in different technology to limit dissemination of sensitive information, we also recognize that there are cases where technology cannot prevent all forms of misclosuress.

To summarize, for the thesis we conducted a series of focus interviews in three organizations to understand the factors contributing to the sensitivity of community information and the means by which the community attempts to fulfill their privacy responsibilities. Three themes emerged from the analysis of these focus group interviews which are described as *privacy awareness, situated disclsoures, and confinement of sensitive information*. These three themes capture the character and complexity of community oriented privacy and expose breakdowns in current approaches.

REFERENCES

1- Adams, A., and Blandford, A., *Bridging the Gap Between Organizational and User Perspectives of Security in the Clinical Domain.* International Journal of Human-Computer Studies, 2005. **63**(1-2): p. 175-202.

2- Ackerman, M.S., *Privacy in pervasive environments: next generation labeling protocols.* Personal Ubiquitous Comput., 2004. **8**(6): p. 430-439.

3- Ackerman, M., and Mainwaring, S., *Privacy issues and human-computer interaction.* Security and Usability: Designing Secure Systems That People Can Use. 2005: Sebastopol, CA: O'Reilly S, pp. 381–400.

4- Acquisti, A., Friedman, A., and Telang, R., *Is there a cost to privacy beaches? Anevent study*. In *Proceedings of the International Conference of Information Systems (ICIS)*, 2006.

5- Aggarwal, C.C. and Yu, P.S., eds. *Privacy-Preserving Data Mining: Models and Algorithms*. Advances in Database Systems. 2008, Springer. 513.

6- Altman, I., *Privacy Regulation: Culturally Universal or Culturally Specific?* Journal of Social Issues, 1977. **33**(3): p. 66-84.

7- Beguin, P., and Clot, Y., *Situated Actions in the Development of Activity.* Activities, 2004. 1(2) 50-63.

8- Bellotti, V. and Sellen, A., *Design for Privacy in Ubiquitous Computing Environments*. In *Proceedings of the Third Conference on European Conference on Computer-Supported Cooperative Work*. 1993: Kluwer Academic Publishers.

9- Bobba, R., et al., *Attribute-Based Messaging: Access Control and Confidentiality.* ACM Trans. Inf. Syst. Secur., 2010. **13**(4): p. 1-35.

10- Bobba, R., et al., *Usable secure mailing lists with untrusted servers*, in *Proceedings of the 8th Symposium on Identity and Trust on the Internet*. 2009, ACM: Gaithersburg, Maryland. p. 103-116.

11- Bellotti, V. and Sellen, A., *Design for Privacy in Ubiquitous Computing Environments*. in *Proceedings of the Third Conference on European Conference on Computer-Supported Cooperative Work*. 1993: Kluwer Academic Publishers.

12- Caine, K., *Linking Studies of HCI to Psychological Theories of Privacy. Georgia Institute of Technology*. 2008.

13- Caine, K., *Exploring Everyday Privacy Behaviors and Misclosures.* Dissertation, Georgia Institute of Technology, 2009.

14- Cavoukian, A., *Privacy by Design, the 7 Foundational Principles.* Information and Privacy Commissioner, 2009 Ontario, Canada.

15- Cheng, H.-T., Lin, C.-L., and Chuinst H.-h., *A Collaborative Privacy-Enchanced Alibi Phone*, in *Advances in Grid and pervasive Computing, Lecture Notes in Computer Science*, Y.-C. Chung and J. Moreira, Editors. 2006, Springer: Berlin / Heidelberg. p. 405-414.

16- Community of Practice. Etienne Wenger. 2004. 03May 2012 < http://www.ewenger.com/theory/ >.

17- Covington, M.J., Moyer, M.J., and Ahamad, M., *Generalized Role-Based Access Control*. In *21st International Conference on Distributed Computing Systems*. 2001, IEEE Computer Society. p. p. 391.

18- Cranor, L.F., *Introduction to P3P*. In *Web Privacy with P3P*. 2002, O'Reilly and Associates. p. 3-11.

19- Cranor, L.F., Arjula, M., and Guduru, P., *Use of a P3P user agent by early adopters*. In *Proceedings of the 2002 ACM workshop on Privacy in the Electronic Society*. 2002, ACM: Washington, DC.

20- Dourish, P. and Anderson, K, *Collective Information Practice: Exploring Privacy and Security as Social and Cultural Phenomena.* Human-Computer Interaction, 2006. **21**(3): p. 319-342.

21- Dourish, P., et al., *Security in the Wild: User Strategies for Managing Security as an Everyday, Practical Problem.* Personal Ubiquitous Computing, 2004. **8**(6): p. 391-401.

22- Dourish, P., *What we talk about when we talk about context.* Personal Ubiquitous Comput., 2004. **8**(1): p. 19-30.

23- Ernst & Young. *Top 11 Privacy Trend for 2011*. 2011. 24 Apr. 2012. < http://www.ey.com/GL/en/Services/Advisory/IT-Risk-and-Assurance/Top-11-privacy-trends-for-2011---9--Privacy-by-design>.

