Masks
Designing for Privacy in Online Education
http://groups.ischool.berkeley.edu

Heather Ford
Alex Smolen

Advisors
Brian Carver
Deirdre Mulligan

MIMS 2011

UC Berkeley School of Information
Foreword

Do you wear a mask? Don’t we all?

Gaining trust from people is often a process of trying on different masks as we work out our roles in different social structures. There are times when we want to hide in a crowd and watch from afar; others when we present a side of ourselves that will be well accepted by others; and still others when we want to be free to explore parts of ourselves that are often hidden. On the Internet, we often work fastidiously to maintain these different masks – whether it is maintaining separate personas on Twitter and Facebook, or continuously grooming any links to our official identity from Google searches.

Despite these protective hacks, we are often forced to expose all early on in a relationship. This is especially apparent in online learning management systems, where students are encouraged to publicly document their learning experience and engage with others in a form that may be persistently available and outside their control.

Masks is an examination of how privacy and identity impact online education. There are three sections to this document. The first is a collaborative review of privacy for Peer-to-Peer University (Privacy and P2PU), but the second two are individual research projects that support the first.

Alex’s individual research (Privacy Design Analysis) focuses on designing for privacy, and how that could inform the design of online education systems. He looked at existing tools like privacy impact assessments, academic research about privacy design heuristics, privacy requirements gathering, and values in design decision-making strategies. With this in mind, he constructed a simple methodology and tool to think through the privacy impacts of a technology.
Heather’s research (The Spaces Between) focuses on norms of information flow in educational environments. She interviewed I School students to get a sense of how they manage personal information and identity in online and physical circumstances. She found the education doesn’t only happen in the classroom – more important are the “spaces in between”, like the hallways and the lounge. There, students can socially construct knowledge, and feel more comfortable with intimate environments outside the purview of authority figures and unknown audiences. For the most part, online education attempts to replicate the classroom, with instructor-led courses and public discussions, failing to recognize the important role that different social contexts play in knowledge creation.

We hope that our work is useful for others who are grappling with issues of privacy in the educational context. While the problem is complex, we feel that privacy compels us to stop and take stock of the real impact of technology on the educational future of so many and that it is useful not only for protecting students but for enabling more meaningful engagement in online educational communities.
# Table of Contents

Privacy and P2PU ........................................................................................................ 6
Introduction ............................................................................................................... 6
Background .............................................................................................................. 8
  Privacy in P2PU .................................................................................................... 8
  Negative Effects .................................................................................................. 11
  Privacy in Education ......................................................................................... 12
Approach .................................................................................................................. 13
  Contextual .......................................................................................................... 13
  Practical ............................................................................................................... 14
Privacy Design Analysis .......................................................................................... 15
  Step One: Determine Actors ........................................................................... 15
  Step Two: Identify data disclosed by actors ................................................... 17
  Step Three: Identify actor goals ....................................................................... 19
  Step Four: Flag problematic disclosures ....................................................... 22
  Step Five: List potential mitigations .............................................................. 24
  Step Six: Determine impacts of mitigations ................................................... 26
Design Alternatives ................................................................................................ 33
  Prototypes of Design Alternatives ................................................................. 34
Conclusion ............................................................................................................... 43
Privacy Design Analysis .......................................................................................... 45
Introduction ............................................................................................................ 45
Related Work ......................................................................................................... 45
Methodology .......................................................................................................... 48
In Practice ............................................................................................................... 51
Tool ........................................................................................................................ 52
  Requirements .................................................................................................... 52
  Implementation ................................................................................................. 54
Conclusion ............................................................................................................. 55
Privacy and P2PU

Introduction

On the 22nd of February 2011, P2PU co-founder Stian Haklev made the following proposal to the P2PU research community by:

I’m hereby proposing that P2PU officially decides that all user interaction data are completely public. This includes not only all the material that has been posted, but also when users access the websites, what they click on etc.

Haklev’s statement was followed by a lively debate about whether or not to proceed with the suggested functionality. One contributor noted that making every click public by default, in a way no other major site on the web did, could lead to a privacy backlash.

In adopting a “completely public” policy can you guarantee that no personal data could be inadvertently released due to an oversight or lack of a review process?

This kind of question is at the heart of current debates about privacy. Some people may see an act of information disclosure like this one as being acceptable and necessary for any kind of social web application, while others are may be adamantly opposed. Some may make bold statements about privacy being dead, and suggest that norms have changed, but privacy concerns seem to be omnipresent in the mainstream consciousness of the information age.

In some cases, the argument for disclosure and against privacy has obvious justifications. If a company is in the business of behavioral advertising, they have a vested interested in promoting the disclosure of information that helps inform them
of consumer behavior. If a government wants to oversee its citizens, it will argue for the safety benefits of surveillance technology. P2PU, in this case, wishes to enable greater access to its system by researchers in order to improve it, but in doing so, is P2PU favoring the interests of outside parties to the detriment of those central to its mission?

We’ve been interested in P2PU as a kind of experimental laboratory for our privacy research for the last semester. As suggested by the mailing list exchange above, and further discussions described below, there is a considerable tension between the values of openness and privacy when the openness principle is taken to its ultimate extreme. Accessibility and openness are central to P2PU’s founding and development, with nearly all exchanges and data on the site being publicly accessible and available under a Creative Commons license. But is the openness principle making students inadvertently feel as if they are being watched, either by an authority figure or an unknown audience? Is this one of the factors contributing to high dropout rates on P2PU? Is the lack of intimate spaces preventing the community building necessary for successful learning?

