Access Control for Home Data Sharing: Attitudes, Needs and Practices

Michelle L. Mazurek, J.P. Arsenault, Joanna Bresee, Nitin Gupta, Iulia Ion, Christina Johns, Daniel Lee, Yuan Liang, Jenny Olsen, Brandon Salmon, Richard Shay, Kami Vaniea, Lujo Bauer, Lorrie Faith Cranor, Gregory R. Ganger, Michael K. Reiter

October 19, 2009

CMU-CyLab-09-013

CyLab
Carnegie Mellon University
Pittsburgh, PA 15213

Access Control for Home Data Sharing: Attitudes, Needs and Practices

Michelle L. Mazurek, J.P. Arsenault, Joanna Bresee, Nitin Gupta, Iulia Ion[†], Christina Johns, Daniel Lee, Yuan Liang, Jenny Olsen, Brandon Salmon, Richard Shay, Kami Vaniea, Lujo Bauer, Lorrie Faith Cranor, Gregory R. Ganger, Michael K. Reiter[‡]

Carnegie Mellon University
Pittsburgh, PA

[†]*ETH Zürich*
Zürich, Switzerland

[‡]*University of North Carolina*
Chapel Hill, NC

ABSTRACT
As digital content becomes more prevalent in the home, non-technical users are increasingly interested in sharing that content with others and accessing it from multiple devices. Not much is known about how these users think about controlling access to this data. To better understand this, we conducted semi-structured, in-situ interviews with 33 users in 15 households. We found that users create ad-hoc access-control mechanisms that do not always work; that their ideal polices are complex and multi-dimensional; that a priori policy specification is often insufficient; and that people's mental models of access control and security are often misaligned with current systems. We detail these findings and present a set of associated guidelines for designing usable access-control systems for the home environment.

Author Keywords
access control, home computing, human factors, privacy, security

ACM Classification Keywords
D.4.6 Security and protections: Access controls; H.3.2 Information Storage: File organization; H.5.2 Information Interfaces and Presentation: User-centered design

1. INTRODUCTION

Digital content is increasingly common in the home, as new content is created in digital form and people digitize their existing content. Devices such as digital cameras, mobile phones and portable music players make creating and interacting with this content easy. Home users are increasingly interested in sharing this content, both inside and outside their homes, across computers and other digital devices [1, 7]. Researchers and corporations are already developing new systems [20, 10, 15] to meet this need.

Providing secure, usable access control for this mobile personal data may be difficult. Studies repeatedly show that computer users struggle with specifying access-control policies [21, 13]. Worse, home users are often technically inexperienced and notoriously impatient with complex interfaces. Large organizations have system administrators to set up and maintain access-control policies, but home users typically have only themselves, family members and friends.

Not much is known about how people think about and interact with access control in the home environment [5]. It is not yet known how much or what kind of access control is required in order for home-data-sharing systems to be usable while providing the protections users desire. As a first step, we conducted semi-structured interviews with 33 non-technical computer users in 15 households to examine their current access-control attitudes, needs and practices. We also discussed hypothetical scenarios with them to understand what their needs and preferences might be in a world where sharing digital files is routine and ubiquitous.

1.1 Key findings

Analysis of the interview data led to several findings. First, we found that people construct a variety of ad-hoc access-control mechanisms, but the procedures they use do not entirely allay their concerns about protecting sensitive data. Second, we found that people's ideal access-control policies can be complicated; they are not always defined exclusively in standard role-based terms but can also incorporate factors like who is present, where the access occurs and what device is being used. Third, we found that a priori policy specification is often insufficient, because it does not align well with many people's social models of politeness and permission. In addition, many participants expressed a desire to update their policies iteratively in reaction to data access requests. Fourth, we found that people's mental models of access control and of computer security in general are often misaligned with current system designs in ways that could leave them vulnerable. From these findings, we distill a set of guidelines for designing usable access-control systems for digital data in the home environment.

1.2 Background and related work

Family dynamics and social norms are important in the home context. When working in small groups, as within a household, people often establish social rules that allow them to function without tight security [2]. Preliminary studies by Salmon et al. show that home users trust the other members of their households and expect them not to pry beyond clearly marked boundaries. Instead of using technology to protect their files, users hide files or store them on devices identified as off-limits to others [19]. In focus groups targeting the implications of ubiquitous shopping technology, Little et al. found that family dynamics play an important

role in information control within a household. Users in that study expressed concern about the effects such technology would have on family social balance [12]. Our work expands on these ideas by focusing on home users' current practices as well as future needs for access control within and outside their households.

Studies of academic and corporate environments have found that users have dynamic access-control policies that can change quickly [3, 8, 16]. Dourish et al. also found that younger office workers tended to articulate more complex security needs [8]. Razavi and Iverson found that students using a personal learning space called Elgg had a strong need for privacy controls. As documents moved through a life cycle from draft to finished, their privacy needs changed too. Users liked sharing some finished documents with wide audiences but wanted to control the visibility of private and work-in-progress documents that should only be seen by a small set of groups. Users also found that managing access controls to match these preferences was too labor intensive [16]. In a lab study, Olson et al. explored how comfortable people were sharing different types of data (including information such as age and salary, as well as digital files) with different types of people. They found that both people and data generally clustered into a small number of intuitive groups based on interpersonal trust relationships. They also recommended, however, that interfaces allow users to choose the level of granularity with which to specify access control and allow for exceptions [14].