24- Fischer, G., *Communities of interest: learning through the Interaction of Multiple Knowledge Systems*, in *24th Annual Information Systems Research Seminar in Scandinavia (IRIS'24)*, S. Bjornestad, et al., Editors. 2001: Ulvik, Norway. p. 1-14.

25- Flechais, I., Riegelsberger,J., and Sasse, M.A., *Divide and Conquer: The Role of Trust and Assurance in the Design of Secure Socio-Technical Systems*. in *Proceedings of the 2005 Workshop on New Security Paradigms*. 2005. Lake Arrowhead, California: ACM.

26- Foster, I., Kesselman, C., and Tuecke, S., *The Anatomy of the Grid - Enabling Scalable Virtual Organizations.* International Journal of Supercomputer Applications, 2001. **15**: p. 25 pp.

27- Garfinkel, S.L. and Miller, R.C., *Johnny 2: a user test of key continuity management with S/MIME and Outlook Express*, in *Proceedings of the 2005 symposium on Usable privacy and security*. 2005, ACM: Pittsburgh, Pennsylvania.

28- Google. *About circles*. 2011. <http://support.google.com/plus/bin/answer.py?hl=en&answer=1047805>.

29- Homeland Defense Journal, 2007.

30- Iachello, G., et al., *Developing privacy guidelines for social location disclosure applications and services*, in *Proceedings of the 2005 Symposium on Usable Privacy and Security*. 2005, ACM: Pittsburgh, Pennsylvania.

31- Jang, J., et al., *Collaborative Privacy Management System*, in *Proceedings of the 2008 International Conference on Information Security and Assurance.* 2008, Washington, DC.

32- Kafura, D., et al., *An Approach to Community-Oriented Email Privacy*. In *IEEE International Conference on Privacy, Security, Risk, and Trust*. 2011

33- Karat, C.-M., Karat, J., & Brodie, C., *Editorial: Why HCI research in privacy and security is critical now.* International Journal of Human-Computer Studies, 2005. *63*(1), p.1-4.

34- Karat, C.-M., et al., *Evaluating interfaces for privacy policy rule authoring*. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 2006.

35- Karat, C.-M., Karat, J., Brodie, C., *Usable Privacy and Security for Personal Information Management.* Communication of the ACM, 2006. 49(1), 56-57.

36- Kiemer, J., et al., *Spybuster - a community-based privacy tagging platform*, in *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*. 2008, ACM: Amsterdam, The Netherlands. p. 491-492.

37- Kolter, J., Kernchen, T., and Pernul, G., *Collaborative Privacy – A Community-Based Privacy Infrastructure,* in *Emerging Challenges for Security, Privacy and Trust*, D. Gritzalis and J. Lopez, Editors. 2009, Springer Boston. p. 226-236.

38- Kolter, J., T. Kernchen, and g. Pernul, *Collaborative privacy management.* Computers & Security, 2010. **29**(5): p. 580-591.

39- Lacey, A., Luff, D., *Qualitative Data Analysis.* The NIHR Research Design Service for Yorkshire & the Humber, 2007.

40- Lampinen, A., et al., *We're in it together: interpersonal management of disclosure in social network services*, in *Proceedings of the 2011 annual conference on Human factors in computing systems*. 2011, ACM: Vancouver, BC, Canada. p. 3217-3226.

41- Lederer, S., Mankoff, J., & Dey, A. K., *Who wants to know what when? Privacy preference determinants in ubiquitous computing*. In *CHI '03.*

42- Liu, D., X. Wang, and Camp, L.J., *Mitigating Inadvertanent Insider Threats via Incentives,* in *Financial Cryptography and Data Security FC'09)*. 2009, Springer: Barbados. p. 16.

43- Liu, H., Krishnamachari, B., and Annavaram, M., *Game theoretic approach to location sharing with privacy in a community-based mobile safety application*, in *Proceedings of the 11th international symposium on Modeling, analysis and simulation of wireless and mobile systems*. 2008, ACM: Vancouver, British Columbia, Canada. p. 229-238.

44- NASCIO, *State CIOs Take Action Now!*, in *Technical Report: National Association of State Chief Information Officers*. 2007.

45- Ni, Q., Bertino, E., and Lobo, J., *An obligation model bridging access control policies and privacy policies*, in *Proceedings of the 13th ACM symposium on Access control models and technologies*. 2008, ACM: Estes Park, CO, USA.

46- Nissenbaum, H., *Privacy as Contextual Integrity*. Washington Law Review, 2004. **79**(1): p. 119-158.

47- Palen, L., and Dourish, P. , *Unpacking "Privacy" for a Networked World*, in *CHI'03*. 2003, ACM: Ft. Lauderdale, Florida, USA. p. 129-136.

48- Petronio, S., *Boundaries of Privacy: Dialectics of Disclosure*. State University of New York Press, 2002.

49- Rafaeli, S. and Hutchison, D., *A survey of key management for secure group communication.* ACM Comput. Surv., 2003. **35**(3): p. 309-329.

50- Smetters, D.K. and Good,N., *How users use access control*, in *Proceedings of the 5th Symposium on Usable Privacy and Security*. 2009, ACM: Mountain View, California. p. 1-12.