We wanted to determine what sort of privacy concerns users may have while engaging with P2PU, and to suggest a set of design alternatives that may help preserve user privacy while incorporating other goals of actors in the system. We started with a contextual analysis of why privacy is important for education and what norms of information disclosures support educational goals. Next, we performed an iteration of the privacy design analysis methodology using information gleaned from interviews and textual analysis of P2PU. This methodology is a structured way of looking at the different types of information disclosures that occur in the system in order to find potential privacy issues. The result of this process is a set of design alternatives that we propose the P2PU community consider.
Background

As background for our analysis, we describe P2PU stakeholders’ understandings of privacy, and put forward a cumulative definition of how privacy is broadly conceptualized. By rooting the analysis in the understandings of the community, we hope to show that P2PU has the seeds of the solution in their discussions on these issues and that we merely provide them with a framework for thinking systematically about what the solutions might be.

Privacy in P2PU

We asked one of the P2PU co-founders to tell us whether privacy was important to P2PU.

I think privacy is important. Period. Not just for P2PU. I think the individual should be in full control about the information they want to share about themselves and their activities. For us, things like selling off peoples’ email addresses that would be absolutely inconceivable at P2PU because we would feel like we’re taking peoples’ personal information and using it in a way... I mean we could obviously have these very sneaky end user agreements where you never actually know what you’re clicking on and what you’re accepting, but we could do all that stuff. But that would be totally counter to the way that our community works where everything is based on trust and if you start abusing that trust we will lose our community and we will lose P2PU. So for us as a project it’s important that people feel that we’re treating their information with respect and care. At the same time, we are in a fairly good position because there isn’t actually that much information that we have that not everyone else has access to already.

Privacy, because it is a complex concept to understand, is often described in this fairly limited way. In this case, the P2PU co-founder describes privacy in terms of very specific types of information – in this case, email addresses – rather than tied to specific social interactions. Another preconception about privacy is that merely notifying users of what information they intend to make public can solve the problem. In the discussion about whether to enable interaction data to be released to the public, Haklev said that one concern could be that “students have not been
properly informed about this.” He said that they could mitigate this by providing “clear information about this in non-legalese (to) be displayed and agreed to when students apply to courses” and provided the example:

I realize that all my interactions with the P2PU site will be logged, and made publicly available. All my contributions will also be licensed under a CC license.

The discussion that followed indicated that the community collectively understood the privacy problem much more broadly. Some pointed out the problem was that the data at stake was not data that is currently available but clickstream data, the traces that a user leaves when interacting with the system. It would be difficult to conceptualize this data, particularly because it is invisible to most users. Only administrators see how long a user interacts with the P2PU website, which pages they visits and where they are physically located when they log into the system, for example. When some community members indicated concern, others kept returning to the current open data on the site, believing that what was being proposed was little more than developing an accessible data file of what was currently available. According to this line of thought, if some people were concerned with this, then they should be concerned with all the other interactional data that was being released. As one community member wrote:

These are open forums already. Everything is already being tracked (try and search Google for a phrase from the forum) and can be analyzed if anyone cares.

Haklev pointed out a second concern in the forum. He said that that doing may “stop people from participating, it would negatively affect learning processes, it would give us a bad reputation.” Responding to a request for exactly what “unintended consequences” might result from aggregating user data and making clickstream data available to the public, one community member wrote that:
The unintentional consequences I think largely involve social factors like privacy, trust, encouraging small group learning and formation, and simple risk mitigation.

This got community members thinking – not just about the proposed initiative but the effects of the current open system and ways to mitigate some of the harmful effects that might be taking place:

I would say that, quite arguably, any chilling effect from "publicity" is already fully in effect. However, people are free to sign up with a pseudonym, and most appear to do that. P2PU could offer a sign-up service that offered to more fully "mask" users online, and I think this would be a good idea. It could also offer a sign-up service that found a way to associate real-world identity to someone (e.g. send them a thumbdrive with a given cryptographic key on it) and I think that could also be a good idea.

This exchange showed that the discussion was driving towards a sophisticated understanding of privacy. Starting off with a very limited view of privacy and relating to specific types of information, the community was starting to understand privacy as tied to specific roles and relationships. We approach privacy from this viewpoint – not as a feature of certain types of information, but as the ability to control information, or who knows what about you (Livingstone, 2006). In doing so, we understand that students must and do disclose personal information in order to sustain intimacy in the learning environment, but they wish to be in control of how they manage this disclosure. From this vantage point, P2PU-ers started to question the open-everything perspective, asking themselves:

Personally I think the question shouldn't be "should the data be open?", but rather, "what data, precisely, should be open, and why?"

The community was clearly grappling with the question of how they can remain true to their identity as an organization committed to openness and accessibility while still respecting the privacy needs of their users. But they needed to know what were
the negative effects of making all data open to the public and how to mitigate the effects of such openness. A balance needed to be struck but they seemed to not have the tools to strike that balance.

**Negative Effects**

We wanted to know whether students believed the current system to be adequate for their privacy needs. In looking at the system, we identified features that may indicate a need for greater privacy. In one course that we looked at, the course organizer asked students to keep a learning journal to reflect on what they were learning:

One of the main expectations for this course is that participants will keep a journal throughout the 8-week period. This journal can be a blog, or it can be a private journal - the purpose is to encourage you to begin and maintain a reflective practice, and it is up to you whether or not you would like to make your writing public.

Only four out of the forty students responded to the course organizer's post. One student responded with a link to a blog that he had set up specifically for the course, another linked to his personal blog in which he had made a category specifically for the course, yet another wrote: “I have a private journal that I keep and if it is okay I will use that” with a fourth (the only woman among the four) saying that she would “use a private journal until I catch up and can figure out how to blog.” This interaction indicates the need for private, or at least more intimate spaces controlled by students outside of what was being offered by the open P2PU platform. As the course organizer indicated in an interview:

Overall, I think that the more open people are, the more they are able to contribute and the more valuable discussion and exchange of ideas becomes. However, personal (private) reflections help participants to make connections and understand the material, and I think this is critical.
**Privacy in Education**

These personal reflections that the course organizer highlighted also reflected a need for a variety of interaction spaces. These interactions could be more useful and common if they were shared with a few trusted peers, or only with the course organizer. In Goffman’s language, personal reflection enables us to “try on” a diverse range of “costumes” or identities in order to choose who we want to be and which communities we want to join (Lave and Wenger). In the formative stages of the learning process it is important that “audiences” are limited in order for the student to feel safe enough to experiment.