Other studies have examined the use of computer accounts at home. Unlike in a structured enterprise environment, where each person has her own account, home users tend to share a single account on the family computer, which obscures the connection between users and data accesses [6, 19]. Account sharing is primarily driven by convenience; the ability to quickly access the computer outweighs the privacy and security concerns that can be solved with multiple accounts [9].

During our study, we asked participants to think about using a *reactive policy-creation system* for their files. A reactive policy-creation system allows a file or resource owner to make a semi-real-time policy adjustment in response to an attempted access that cannot otherwise succeed. For example, in a deployment of the Grey system [4], users send messages with their cell phones to the "owners" of doors they wish to open in a university building. Door owners can respond with their own cell phones and grant access. Bauer et al. found the Grey system enabled users to construct policies closer to their ideal policies than did keys [3].

In the remainder of this paper, we describe the methodology of our study, discuss the key findings in more detail and present system design guidelines based on our findings.

2. METHODOLOGY

We designed this study to increase our understanding of how home users think about controlling access to their digital files. We did not start out with any hypothesis; instead, as we conducted and analyzed interviews, we developed theories about home users' preferences, needs and mental models for access control. We gathered data using semi-structured, in-situ interviews and then analyzed it using an iterative method. Our methodology is very similar to the Grounded Theory methodology used by Razavi and Iverson [16].

2.1 Participants

We recruited participant households from a medium-sized city and surrounding areas using a variety of methods. We posted on Craigslist, sent e-mails to university distribution lists, hung advertisements at local grocery stores and recruited families at children's soccer games. Households were prescreened to include those with a range of digital devices storing at least a moderate amount of personal data, but to disallow households that included computer programmers or software engineers. We targeted three types of households: romantic couples, roommates and families with children. In total, we interviewed 33 people, ranging from elementary-school students to retirees. Participant households, which included five couples, six families and four sets of roommates, were paid $50 in compensation.

In this paper, we refer to participants using a naming scheme that identifies their household type (C for couples, R for roommates or F for families), household number within that type and member letter. For example, participant R2A belongs to the second roommate household (R2) and is the first of that household to be interviewed (A).

2.2 Interview protocol

Interview sessions were conducted in participants' homes using a semi-structured framework. Households were first interviewed as a group, and then each participant was interviewed separately. All interviews were structured around a predetermined set of questions designed to cover a wide range of access-control-related topics. The questions were intended to encourage participants to discuss their past experiences along with their current behaviors, thoughts and feelings. If a participant mentioned an interesting topic not in the questions, the interviewer probed further, but otherwise kept to the question list. At least two interviewers attended each session, and the interviews were recorded, resulting in more than 30 hours of videotape.

Group interviews lasted approximately half an hour each and included all available members of the household. In these sessions, we asked participants about how they currently protect important information both on paper and digitally. To guide participants' thinking, we asked them to draw maps of their homes and illustrate which devices and rooms they considered to be public or private. We also asked about current formal and informal rules and policies about who can use which devices under which circumstances.

Individual interviews with each participant lasted between 30 minutes and 1.5 hours. The goal of the individual interviews was to understand how participants define their ideal access-control policies and what features they would find useful to implement desired policies. The interview protocol had three major components, which are discussed below.

First, we asked participants to describe past experiences when they were concerned that others might view or modify data in an unwanted way. This section was used to prime the pump: to focus the participants on why and when access control policies would be important to them.

Second, we used the map they drew in the group interview to walk participants through a list of the types of digital data they own, asking generally about which types are more private and which types are more public. This list was used to guide the rest of the interview.

In the third section, we presented participants with ten scenarios in order to learn whether and to what degree various dimensions of policy definition would be useful for them. The scenarios tested potential policy specification factors including: who attempts the access, the location of the accessor, the device used for access, whether or not the file owner is present during the access, the time of day access is attempted, the location of the file owner and the incidence of social events. We also asked participants to respond to possible access-control features including privacy indicators, a detailed access log and reactive policy creation.

We prompted participants with specific events and people drawn from their sets of data in an attempt to discern their general attitudes toward specific access-control mechanisms. For example, in order to assess whether participants wanted to restrict assess based on person, we picked two people that the participant had mentioned – a close friend or family member and someone they were not close with – and asked: "Imagine _____ could view all of your files and data. What would you not want them to see or change?" We gauged the strength of these preferences by asking participants how upsetting a violation would be, using a five-point Likert scale ranging from 'don't care' (1) to 'devastating' (5). We noted the scenarios that resonated with participants and elicited strong examples.

2.3 Data analysis

As we completed early interviews, we began the process of data analysis by applying a coarse-grained coding method to assess the general level of positive response to each scenario. We recorded whether each participant was generally very interested, somewhat interested or uninterested in each axis of control, along with high-level explanations of their responses. Using this method, we recognized several interesting patterns that helped focus later sessions. Additional questions and clarifications related to these patterns were added to the interview protocol.

Once interviews were complete, we iteratively coded each interview. The coding was collaborative, with team members working together to validate each other's results. We transcribed all the videos and then applied a more detailed process of topic and analytic coding as described in [17, 18]. We recorded codes for each aspect of participants' answers in a searchable database designed for easy cross-referencing of participants and topics. As new concepts emerged, we revisited previously analyzed transcripts to see how the new concepts related. By grouping individual codes into larger categories, we were able to formulate broader theories about participants' access-control concerns, needs and preferences.

The results of our study are purely qualitative. We sometimes report the number of participants who fall into a given category to provide context; this is not intended to imply statistical or quantitative significance.