51- Solove, D., *A Taxonomy of Privacy.* University of Pennsylvania Law Review, 2006. **154**(3): p. 477-560.

52- Sparck-Jones, K., *Privacy: What's different now?* Interdisciplinary Science Reviews, 2003. (28): p. 287-292.

53- Squicciarini, A.C., Shehab, M., and Paci, F., *Collective privacy management in social networks*, in *Proceedings of the 18th international conference on World wide web*. 2009, ACM: Madrid, Spain. p. 521-530.

54- Squicciarini, A., F. Paci, and Sundareswaran, S., *PriMa: an effective privacy protection mechanism for social networks*, in *Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security*. 2010, ACM: Beijing, China. p. 320-323.

55- Srivastava, A. & Thomson, S. B., *Framework Analysis: A Qualitative Methodology for Applied Policy Research*. 2009, Research Note JOAAG, 4 (2)

56- Stanton, J. M., *Information technology and privacy: A boundary management Perspective*, in *S. Clarke, E. Coakes, G. Hunter, & A. Wenn, Socio-Technical and Human Cognition Elements of Information Systems*. 2002, London: Idea Group. p. 79-103.

57- U.S Department of Education. *Family Educational Rights and Privacy Act*. 2011. 26Apr.2012< http://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html >

58- Warren, S., D., and Brandeis, Louis, D.,, *The Right to Privacy*. Harvard Law Review, 1890. **IV**(5).

59- Wenger, E., *Comunities of Practice. Learning as a social system.* Systems Thinker, 1998.

60- Whitten, A. and Tygar, J.D., *Why Johnny can't encrypt: a usability evaluation of PGP 5.0*, in *Proceedings of the 8th conference on USENIX Security Symposium - Volume 8*. 1999, USENIX Association: Washington, D.C.

61- Zilberman, P., Shabtai, A., and Rokach, L., *Analyzing Group Communication for Preventing Accidental Data Leakage via Email*, in *Proceedings of the 2010 Workshop on Collaborative Methods for Security and Privacy (CollSec'10)*. 2010: Washington, D.C.

Appendix

Focus Group Interview Questions

1. Preamble

Introduction

Good morning. We appreciate your agreeing to meet with us. We are trying to understand how people in a group think and act about sensitive information for which they share responsibility.

By sensitive we mean information that, for whatever reason, has limitations on its dissemination or disclosure. We want to understand what actions are taken to achieve these limitations. We also want to learn about the group structure, the role of the group members and how they interact with each other. Your answers will help inform the design and study of systems dealing with sensitive information.

Interview Procedure

The interview will last about 45 minutes to an hour. We will ask you a series of questions related to the group structure and another series of questions related to the group interaction and the actions taken to control the disclosure of information. For all of these questions there are no right or wrong answers, therefore we encourage you express your opinions freely and openly.

Now, we would ask you to read the consent form and sign it before we start the interview. All information and audio will be stored in a password-protected computer. Files will only be named with your participant number that is written at the top right of this form. Do you have any questions before we begin?

2. Main Questions

A. We are interested in studying a group of individuals who share responsibility for some sensitive information. Can you tell us about such a group you are or were part of?

a) What is the purpose of the group?
b) How was it created?
c) How did you become a member of this group?
d) How do you know who else is in the group?

e) Are there ways to remove a member from or add a new member to the group? If so, can you describe how this happens?
f) How frequently does the membership of the group change?
g) How big is the group?
h) Is it a group that meets regularly or not? Online or face to face?

B. Can you give examples of sensitive information that the group share responsibility for?

a) How did the group become responsible of this information?
b) Who is the sensitive information shared with?
c) How do you have access to the sensitive information
d) How does it flow within the organization?
e) How does the group make decision on how to keep this information private? Was someone responsible to make the decision for the group? Or did the group decide in some way
f) For the duration of the management of the information, were there changes to the privacy requirements? If so, why did you make changes in them? Why not?

C. Thinking about the actions taken to limit the disclosure of sensitive information, was there a case when the sensitive information needed to be disclosed to someone outside the group?

a) Can you describe the circumstances?
b) How was it decided whether the sensitive information could be disclosed in this way?
c) Was anyone notified?
d) Were other group members aware of this disclosure?
e) Are there procedure or written directions or guidelines to follow in such cases?
f) Aside from the limitations what other factors did you take into account ?

D. Thinking about the actions taken to limit the disclosure of sensitive information, was there any other situation where the information was disclosed outside the group.

a) How did someone in the group become aware of that?
b) How was it handled?
c) What were the consequences of the inappropriate disclosure?
d) Are there any procedures or written directions or guidelines to follow when this happens?

E. In deciding on the limitations what factors did you group consider?

a) What were those other factors?
b) What were the tradeoffs among these factors?
c) How was the decision made?

3. Optional Questions

A. Did your group member overlap with other group dealing with similar information?

a) Can you describe the other group?

b) How many people are involved in the overlap?

B. Technology: what technologies have you used that help to enforce the limitations. What features help or hinder your use of these technologies to enforce the limitations.