The need for safe spaces for experimentation in education is echoed by our ethnographic research of a graduate school in the United States to understand their relevance to learning and privacy. In this study, we find that the “spaces between” – the classroom before and after class, the student lounge and computer laboratory – play an important role in the learning experience because it is where students can construct knowledge with their peers and practice the performance of new identities. The fact that these spaces are located outside the purview of those in authority and that they enable students to choose who they can be intimate with is consequently central to the success of these spaces. When private spaces are unavailable, students “hack” the system by opening up their own private channels resulting in harms including exclusion, identity crises and self-censorship.

In a similar way, we can see the example of the learning journal as a way in which P2PU users need to hack the system for privacy reasons. Without the technology to provide more intimate spaces, users turn to blogs, systems that enable only completely public viewing or completely private accessibility, thus excluding the opportunity for more intimate social spaces. In addition, the use of private blogs means that those who don’t have the know-how about how to set up those blogs are excluded, forcing them to use analogue technology when what they want to do is be able to be part of the learning community. As the female student wrote about her
personal diary, "It is only temporary until I figure this out and get caught up. I do want to share and be a part of it all."

We wanted to enable a system on P2PU that didn’t exclude students in this way. We wanted them to be able to meet the needs of more intimate spaces while retaining their goals of openness and accessibility. To do this, we needed a more systematic, transparent approach.

**Approach**

There are a variety of approaches for identifying and describing privacy concerns in a system. They range from the privacy impact assessment used by the federal government to sophisticated privacy modeling languages. As part of our research, we developed a methodology and tool for eliciting privacy design issues and mitigations. For P2PU, we were wanted to ensure our analysis met three criteria: that it was a) contextual b) holistic and c) practical.

**Contextual**

Although innovative online systems like P2PU are new and change the power dynamics of learning in fundamental ways, they still fall within the context of education. Helen Nissenbaum’s theory of “contextual integrity” suggests that privacy is about preserving norms of information disclosure. This proved useful in understanding the role of contexts – what Nissenbaum calls “structured social settings” – in designing for privacy. Nissenbaum believes that we need to look to formally established spheres (in this case, the tertiary educational institution) in order to understand the key characteristics of its corresponding socio-technical systems and practices (in this case, the P2PU system) (Nissenbaum, 2010). In addition to being a site based on educational goals, P2PU also has functions similar to a social networking site, and more broadly is a site on the World Wide Web. Understanding these multiple contexts of P2PU helps us establish the norms governing information flow, a central tenet of our understanding of privacy.
Holistic
The predominant approach for dealing with privacy has, until recently, been the “notice and consent” model. Companies tell users what they do with their information and the user makes a choice about whether to use the system or not. But this is not always an effective approach since people rarely read these privacy policies. If P2PU were simply to update its privacy policies to include information disclosures that its users may find objectionable, it would not be serving its users, since these statements may go unnoticed.

Our methodology looks at a comprehensive set of disclosures and potential design alternatives. It considers information disclosures that are problematic not just from a legal viewpoint, or from a security viewpoint, but from a broader social perspective. Finally, and perhaps most importantly, our analysis is not solely focused on achieving privacy goals. It also designed to respect the other goals of P2PU stakeholders. When a design alternative represents a tradeoff in these goals, then at the very least the costs should be well understood and transparent. This is critical in building privacy solutions that are grounded and relevant to the system’s priorities.

Practical
We believe that much debate about privacy occurs in silos. Privacy advocates talk in their circles about their dissatisfaction with current systems in terms of privacy needs, while system designers often find their recommendations irrelevant and out of touch. Within organizations, the person responsible for privacy may reside in the legal department and not share a common language with the person responsible for implementing the technology. We wanted to develop a way to bridge the gap between theoretical privacy discussions and real-world system implementation. This tool attempts to bridge that gap with policy-focused input and technical design outputs.
Privacy Design Analysis

To analyze P2PU privacy and generate a set of design alternatives, we turn to the aforementioned privacy design analysis methodology. This section discusses the results of that methodology applied to P2PU.

Step One: Determine Actors

The first challenge in determining potential privacy issues in P2PU is to understand the actors and their roles in the system. Understanding the different actors helps elicit who may be disclosing data, and to whom they may be disclosing it. Correctly aggregating groups of actors can be somewhat difficult because there may be slight differences between actor types who may or may not be material in the analysis. Generally, if groups have significantly different goals, or there are power differentials between two groups, then they should be considered separately.

In the context of P2PU, the following groups of actors were identified:

Users

Users in P2PU are people who register for the site and contribute, both as course facilitators and as students. In initial runs of this analysis, these two groups were considered separately, but it soon became apparent that their privacy concerns and goals largely overlapped. Since this is a peer-to-peer model of education, there isn’t the same power differential between educators and learners. People who participate in classes are encouraged to lead classes, emphasizing the fluidity of the educational roles and the unity of purpose between the two. Some users may participate and lead a large number of courses, while others may register but not actively participate in courses. However, since they have identified themselves to the system, they are no longer considered the “public”. Additionally, some P2PU organizers may be users, but their goals in that context are different.
Users are the group whose privacy concerns are explored in our analysis. Generally, users are the ones who disclose data to the system and other actors are those who receive data from the system.