3. PEOPLE NEED ACCESS CONTROL

Unsurprisingly, we observed that people have data they consider important or sensitive, and they want to ensure this data is protected. We discuss this result here for completeness as well as to shed light on specific concerns participants raised and on the sometimes-surprising ways in which they accomplish their access-control goals. In the first subsection, we demonstrate that people have data they classify as sensitive, and they find the idea of unauthorized people accessing this data disturbing. Next, we provide evidence that these concerns are not just hypothetical; several participants described incidents where their data was put at risk or exposed. Finally, we discuss ways that people construct their own ad-hoc access-control mechanisms using both technical tools and social norms.

3.1 People have data they classify as sensitive

Almost all participants want to limit access to their personal data. When we asked participants to imagine a breach of their ideal policy, we found preferences for access limitations are often very strong. Eighteen participants across 13 households classified at least one hypothetical policy violation as a 4 or 5 on our Likert scale. These devastating or near-devastating scenarios included unauthorized access (read, modify or delete) to financial data, schoolwork, e-mail, hobby or activity files, work files, text messages, photos, home videos, journal files and home musical recordings.

Many participants considered unauthorized access by strangers, acquaintances, bosses and teachers to be highly undesirable. Perhaps more surprisingly, several were equally disturbed by situations involving closer relationships like parents, children, family, friends and even significant others.

Examples of these critical violations (along with their Likert scores) include F4A's children seeing her finances (4); C1B's boss seeing her photos (4); R4A's boyfriend seeing her personal documents or work files (4), or modifying them (5); and nine-year-old F5D's friends seeing her e-mail (4).

3.2 People's concerns are not just hypothetical

Our results reveal that not only do people have data they want to protect, but that their current protection mechanisms are not always adequate (or are not perceived to be adequate). We asked participants to recall situations where they were concerned their sensitive data might be at risk, as well as situations where their ideal access policy was actually breached and their data was accessed improperly. Twenty-two participants could recall specific instances of concern about someone viewing or modifying their data without permission; only six reported they had never had such concerns.

Nine participants reported actual policy breaches of varying degrees of severity.

Participant F4A, a divorced mother of two teenage boys, reported concern about her sons accessing her e-mail when she leaves her account logged in on a family computer. "Maybe someone sort of e-mails you a sexy e-mail, or something, and I wouldn't want the kids to see it."

R4B was upset when he caught a roommate in his bedroom, using his computer without permission. F2B said her roommate sometimes grabs her phone and looks through pictures on it without asking, which is "kind of uncomfortable." R4C has also had private photos exposed on more than one occasion, including one incident where his girlfriend "stumbled upon an ice skating video of me and my ex. And it wasn't anything, but it was an awkward moment."

R2A, a law student, once lent her computer to her adolescent sister, who inserted random words into a class assignment. R2A turned in the altered paper without noticing and later had to apologize to the professor. Participant F1A reported a less serious but still annoying instance of data modification: his wife accidentally deleting shows from their DVR before he had watched them. "It's frustrating, because you're expecting to see it.... But what can you do, it's already done, it's gone."

3.3 People use a variety of access-control mechanisms

Because people are concerned about limiting access to their sensitive files, they take precautionary measures to reduce the risk of exposure. We found that while some people use standard tools designed for access control, many others have developed ad-hoc procedures. These procedures include both technical and social mechanisms whose actual efficacy may vary, but which participants find reassuring. In total, 30 of 33 participants, including at least one in every household, reported using precautionary measures, several of which are discussed below.

Use accounts, passwords and encryption. Seven participants use passwords, encryption or separate accounts for access control. Four said they are careful to log out or lock the computer when they walk away. R4C said, "I guess I'm a security junkie with my phone. Encrypting my text messages, it's not really necessary. But it makes me feel comfortable." Like most participants who used passwords, R4A protects her laptop rather than individual files. She said she uses the password "just in case when we have guests over, that nobody thinks that, 'Well, it doesn't have a password, that means I can use it.' Just to better my chances of not having my identity or secret information taken."

Limit physical access to devices. In most participants' configurations, data boundaries are device boundaries; anyone using the device has access to all the data stored on it. As a result, many participants are cautious about lending their devices to others, even for common tasks like checking e-mail or browsing the web. Most participants allow only people they trust to access their devices. As 15-year-old F4C said, "Obviously I don't let anyone who walks through the door on to my computer, but if someone's on my computer I trust them." A few participants only allow others to use their devices if they are present to supervise, and another few don't allow it at all. Some participants shut down or put away their devices in order to discourage others from using them. One participant keeps her most important data on an external hard drive, which she physically hides from her roommates.

Hide sensitive files. Participants also attempt to hide files within the file system: A few have named sensitive files obscurely for concealment, and others bury sensitive files in layers of directories. According to R2A, "If you name something '8F2R349,' who's going to look at that?" C2B said, "[My husband] is a good hider of things.... If someone was trying to find something specific and he had it hidden, it would take them a while."

Delete sensitive data. Six participants have deleted sensitive files to prevent others from seeing them. F1A has deleted pictures of his two-year-old daughter from his cell phone for this reason: "If I didn't want everyone to see them, I just had them for a little while and then I just deleted them." A few participants have closed existing Facebook accounts because of privacy concerns.

4. PEOPLE NEED FINE-GRAINED ACCESS CONTROL

In practice, many current access-control systems designed for home users favor simple, coarse-grained access policies. In Windows XP, the default "My Documents" and "Shared" folders divide a user's files into those accessible only to her and those accessible to everyone on her network. Although more fine-grained controls are available, they may not be sufficiently usable, as evidenced by participants' hiding files in the directory structure or giving them confusing names. Apple's iTunes offers options for sharing the user's entire library, sharing only selected playlists and requiring a password for the shared files. This configuration does not allow users to share different subsets of music with different people. Facebook supplies rich, customizable access controls for photo albums, but there is no differentiation between reading and writing. Any user who can view a photo can also tag it and leave comments on it. The HomeViews system [10], designed to enable easy data sharing for home users, is limited to read-only access.

Our results indicate people's policy preferences may be incompatible with coarse-grained control mechanisms in several ways:

- Some participants' policies include fine-grained divisions of people and files.
- Additional dimensions of policy specification beyond person and file are also important in some circumstances.
- Even when individual policies are relatively simple, comparing across participants shows little consistency; there is no small set of default policies that could meet most people's needs completely.

In the following sections, we discuss each of these complicating factors.

4.1 Fine-grained division of people and files

Early in our individual interviews, we asked participants to explain which people they would allow to access which files. We found that many participants specified complex groupings for both dimensions.

For some participants, an ideal policy specification required many different categories of files, some of which had fuzzy boundaries. C5B, for example, made several kinds of distinctions among her photos. In her first attempt to categorize her photos, she divided them simply into photos she was willing to publish and those she wasn't. After further thought, she divided the restricted photos into four separate categories: truly private photos as well as separate groups to share with family, sorority sisters, and general friends. Even these distinctions did not prove entirely adequate – there were some pictures she might only want to share with those people who were in them. She also said her boyfriend could see some of the truly private photos, but not others, particularly those involving ex-boyfriends. R4C had a similarly complex division of his photos into sometimes overlapping categories; he also mentioned different photos taken at the same event that should carry different restrictions. Currently, both of these participants manage photo-sharing by over-restricting; if they don't feel they can control access to a photo precisely enough, they decline to share it at all.

The need for multiple policy subdivisions is not unique to photos; other participants specified similar distinctions within categories like music, videos, school files and work files.

Our results also indicate people, like files, cannot be easily divided into just a few groups. Popular person designations included significant other, friends, family, co-workers and strangers, but these groups often required additional subdivision. Several participants differentiated policy for one or two "best" friends; others made distinctions among close friends, casual friends and acquaintances. Within families, policy varied for siblings, parents and children. R2A said she is "far more willing to show my sister things than my parents." At work, participants make distinctions between bosses and colleagues as well as within groups of colleagues. C5A even differentiated among strangers: "I think I would feel less embarrassed if I knew someone 100 miles away was looking at it [sensitive files] rather than someone on the bus."

Figure 1 summarizes one participant's ideal policy, indicating which files she would share (white), restrict (black), or sometimes share (gray) with which people. As this fairly typical policy makes clear, access decisions are not binary across people or file types. The presence of many gray squares indicates a finer level of detail would be required to completely specify this policy.

4.2 Dimensions beyond person and file

Other factors besides the person requesting the access and the file being accessed also inform participants' ideal policies. We asked participants to think about the differences between read and write permissions, as well as whether or not the participant was present during the access, the participant's location, the location of the accessor, the device used for access and the time of day of the access. Each of these factors was meaningful to at least a few of the participants.

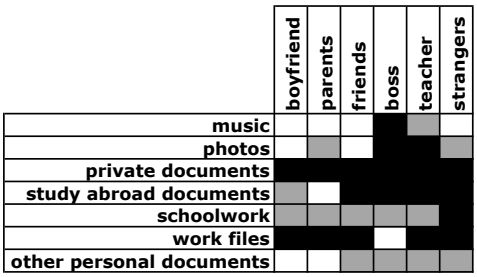

Figure 1. This figure shows a high-level view of participant R4A's ideal policy. This policy is complex, and is not binary across people or files. White squares indicate a willingness to share; black squares indicate restriction; and gray squares indicate a willingness to share some files under some circumstances.

Distinguishing read access from write access. Many participants in our study described important policy differences between read access and modify/delete access. F4C said no one else should ever be able to modify any of his files; C2A and F1B were willing to grant their bosses only read access to some files. A few participants described general categories of files they were not concerned about sharing, but that they would want to protect from modification or deletion, including music, game files, schoolwork and photos.

This read-write distinction extends to highly trusted people such as family members and significant others. In one of several examples, middle-school-student F5C was willing to share almost all of her files with her family members, but did not want to grant modify or delete permissions. Similarly, R4A was willing to share highly sensitive files such as financial information and photos with her boyfriend, but did not want to grant him write access to any files.

On the other hand, a read-only system would not be sufficient for some participants, who see value in allowing others to edit their files sometimes. C5A wanted to let his mother improve his resume, and F2B would allow friends to provide feedback on scholarship essays. F5B would let her clients update business files they send her, and R2B expressed interest in allowing collaborators to edit files related to the projects they work on.

Presence. Policy specification based in part on whether or not the file owner is present resonated with a majority of participants. Participants believed that being present would allow them to exercise additional control over who accessed what, as well as providing social pressure to encourage good behavior. F3B, a nine-year-old boy, said, "If I was next to [my friend], I would know which files he would be bringing

up, but away from him I wouldn't have a clue what he was doing on my computer." R4C said, "If you have your mother in the room, you are not going to do anything bad. But if your mom is outside the room you can sneak." According to C3A, "If I'm in the house then it's likely that I'm spending time with them. If I'm not with them, I can find them and say, 'Hey! What are you doing on my computer?'"