**Organizers**

Organizers are the administrators, developers, and backers of the P2PU project. These are people who are stakeholders in P2PU’s success and contribute to the project in some material way. They are distinct from users because they are responsible for defining the P2PU architecture and practices, and are in many cases privy to the data submitted by users. Similar to users, there are many different “types” of organizational actors, such as developers, administrators, and evangelists, but they generally share the same set of concerns about P2PU.

**Researchers**

In the context of P2PU, researchers are the people who intend to use data from P2PU in order to create and publish research about online education. Since P2PU is an “open” educational environment, it represents a fertile ground for data about how students engage in online education. This data is useful for researchers who hope to understand and improve online education system including P2PU. Researchers tend to belong to educational institutions, and may interact with organizers through the P2PU researcher mailing list.

**Public**

The public actor is defined as the people browsing the Web who visit P2PU. They may want to browse a course anonymously, or they may be looking for information about a user. Anything disclosed on the Web is available to the public, and may be available to entities such as search engines and web archives that store and index publicly available material.
Others actors not included
In the context of defining the relevant P2PU actors, several actors were considered but ultimately excluded because they didn't directly affect the educational disclosure relationships that we were primarily concerned with. These included third parties like governments and ISPs as well as local network users who may be able to view users traffic. While a more extensive privacy analysis may consider these actors, and the potential privacy threats they pose, these concerns are not especially specific to P2PU and there judged to be less important in the analysis.

Step Two: Identify data disclosed by actors
The next step in the process is to identify the sensitive data that is disclosed by the actors in the system. In the case of P2PU, the set of sensitive data is composed only of data that is disclosed by users. While presumably organizers, researchers, and the public disclose some form of data to the P2PU system, none of this data has the same sensitivity as user data. This is primarily due to the structure of the power relationships – organizers create the rules of the system and have access to all (or most) of the data, researchers don’t necessarily disclose any information directly to the system, and the public do not directly identify themselves to the system, so any data they leave behind is mostly anonymous. Users, on the other hand, disclose significant amounts of data to the system. This data is grouped into a set of categories to ease the analysis process – if data has the same general set of concerns, then it makes more sense to treat it as a single type of data. The following categories of data were identified:

Profile information
Profile information includes user name, full name, email address, location, bio, and profile image. All of these pieces of data may be used to uniquely identify or otherwise describe a user in real life. One piece of data not included in this category is the password that should never be disclosed to anyone, under any circumstances,
and its disclosure represents a pure security concern as opposed to a privacy concern.

**Social connections**
Social connections are tracked in P2PU similarly to other social networking and media services, where users can follow and be followed by other users. This information includes both the number of followers and followees, as well as the list of users who are following and followed. Users in P2PU are enrolled in courses that they “follow” so the social connections includes information about courses that users in P2PU are involved in (either as course organizer or student).

**Posts**
Posts are messages that students publish as members of a course. They may be in response to a task, and are typically a response to some sort of question. In the previous architecture of P2PU, this functionality was structured using a forum-based interface, while in the current interface it is based on a threaded comment format. This is the primary place for learners to interact as a group on the P2PU platform.

**Private messages**
In addition to posting on course pages, users can send private messages to each other. These messages work similarly to Facebook messages or Twitter direct messages, which are understood to be private except to those whom they are addressed to.

**Clickstream activity**
Clickstream activity is the set of data that users disclose as they use the P2PU site. This includes pages visited, time that pages were visited, login and logout activity,
and other information about a user’s session. Anything that ends up in a web server log could potentially be part of the clickstream activity data set.

**Step Three: Identify actor goals**

The next step is to identify the goals of the users of the P2PU system. In many cases accessing the publicly available documentation from the P2PU project websites and mailing lists enabled us to fairly easily identify these goals. In other cases, our understanding was shaped by interviews that we conducted with P2PU staff and organizers. The following set of goals attempts to list the major concerns, even when these goals are somewhat abstract. In the case where goals are conflicting, they can be broken down into sub-goals. This decomposition proved unnecessary in this case.

**Users**

*Learn about subjects from peer contributions*

The primary goal for P2PU students and coordinators is to gain competency in particular subject areas, and secondarily to broaden their social network. According to P2PU co-founder, Phillip Schmidt, users were initially equally interested in the idea of peer to peer learning as they were in learning, whereas now, ‘people are coming to the site specifically because they are much more interested in actually learning something’. Users therefore see P2PU as a way to get an education and broaden their qualifications.

*Organize courses that are engaging and enable relationships to be formed*

In addition to learning, some users want to organize courses for other users. Because the majority of course organizers are not paid by P2PU, their motivations are largely philanthropic. For this reason, a key goal for course organizers is to run a course that is engaging and that elicits responses from students. Many course organizers are doing this work because they are passionate about a particular topic
and want that passion to transfer to students. When students drop out or don’t respond they may become frustrated or disconcerted.

Avoid revealing embarrassing or otherwise harmful information online
Privacy concerns are prevalent among users of social networking services. Students have specific concerns regarding how potential employers and respected peers view their competency in educational settings. Also, since P2PU has a privacy policy that dictates what it can and can’t do with users’ information, users will certainly have an expectation that their information will be used in a way that doesn’t violate their privacy. This goal is central to the analysis: if no one had the goal of protecting their privacy, then the analysis wouldn’t yield any results.

Organizers

Create an open environment for learning in order to enable participation, accessibility, innovation and accountability.

P2PU describes openness as its primary core value, followed by community and peer learning. Openness is not an end in itself for P2PU; openness is seen as a way to enable participation, accessibility, innovation and accountability and it is these sub-goals that we need to be aware of when assessing the impact of information sharing on these goals.

Bring learning materials to as many people and to as diverse of a range of people as possible.