For a few participants, being present provides additional benefits. Three said being present would allow them to make a last-minute decision to share something. Others said they wanted to be there to witness the accessor's reaction, explain things and correct any misunderstanding. F6A wanted to make sure some opinionated journal writing wouldn't be misunderstood: "I could explain myself! Totally! If only I had a crystal ball for all the times somebody got upset with me and I didn't know it. If only I could have been there, then I could have told them: No, I am a lover, not a hater!" He also mentioned being present to explain things to his children: "Most movies I want to be there with [my son] ... in case he has questions or it's too scary. I can calm him down."

Location. We asked participants how location – their own or that of the file accessor – would affect their ideal access policies. A slight majority of participants said they felt safer sharing data in their own home than in other environments. C5A said, "I don't want them to look at my e-mails or texts. But if they were here, I wouldn't care if they wanted to look at my e-mail. I don't know why, but I just feel more comfortable doing these things at home than being out in public with my information." Eight participants did not want to share any files in public places like buses or coffee shops. According to F1A, "Chang[ing] the settings as I move? That makes sense.... Going to work with the laptop vs. being at home – you might put it on extreme lockdown." In general, participants' responses to this question reflected their ideas about who was likely to be in a given location. R1C said, "At studio [at school] I am more hesitant to share my files if I am not there. In the apartment I can trust them with music or movie files. There is a mutual trust with people you live together with."

To many people, the accessor's location could be a proxy for trust: guests in the participants' homes were presumed to be trusted to some degree. Several participants said they would share more with people in their house; a few others would share more with people who were in their bedrooms, an even higher marker of likely trust. According to C2A, "I feel that if they are in my house I can control them a little more. If they are in their house, they have a freedom to do whatever they want and there is not a chance of me walking in on them."

These ideas, though popular, were not universal. Several participants said their own location would make no difference. According to F2B, "If there's a way to have a certain setting for a specific individual and have that setting not change based on location, then I wouldn't mind having the same access rights for my friends when I'm home or at school." Policies based on the location of the accessor also didn't make sense to many participants. As F6A said, "Just because you are inside my house I would not categorize my files differently than if you were not there."

Device. We also asked participants whether or not the device used for access would affect their ideal policy. Most said the device had no effect. As with accessor location, however, a large minority did find the device used for access meaningful. To several participants, including R4A, devices with smaller screens are preferable for accessing sensitive files, as "it feels more private on a smaller screen." In contrast, others worried that a private device like a mobile phone might promote sneakier behavior than a public device like a television. According to F1B, "Maybe it's worse doing it on the laptop [than the TV], because of being a bit more private about it."

Time of day. We asked participants if their policies would vary according to the time of day when access was attempted. To a large majority, this idea did not make much sense; as C2B said, "It doesn't matter the time of day.... The things that I don't want you to see, I don't want you to see at any time. And no time would be worse than another time."

A few, however, did find this possibility interesting. Some responded to time of day as a proxy for presence or awareness; they did not want to share files while they were sleeping, because they could not know about or control the transaction. Said C3A, "If it's bedtime and I'm in bed, then I don't really get to see what people are looking at if I wanted to." F3A wanted to restrict her young sons' access to files at night, when they are supposed to be asleep.

4.3 Policies vary across people and households

As demonstrated in the previous sections, some individuals' ideal policies are complex. Even when individual policies are relatively simple, however, comparing policies across participants can introduce complexity. We found that policies specified by our participants rarely overlapped, meaning no standard set of default rules can be expected to meet most people's needs.

We spoke to many participants with relatively uncomplicated policies, but found these policies rarely matched. R2B wanted to tightly restrict financial and work files, was willing to share e-mail and photos with most friends, and was not at all concerned about sharing music. R4B, by contrast, did not consider his photos private but was concerned about sharing e-mail and music, even with friends. F5B was very interested in restricting e-mail but not concerned about sharing financial or work files. F6A did not consider anything except some financial information private. Figure 2 illustrates some of these variations.

Another important area of difference was reflected in participants' basic attitudes toward privacy. Many participants, including R2A, C4B, C5B and R4C, started from the presumption that everything should be private and then named specific items to share with specific people. According to R2A, "Basically, it's my stuff; if I want you to have it I'll give it to you. If you want access to it, then ask." C5B

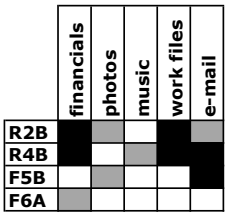

Figure 2. Participants' policies vary widely. This figure shows some of the variation among participants' ideal policies. White squares indicate file types that are generally unrestricted; black squares indicate file types that are highly restricted; and gray squares indicate file types that are partially restricted.

said, "The files that I do have are private.... If I share it with you there's a specific reason." Other participants started from the opposite position: sharing everything except a few specific exceptions. C3A, F4C, F5A and F6A were among this group. According to F5A, "I don't really have ... private files.... There's nothing that I am hiding from anybody." C3A said, "I'm not really that private. There's not a whole lot of stuff that I really want to keep from people aside from financial stuff."

In many cases, we found broad agreement on a general principle but enough variation in the details to make defining a satisfactory default policy difficult. For example, most participants identified one or two most trusted people – often a best friend or a spouse – to receive the most access. Within this group, about half were willing to grant this closest person complete access to everything; the other half wanted to grant access to most things but restrict access to some things. The specific exceptions varied from participant to participant and included everything from e-mail, photos and text messages to financial documents, work files and even web-browser history.

5. AWARENESS AND CONTROL

In the previous section, we showed that participants responded positively to several policy dimensions beyond person and file, including location, presence and device being used. Participants gravitated toward options they perceived as providing the most visibility into and control over accesses to their files. As C5B explained, "I guess I'm not a terribly private person, but I think if someone's going to be meddling in your things, you should be able to know what exactly they're looking at."