P2PU wants to enable broad access to learning. This is demonstrated in its “learning for everyone, by everyone” tag line. According to Phillip Schmidt, P2PU should become useful to a more diverse audience. “I think we need to be much more useful for people who don’t live on the web already; who don’t use these tools everyday; who are not geeks already (and geeks here in the broad sense – not just technology, you know: learning geeks). People who would choose p2pu as an opportunity to learn as an alternative existing formal education. So people who would usually go to
formal education, we should be able to give value to them so that they can learn with other people in an open way.” Thus, diversity is as important a goal as increased numbers of students for P2PU.

Ensure learning material is high quality
P2PU organizers want to ensure that the learning material on P2PU is high quality. This is specifically mentioned on the web page, which suggests, “P2PU enables high-quality low-cost education opportunities”.

Support research that leads to better understanding of open educational resources and P2PU specifically
P2PU “strongly supports both research activities under the auspices of P2PU as well as research activities about P2PU” and priorities research that “drive(s) understanding and lead(s) to action”. According to the organization, “we want research to improve and increase open social learning opportunities. We support research that leads to better understanding of the overall OER ecosystem, but we have a strong interest, and can commit more resources to research that also directly helps P2PU.”

Research community
Conduct research and experiments that provide insight and guidance
Researchers who use P2PU for research material are intending to create and publish research about online education that leads to increased insight and advancements in their field.

Follow ethical guidelines
According to the P2PU research guidelines, “[r]esearchers agree to be respectful of the privacy and time of everyone involved, especially individual users who are approached for data collection.”
Public

Benefit from courses without directly participating
Members of the public may visit and use P2PU in order to learn about subjects. The assumption here is that others learn from watching the interaction as well as reading the syllabus and materials. According to Phillip Schmidt, "[W]e don't know if they're learning anything but we do know that a lot of people click on the courses who are not in the courses. If you look at the course participation even within the people who sign up for the course many more sign up than actually participate but still some of them log into the site occasionally."

Step Four: Flag problematic disclosures
After identifying the set actors, goals, and data, the next step is to cross-reference all of the possible disclosures and determine which disclosure may potentially violate expectations of privacy.

The process of determining which disclosures are problematic isn't concretely defined. First, we considered if the disclosure occurred in the system. So, for instance, while it would be a privacy problem if users’ passwords were published on the Internet, P2PU didn’t support this disclosure. Second, we considered whether or not a disclosure was consistent with information disclosure norms. To determine if it was consistent, we contrasted disclosures with existing social networks like Facebook and Twitter as well as thinking about norms of information flow in the educational context. So, for instance, when people are on Twitter, they are publishing information to be consumed by others. However, when they interact on P2PU, they may be more interested in figuring out the answer to a question. Their goals and expectations about how that data will be disclosed will differ.

Out of the set of possible disclosures, the following four types of disclosures were identified as potentially problematic:
Users disclosing profile information to the public

Profile information includes identifying information like email address and full name along with biographical information, a profile picture, and links to user’s websites (along with automatically polled content from any RSS feeds). While most of this information is optional, a user may disclose this information without realizing that it will be disclosed publicly, and if archived or indexed may exist permanently.

A similar disclosure that demonstrates why this may be a privacy concern is when Facebook decided to make certain aspects of a profile public. People assumed that their detailed profile information would only be available to friends, but it was now available to anyone online. Users were upset, and Facebook changed some defaults to make profiles more private by default.

Users disclosing social connections to the public

In addition to profile information, the user’s social connections are displayed publicly. This is not optional and a user has no notice that this information is made public. In the educational context, social connections could be seen as potentially sensitive if students have few friends, or if they are friends with those in authority. There have also been cases of using social graphs to predict sensitive information about users (link to studies). Additionally, Google Buzz revealed users’ social connections by making followers out of people that users emailed. This violated people’s sense of privacy, even though no message contents were disclosed.

Users disclosing posts to public

The fact that users disclose their posts to the public is a central tenet to P2PU’s open architecture. However, users may feel uncomfortable posting in an educational context if all of the posts are publicly available, linked to their public profile, and archived permanently. Clearly there is some form of tension here, as users may not
necessarily feel a privacy violation when their posts are publicized, but may not feel comfortable posting in the first place. Even if they feel they must post, they may not include any sensitive information (for example, if they're struggling with material) that would help to bring about social cohesion necessary for the growth of community and trust.

**Users disclosing clickstream activity to research community**

P2PU has discussed the concept of releasing clickstream data for all users to the P2PU research community. This would allow the research community to perform a variety of experiments with the data, but it may also lead to problematic disclosures since users are unlikely to realize that their data will be used in this way. Privacy debacles have occurred in both the AOL query research and the Netflix recommendation research when user data was made available to a broad research community, even when that data had been anonymized. The fact that this “invisible” data is being released to third parties may also produce a chilling effect on social interactions among peers who feel they cannot use the platform for social exchanges. This may decrease intimacy and sense of solidarity that motivates students to continue the learning process in order to engage in a process of “becoming” (Lave and Wenger).

**Step Five: List potential mitigations**

The disclosures that lead to a sense of violated privacy in P2PU can be augmented with controls that either change user expectations or reduce the impact of the disclosure. In this step, each disclosure was considered against a set of potential design alternatives that would mitigate the harm of the disclosure. These potential disclosures include preventing the information from being disclosed at all (don’t disclose), removing the identifying information from the data so that it can’t be traced back to an individual user (anonymize data), allowing users to provide different identifying information that may or may not be tied to their real identity (allow pseudonyms), aggregating data so that a single user’s data is not separable
(aggregate data), providing notice to the user about the disclosure before it occurs (provide notice), and allowing the user to choose whether or not the data is disclosed (offer choice).

This list of mitigations is simply a common set of design choices for preserving user privacy. Certainly, there are design choices outside of this list that may be appropriate for certain problematic disclosures. There was no hard and fast rule that made us consider only these mitigations. For instance, there could be contractual obligations between actors to use data ethically, or more exotic technical concepts like differential privacy could be used. However, we were able to find some commonly used design alternatives that effectively addressed these problematic disclosures.

Each disclosure was examined against these alternatives, and the set of potentially applicable alternatives are discussed below:

**Users disclosing profile information to public**

*Don't disclose:* Profile information could be forever hidden from the public. This is similar to some forums and online communities, where you must be a member to view other users’ data.

*Provide notice:* P2PU could provide more explicit information about what data will be shared publicly.

*Offer choice:* Users could have some form of control over what profile information is shared publicly.

**Users disclosing social connections to public**

*Don't disclose:* Social connections could be hidden from the public.

*Aggregate data:* Rather than displaying a user’s lists of friends, P2PU could only display the total numbers of connections.
Provide notice: P2PU could let users know that their social connections will be displayed to the public.

Offer choice: P2PU could offer users the option to show or hide their social connections from the public Web.

Users disclosing posts to public
Don’t disclose: All posts could be hidden from the public until they signed up for P2PU.
Anonymize data: Posts could be published anonymously, allowing users to comment without disclosing their identity.
Allow pseudonyms: Posts could be published pseudonymously, allowing users to selectively disclose their identity.
Offer choice: Some posts could be made private or published to a selected audience, rather than public only.

Users disclosing clickstream activity to research community
Don’t disclose: Clickstream data could be withheld from the research community.
Anonymize data: Anonymized clickstream data (stripped of identifying information) could be made available to the research community.
Aggregate data: Aggregated clickstream data could be made available to the research community.
Provide notice: Users could be given notice that their clickstream data is used for research purposes.
Offer choice: Users could be given the opportunity to choose whether or not their clickstream data is captured and used for research purposes.

Step Six: Determine impacts of mitigations
In this step, we list the mitigation, then the list of goals that are supported or harmed by the mitigation.
Profile information

Only display users’ profile information to other users

Supported: (Users) Avoid revealing embarrassing or otherwise harmful information online

Harmed: (Organizers) Create an open environment for learning in order to enable participation, accessibility, innovation and accountability

Harmed: (Public) Benefit from courses without directly participating

If users don’t disclose their profile information to the public, then they will be less likely to reveal harmful information. But if they want to share their learning profile with others who aren’t registered P2PU users, then they may be less motivated to contribute. Also, if users are unable to disclose their profile information publicly, even if they want to, it impacts the ability of P2PU to foster openness. Learning about users may help the public understand how a user’s identity affects their comments but it is unlikely that (unless they are a researcher, covered separately below) they will be able to make anything more than superficial judgments about the content or the community from reading profile pages.

Offer choice when users disclose profile information to public

Supported: (Users) Avoid revealing embarrassing or otherwise harmful information online

If users are given the choice of whether to reveal their profile publicly, they are given more control over their information, potentially lessening the privacy harms. Users may be more willing to share more sensitive information (for example, their learning styles, what they struggle with, what they excel at), which could advance social cohesion.

Provide notice when users disclose profile information to public
Supported: (Users) Avoid revealing embarrassing or otherwise harmful information online

If users are made aware that their profile information will be made public, they will be more likely to avoid revealing embarrassing or otherwise harmful information on P2PU.

Social connections

Only disclose social connections to other users

Supported: (Users) Avoid revealing embarrassing or otherwise harmful information online

Harmed: (Public) Benefit from courses without directly participating

Harmed: (Organizers) Create an open environment for learning in order to enable participation, accessibility, innovation and accountability

If this information isn’t disclosed, it lessens the choice of a problematic disclosure. However, like preventing disclosure of profile information, not revealing the social graph impacts the public’s ability to access a learner’s social connections.

Aggregate data when users disclose social connections to public

Supported: (Users) Avoid revealing embarrassing or otherwise harmful information online

Harmed: (Public) Benefit from courses without directly participating

Harmed: (Organizers) Create an open environment for learning

Instead of displaying complete “connections” lists, P2PU could only display the number of connections. If social connection data is aggregated, it has the same (although somewhat lessened) impacts as not disclosing the data at all.
Provide notice when users disclose social connections to public
Supported: (Users) Avoid revealing embarrassing or otherwise harmful information online

If users are made aware that their social connection data will be made public, they won't be surprised to find it available. The data will still be available to the public.

Offer choice when users disclose social connections to public
Supported: (Users) Avoid revealing embarrassing or otherwise harmful information online
Harmed: (P2PU) Create an open environment for learning

If users are given the choice as to whether or not to disclose social connections, their ability to control their data is supported. However, some may feel that this inhibits the community since it is useful to know who is connected to whom in the group. If a user chooses the hide their social connections, then no one will see the data, which would suggest that the public's goal of learning from others' material, is neither supported nor harmed.

Posts

Don’t disclose when users disclose posts to public
Supported: (Users) Avoid revealing embarrassing or otherwise harmful information online
Supported: (Users) Learn about subjects from peer contributions
Harmed: (Public) Benefit from courses without directly participating
Harmed: (Organizers) Create an open environment for learning
Harmed: (Organizers) Bring learning material to as many people as possible
If users’ posts are only displayed to “members” of P2PU, then privacy goals are supported, and users may feel more comfortable contributing to courses, which supports their learning goals. However, the ability of the public to learn from P2PU, as well as the organizational goals of openness and accessibility, may be harmed by the lack of disclosure since not everyone can access users’ posts.

**Anonymize data when users disclose posts to public**

*Supported:* (Users) Avoid revealing embarrassing or otherwise harmful information online

*Harmed:* (Users) Organize courses that are compelling and informative

*Harmed:* (Organizers) Ensure learning material is high quality

If users are allowed to comment anonymously, then they may be more likely to provide truthful comments. However, the communal learning goals may be harmed since anonymous comments may have the effect of devaluing the comments, leading to “trolling”. The ability to maintain high quality content, important for both course facilitators and P2PU organizers, may thus be harmed if anonymous comments are enabled.