While it is not surprising that participants are looking for more control over their data, their ideas about what control means and how to achieve it show unexpected variety and depth. In the following subsections, we describe three specific manifestations of this desire for control: a preference for being asked, a need for iterative policy creation and refinement, and an interest in knowing not just who is accessing their data but why.

5.1 Permission and control

Participants often think of digital data sharing in terms of asking and granting permission. To many of these participants, setting policy a priori does not feel the same as granting permission. We found that mechanisms such as being physically present and responding to system-prompted access requests can provide a stronger sense of permission-based control and therefore increase people's comfort with data sharing.

Many participants wanted an access-control mechanism that reflected standard social conventions of asking before using someone else's things. Three participants said no one should access anything without their express permission. C2B said, "In general I want to be asked. I'd prefer to give [my files] to them. I would not want someone to just look at them." C3A agreed, saying "I'm very willing to be open with people, I think I'd just like the courtesy of someone asking me.... If you ask someone nicely for pretty much anything, people will be more than willing to help you out." When we asked R4B about who is allowed to use his devices, he answered, "Any friend of mine who asks." According to C4A, "Without my permission, without my directly sending it to you, I wouldn't like you to look at ... the financial files or my e-mail. That's my personal stuff. Not that there's a lot of high-security stuff going on with the financial stuff, there isn't ... but those are supposed to be secure areas without my permission."

To many of these participants, the idea of specifying in advance which people can access which files does not seem to convey a sufficient sense of control, possibly because they don't understand the idea of policy specification or they don't trust that policy will be enforced correctly. Several participants, when asked to describe their ideal policies, responded that no one should be able to access their files without their permission. We asked R2A what her boss should be restricted from seeing, and she responded, "Ideally I wouldn't want him to see anything except what I give him access to." Along the same lines, as discussed previously, several participants expressed concern about allowing access while they were sleeping. As R3A said, "I can't be giving you permission while I sleep because I am sleeping." Responses like these suggest that, to many participants, setting an access policy does not seem equivalent to granting permission.

Five people said that when they are physically present, they can control which files can be accessed; this provides additional confirmation that participants are not comfortable with only setting up a policy ahead of time. According to C2A, "We don't get company that much, and we are usually constantly with our company. If they were viewing something, I would be there at all times, guiding them through where they should go or not." C1B said being present would affect her policy, "because I could say, 'These are the things that you could see.'"

Participants responded positively to the idea of a reactive policy-creation system in part because they felt it would extend social conventions of permission into the digital world.

C4A said a reactive system "sounds like the best possible scenario.... It would make me feel much more comfortable if people asked before they could modify or view the files at all. I like that a lot." Others said they would use a reactive system even for files they expected to rarely or never grant access to. C5A was open to making his financial documents – designated as highly restricted – available via such a system. "I don't think I would mind, if it asked me permission first. Say if an employer needs to see it or something. I can't imagine too many people want to look at my stuff."

5.2 Iterative policy creation

For many participants, one important aspect of controlling access to their data was the ability to fine-tune policy easily and repeatedly. In general, we found a strong response to the idea that policy specification should be iterative. Some participants said they might want to make decisions about access at the last minute, if certain files became relevant to a conversation. R2B said she might change her policy "if there is something particularly relevant that I wanted to show, that I wouldn't normally want everyone to have unfettered access to." As discussed in the previous subsection, participants like C5A were interested in using a reactive policy-creation system to grant permission even to files they had not previously planned to share.

Three people placed particular emphasis on the ability to review policy and remove authorizations. C2B said, "I would like to be able to go back on there and say, 'You said yes to all these people to view these things,' but if for some reason I no longer want them to do that, I could say 'denied' now and take them off the list."

Participants were also interested in fine-tuning their policies based on observed activities. Nine participants were interested in using a detailed access log to check for unexpected or undesirable access patterns and then change policy accordingly. C1A said, "It's nice to know who is accessing data more frequently. It opens the question: Are they the only ones viewing them, or are there other people standing next to them?" R1A added, "If someone has been looking at something a lot, I am going to be a little suspicious. In general, I would [then] restrict access to that specific file."

We also found evidence that at least some ideal policies change over time. C2B wanted to temporarily limit her sister's access when they fight. "She's not talking to us right now.... She's one of those people who, if you get mad at her, ... she'll rip up all the pictures of you. ... She could erase stuff on my computer."

5.3 Not just who, but why and for what purpose

Participants wanted to know not only who was accessing their files, but also why. C4B said, "Before you even touched anything, I would have to find out why you're doing it." F2A said she would like to use a reactive policy-creation system "if I know the purpose" for the request. F4B said a reactive policy-creation system "would be very useful, especially if maybe when they sent that they could add a message as to why they needed to see it." This was especially true for write permission – C5B said that she might grant permission to modify a file, "but I think I'd probably have to get into contact with them and ask them why they wanted to."

This interest extends to knowing how files will be used. F4B said, "I feel more comfortable if they're with me or I can see them, because then I have a better idea of what they're doing with whatever files they're seeing." He also mentioned a similar concern related to the device used for access: "If it was something portable, if they're using their phone, I might be worried about who else was watching." F5A felt more comfortable sharing files in his home, where he assumed it would be impossible to show files to an unauthorized third party without his noticing. F3A wanted to limit the devices used for access out of concern about people making copies of her files: "Probably I wouldn't want them to be able to save my information on their computers. 'Cause from my devices they would be able to view it but not save it."