**Allow pseudonyms when users disclose posts to public**

*Supported:* (Users) Avoid revealing embarrassing or otherwise harmful information online

*Supported:* (Users) Learn about subjects from peer contributions

If users can employ pseudonyms in addition to their primary identity, their privacy will be protected. This may also help them learn by encouraging them to contribute without making them feel like their contributions will be permanently associated with a single official identity. Since pseudonyms are semi-persistent, they can still be used to accrue a reputation and have been effective in many online communities as a compromise between unique persistent identities and complete anonymity.
Offer choice when users disclose posts to public
Supported: Learn about subjects from peer contributions
Supported: Avoid revealing embarrassing or otherwise harmful information online
Harmed: Benefit from courses without directly participating

Users who are able to choose whether or not their posts are publicly available may be able to learn more effectively since they will feel more comfortable contributing. They may be able to ask questions or post information that would be embarrassing if posted publicly. However, the public’s ability to access courses may be harmed, since some interactions would take place in a back channel.

Clickstream data
Don't disclose when users disclose clickstream activity to research community
Supported: (Users) Avoid revealing embarrassing or otherwise harmful information online
Harmed: (Users) Conduct research and experiments that provide insight and guidance

If users’ clickstream data is never disclosed, then they are at less risk of problematic disclosure. However, the researchers’ capabilities for providing insight on the community are harmed since clickstream data could prove useful in analyzing learning patterns that could improve learning.

Anonymize data when users disclose clickstream activity to research community
Supported: (Users) Avoid revealing embarrassing or otherwise harmful information online
Supported: Follow ethical guidelines
Harmed: (Researchers) Conduct research and experiments that provide insight and guidance
If clickstream data is anonymized, it may help to protect users’ privacy and researchers follow ethical guidelines, although it is important to note that anonymized data may, in many cases, be re-identifiable. Additionally, the anonymized data may not be as useful for researchers, since it may not be as accurate.

**Aggregate data when users disclose clickstream activity to research community**

*Supported:* (Users) Avoid revealing embarrassing or otherwise harmful information online  
*Supported:* Follow ethical guidelines  
*Harmed:* (Researchers) Conduct research and experiments that provide insight and guidance

Similarly, if clickstream data is aggregated then users’ privacy is protected and researchers ethical goals are preserved, but it may impact the researchers’ ability to effectively analyze the data and reach conclusions about user behavior.

**Provide notice when users disclose clickstream activity to research community**

*Supported:* Conduct research and experiments that provide insight and guidance  
*Supported:* Follow ethical guidelines  
*Harmed:* Avoid revealing embarrassing or otherwise harmful information online  
*Harmed:* Create an open environment for learning in order to enable participation, accessibility, innovation and accountability

If users are given notice that their clickstream data is being used for research purposes, but are obligated to provide it, then researchers’ goals are served but users may either feel as though sensitive information is being shared
inappropriately or that they are unable to effectively participate in P2PU because they feel uncomfortable disclosing the data.

**Offer choice when users disclose clickstream activity to research community**

*Supported:* Avoid revealing embarrassing or otherwise harmful information online

*Supported:* Create an open environment for learning in order to enable participation, accessibility, innovation and accountability

*Supported:* Follow ethical guidelines

*Harmed:* (Researchers) Conduct research and experiments that provide insight and guidance

If users are simply given the choice as to whether or not to contribute their clickstream data, then users’ privacy is protected and the “openness” of the community is upheld. Researchers goals may be somewhat harmed because of the bias introduced by self-selecting participants, but subject choice is commonly a requirement of research ethics guidelines.

**Design Alternatives**

After we analyzed how each mitigation impacts goals, we chose which ones represented the best compromise between stakeholders’ interests. The goal of the previous steps is not to arrive at a conclusion for what to build. Instead, it is meant to guide the analyst through the process of enumerating potential privacy concerns and to understand the ramifications of design choices. There are several factors that could influence the choice of design alternatives. For example, actors who are most vulnerable in the system should have their goals prioritized. This helps preserve the ultimate goal of the initiative, which is to serve its users. This is encompassed by the broader ethical goal of user empowerment. Also, not all of the alternatives are exclusive – some may be useful to consider in tandem. For instance, notice and choice are commonly used for privacy settings in applications.
Offer choice to protect users disclosing profile information to public
A user should be able to choose what information in their profile is disclosed to the public, since this information may be indexed by search engines, archived permanently, and otherwise available to an extremely large invisible audience.

Provide notice to protect users disclosing social connections to public
A user should be made aware that their social connections, including enrolled courses, are publicly archived and accessible.

Allow pseudonyms to protect users disclosing posts to the public
Users should be able to use pseudonyms while interacting on the P2PU system. This will allow more freedom in the learning environment and support users’ privacy goals while upholding the ultimate goals of an open system.

Offer choice to protect users disclosing posts to the public
Users should be able to choose whether or not their posts are accessible publicly or only available to a select group of users. This would facilitate private conversations among users as well as public conversations.

Offer choice to protect users disclosing clickstream activity to research community
Clickstream activity disclosure should be a configurable feature. This would allow users to choose whether or not they want their clickstream activity disclosed, and make it apparent to users when their clickstream activity is being recorded, and what data is in their clickstream activity.

Prototypes of Design Alternatives
After identifying the set of design alternatives, and considering them in the light of education privacy research, we designed the following set of prototypes:
Identity Selector

The “Identity Selector” is a feature that supports pseudonyms in interactions on P2PU. Rather than having a single identity that is used for all interactions, users can create multiple identities, or pseudonyms, and choose which identity to use for each interaction. This is a balance between a single, persistent identity and a user being anonymous. A user can explore different identities freely without worrying about a single “mask”, but each persona is tractable and can accrue reputation.