6. MENTAL MODELS AND SYSTEM DESIGNS DON'T MATCH

Our interviews revealed that, in many cases, a mismatch exists between people's mental models related to access control and current system designs and operations. This mismatch often occurs because users carry assumptions from the physical world into the digital world, where these assumptions may no longer be valid or are not adequately addressed by system designers. These assumptions affect the ad-hoc access-control mechanisms people create as well as the factors that make them feel secure.

Hiding files in the file system. Several participants attempt to hide sensitive files by either naming them obscurely or storing them in multiple layers of file system directories. These ideas arise from physical-world practices of hiding important items or labeling file folders to avoid suspicion. The couple in household C4, for example, keeps their most important papers in a small, hidden box; only less important papers are kept in the file cabinet, which is used as a decoy. The increasing availability of search tools, like Spotlight, Windows Search and Google Desktop, that allow fast, accurate discovery of desired content regardless of file name or directory structure may invalidate this approach.

Preventing violations with presence. Based on physical-world experiences, many users believe being physically present can prevent policy violations. R4A, for example, said, "When I let people use my laptop, I'm usually near them, because it makes me feel comfortable that if anything were to happen, ... I'm right there to say, 'OK, what just happened?' So I'm not as worried." Participants note that their presence may increase social pressure against behaving badly. They also believe they will be able to notice policy violations and react quickly enough to prevent problems. Computer policy violations, however, are often faster or less obvious than physical-world break-ins, which may complicate detection even if the file owner is in the same room as the offender.

Device boundaries. Many participants base their access-control measures on the idea that device boundaries and data boundaries are the same – anyone using a device can ac-

cess all the files on it, and no files can be accessed without physically touching the device where they are stored. As the increasing ubiquity of networking continues to blur distinctions between devices, this heuristic becomes less and less accurate. We also observed that users who subscribe to this model do not take advantage of tools like separate accounts or per-file encryption to segregate files within a device.

Location as a proxy. Some participants used the file accessor's location as a proxy for trust. For instance, based on the premise that only trusted people come into their homes, users would allow anyone within the home a high level of access to their data. It's not clear, however, that location is a particularly accurate proxy. C5B first said she would trust people in her house to access most files, but quickly changed her mind. "I guess originally my assumption would be ... if they were in the house, I'd know them, and they'd be close enough of a personal friend for them to actually be invited into my home. But then I was thinking, we've had plumbers here, guys laying carpet, stuff like that.... People are strange and might be snooping." In future work, it might be interesting to investigate whether the imprecision of this mechanism outweighs its convenience in real-world scenarios.

Infallible logs. Several participants wanted to use a detailed access log or notifications to verify enforcement of policy as well as to confront violators about their actions. F1A said, "[If] I all of a sudden got a thing [alert] on my phone, beep beep, somebody logged in to your account and is looking at it, yeah, I think that'd be great." According to C2A, a log would mean "I can call them on it [a violation] afterwards, and I would have proof of it." These statements rest on the assumption that even if the access-control system is sufficiently broken as to allow policy violations, the log or notification system would remain correct. This assumption seems dangerous, because an attacker sophisticated enough to bypass a reasonably robust access-control system may also be savvy enough to prevent her activities from being logged.

7. GUIDELINES FOR SYSTEM DESIGNERS

Based on the results of our study, we have generated several guidelines for developers of access-control systems aimed at home users.

Allow fine-grained control. We found that participants' ideal policies were often complex and varied, and they were not always defined strictly in terms of files and people. It is important to keep in mind, however, that not all policies are fine-grained, and not everyone wants to specify a detailed policy. An access-control interface should be designed to allow easy policy specification at multiple levels of granularity, according to the user's preference.

Plan for lending devices. We found that participants, especially those living with roommates, are often asked to lend their computers to others who want to check their e-mail or browse the web. Participants are often uncomfortable with these requests because they worry that the borrower will access private files or overwrite important data, either accidentally or on purpose. Karlson et al. suggest lightweight, limited-access guest profiles for mobile phones, with an emphasis on switching to this mode discreetly to avoid the appearance of distrust [11]. We suggest applying a similar approach to laptops and other devices.

Include reactive policy creation. Response to a hypothetical reactive policy-creation system was overwhelmingly positive, with 27 participants expressing interest in using such a system in at least some circumstances. R3A said, "I'd like that, it's useful. Only you can decide. That's something I would use." F4C answered, "That would be good.... Because then it would be easy access for them while still allowing me to control what they see."

Include logs. The majority of participants in our study also reacted positively to the idea of a detailed access log that would record all access attempts and their results. Some participants were interested in a log only out of curiosity, while others said that log contents might influence them to modify their policies. Six participants said they might share more if a log were available, including C4A, who said she would be "not a lot more open, but better than what I usually share." We recommend including a log or even a semi-real-time notification system designed to be human readable and to support policy changes based on log contents.

Reduce or eliminate up-front complexity. We found that although some participants' ideal policies are complex, defining fine-grained policies up front is difficult. Several participants, including C2A, reported that setting up a detailed access policy would be too much work. "If I had to sit down and sort everything into what people can view and cannot view, I think that would annoy me. I wouldn't do that." Even defining broad categories of access is seen as troublesome; participant R4C acknowledged he would not "go through the trouble of setting up a guest account" even to protect important files. As discussed earlier, some participants also had difficulty specifying an ideal policy ahead of time and expressed interest in making last-minute policy decisions. We recommend reactive policy creation, either alone or in combination with preset policy, as one possible mechanism to significantly reduce or even eliminate the up-front cost of setting up fine-grained polices.

Acknowledge social conventions. A new design for an access-control system should take into account users' interest in the social convention of asking for permission. This is another instance where reactive policy creation could be helpful.

Another social convention for which we found strong interest was the idea of plausible deniability. Participants do not want to appear secretive or sneaky; as R4A said, "I don't want people to feel that I am hiding things from them." Several participants felt nervous about admitting they had private data, and often felt compelled to justify it. C4A said, "Not that I have anything wrong or anything that can even be considered wrong, but I still want ... my privacy." Designers should take this into account and build into any new system a means of restricting files unobtrusively.

Support iterative policy specification. We found that ideal policies change over time, and users need to be able to easily review and refine their policies. We recommend creating an interface that allows users to see their current policy, review the resulting access record and make changes as needed.

Account for users' mental models. We discovered many instances where users' mental models of computer security in general and access control in particular are not well aligned with computer systems. New access-control systems should attempt either to fit into users' pre-existing mental models or to guide users to develop mental models consistent with the systems' behavior.

8. ACKNOWLEDGMENTS

We thank the members and companies of the PDL Consortium (including APC, DataDomain, EMC, Facebook, Google, Hewlett-Packard Labs, Hitachi, IBM, Intel, LSI, Microsoft Research, NetApp, Oracle, Seagate, Sun Microsystems, Symantec, and VMware) for their interest, insights, feedback, and support. This material is based on research sponsored in part by the National Science Foundation, via grants #CNS-0831407, #CNS-0756998, and #DGE-0903659; by CyLab at Carnegie Mellon under grant DAAD19-02-1-0389 from the Army Research Office; and by gifts from Intel.

9. REFERENCES

1. ABI Research. Home networking end-user snapshot: Consumer adoption of home and media networking, June 2008. http://www.abiresearch.com/research/1000323-Home+Networking+End-User+Snapshot.
2. D. B. Baker. Fortresses built upon sand. In *NSPW '96: Proceedings of the 1996 workshop on New security paradigms*, pages 148–153, 1996.
3. L. Bauer, L. F. Cranor, R. W. Reeder, M. K. Reiter, and K. Vaniea. A user study of policy creation in a flexible access-control system. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, 2008.
4. L. Bauer, S. Garriss, and M. K. Reiter. Distributed proving in access-control systems. In *Proceedings of the 2005 IEEE Symposium on Security & Privacy*, 2005.
5. K. Beznosov and O. Beznosova. On the imbalance of the security problem space and its expected consequences. *Information Management & Computer Security*, 15:420–431, 2007.
6. A. Brush and K. Inkpen. Yours, mine and ours? Sharing and use of technology in domestic environments. In *Ubicomp*, 2007.
7. D. Dearman and J. S. Pierce. "It's on my other computer!": Computing with multiple devices. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, 2008.
8. P. Dourish, E. Grinter, J. Delgado de la Flor, and M. Joseph. Security in the wild: user strategies for managing security as an everyday, practical problem. *Personal Ubiquitous Computing*, 8:391–401, 2004.
9. S. Engelman, A. Brush, and K. Inkpen. Family accounts: A new paradigm for user accounts within the home environment. In *Confereence on Computer Supported Cooperative Work*, 2008.
10. R. Geambasu, M. Balazinska, S. D. Gribble, and H. M. Levy. HomeViews: Peer-to-peer middleware for personal data sharing applications. In *Proceedings of SIGMOD International Conference on Management of Data*, 2007.
11. A. K. Karlson, A. B. Brush, and S. Schechter. Can I borrow your phone? Understanding concerns when sharing mobile phones. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 1647–1650, 2009.
12. L. Little, E. Sillence, and P. Briggs. Ubiquitous systems and the family: thoughts about the networked home. In *SOUPS '09: Proceedings of the 5th Symposium on Usable Privacy and Security*, 2009.
13. R. A. Maxion and R. W. Reeder. Improving user-interface dependability through mitigation of human error. *Int. J. Hum.-Comput. Stud.*, 63:25–50, 2005.
14. J. S. Olson, J. Grudin, and E. Horvitz. A study of preferences for sharing and privacy. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, 2005.
15. V. Ramasubramanian, T. Rodeheffer, D. B. Terry, M. Walraed-Sullivan, T. Wobber, C. Marshall, and A. Vahdat. Cimbiosys: A platform for content-based partial replication. Technical Report MSR-TR-2008-116, Microsoft Research, August 2008.
16. M. N. Razavi and L. Iverson. A grounded theory of information sharing behavior in a personal learning space. In *CSCW '06: Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 459–468, 2006.
17. L. Richards. *Handling Qualitative Data: A Practical Guide*. Sage Publications, 2007.
18. L. Richards and J. M. Morse. *Readme First for a User's Guide to Qualitative Methods*. Sage Publications, Thousand Oaks, 2007.
19. B. Salmon, F. Hady, and J. Melican. Learning to share: A study of sharing among home storage devices. Technical Report CMU-PDL-07-107, Carnegie Mellon University Parallel Data Lab, October 2007.
20. B. Salmon, S. W. Schlosser, L. F. Cranor, and G. R. Ganger. Perspective semantic data management for the home. In *7th USENIX Conference on File storage Technolgies (FAST'09)*, 2009.

21. S. Voida, W. K. Edwards, M. W. Newman, R. E. Grinter, and N. Ducheneaut. Share and share alike: exploring the user interface affordances of file sharing. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, 2006.