Figure 1: Identity selector

In the first screenshot, there is a new dropdown list in the upper right corner where a user can select an identity. In this case, the identity selected is “alsmola”.
By selecting the identity list, the user can choose from a list of “masks”, including the possibility of being anonymous. The masks described here are helpful for showing why someone might want different identities. Someone who is learning or teaching material that could be considered “pirating” may want to keep their behavior disassociated with their other identity, while someone who is learning a new subject (e.g. “python n00b”) may also want to keep a separate face to the world. The option of anonymity may or may not be a desirable feature. It could be excluded if it was seen to encourage bad behavior.
Once the user switches to a new identity, s/he has a new set of social connections, a new status feed, and so on. From the outside, this is no different than a new user, but the fact that it is supported by the interface encourages students to create new identities and use them to take on new challenges that they may otherwise be intimated by.

**Clickstream recording**

Rather than silently disclosing all of the clickstream data to researchers, P2PU should allow users to “opt-in” to clickstream recording and notify them when their clickstream is being recorded. Furthermore, users should be allowed to see what data is being submitted in a format that is readable and comprehensible.
Figure 4: Clickstream recording

In the first screen, a new privacy section is available in the profile editor. One of these options is the ability to opt-in (or opt-out, depending on the configuration) to the clickstream data program. Another option is to hide profile information from the public. This is not part of this prototype, but a simple design alternative identified in the process above that would fit here well. An additional option could be used to hide social connections.
If the user chooses to enable clickstream recording, they will see a recording option in the upper right corner of their screen. This serves two purposes: first, users will be given notice that their actions are being recorded, and second, it will give them the option to enable and disable recording.
Figure 6: Clickstream recording

The user can disable the clickstream recording, similar to going into browser “incognito” mode, through the interface.
Figure 7: Clickstream recording

The user can also view all of the data recorded about them in a clickstream data viewer, which would show them in a clear format what data would be available to researchers. This is similar to a browser history mechanism. The user could also delete data if they didn’t want it to be submitted for research.

**Private discussion**

Users should have the option of participating in discussions that do not take place in public. This allows a greater sense of freedom while preserving some of the openness of the community by also allowing public interactions.
Figure 8: Private conversations

This prototype supplants the existing private message feature, which is one-to-one, with a multi-party private conversation feature. The user could select a group of students to invite to a back channel conversation.
Figure 9: Private conversations

The user will then be able to engage in conversation with a select audience, out of the view of the public and other students.

**Conclusion**

We believe the result of this analysis is a thoroughly considered set of recommended features for P2PU to support user privacy and broader goals of delivering education to a wide audience. The privacy design analysis methodology is a structured way of identifying privacy concerns and design alternatives that offers a glimpse into the effects of data disclosure that may not have been apparent at first glance. The resulting design prototypes represent a justified set of features that help protect user privacy while supporting the goals of the other actors in the system.
While we realize that not all of the features are likely to be acceptable in their current form, but hope that at the least they will inspire thinking about how privacy concerns could be addressed and that aspects of them find their way to P2PU.
Privacy Design Analysis

Introduction

While many people have privacy concerns, defining what privacy means and whether or not a system violates user privacy is a confounding task. There are many conceptions of privacy that are, upon analysis, an oversimplification. For instance, some have suggested that in the modern era “privacy is dead” – that all, or most, of our personal information should be available online to everyone. Other notions of privacy see a dichotomy between information that is public and information that is private, with a bright line barrier between them. Privacy theorists have established a more nuanced and comprehensive view of privacy (Dourish & Anderson, 2006; Nissenbaum, 2004). In this view, privacy is both contextual and experiential. A privacy violation occurs when information flows in a way that violates peoples’ expectations based upon established social norms.

People are troubled by threats to a widely recognized expectation of privacy, yet there are few tools available to technologists who would want to consider the privacy impact of their innovations. Existing approaches, like privacy impact assessments, tend to rely on coarse distinctions between private and public information and offer little guidance to the assessor for surfacing issues – especially when this requires creativity and antagonistic thinking. The following discussion of a new methodology and tool for privacy analysis builds on existing approaches but is more methodical and encourages a comprehensive review of potential violations.

Related Work

An effective and practical method for identifying and responding to privacy issues in technology as it is being built would be a boon to technologists. Companies have been forced to make hasty changes to product features and issue apologetic press releases when their products violate privacy norms. The FTC has recently discussed
their motivation to prevent companies from tracking consumers without their consent\(^1\). The growth of the social web has seen the rise of new companies that are built on the private information of others. A methodology to quickly describe and qualify privacy threats could facilitate privacy by design, where technology is built to respect personal information disclosure practices.

One approach to privacy analysis in somewhat common usage is the privacy impact assessment. Privacy impact assessments are structured analyses of how information systems handle personal data used primarily by the US Federal Government since becoming a requirement in the E-Government Act of 2002\(^2\). Based on Fair Information Practice Principles\(^3\) (FIPPs), privacy impact assessments encourage minimizing the disclosure and storage of sensitive information. While privacy impact assessments are a useful tool for high-level compliance goals, they do not provide guidance to the assessor about what types of information may be considered sensitive beyond personal data, nor how different design alternatives may lead to more privacy-preserving systems that achieve the same goals. They are also insensitive to context – the same information in two different systems is treated identically. While the overall goals of privacy impact assessment are useful, they may not be detailed or contextual enough to inspire “privacy by design”.

There are three lines of academic privacy research that inform this methodology. The first, based primarily on HCI research, focuses on heuristics – questions or principles that are meant to guide the expert reviewer through the process of performing a privacy review (Bellotti & Sellen, 1993; Hong, Ng, Lederer, & Landay, 2004; Langheinrich, 2001). The second, stemming from the field of requirements engineering, emphasizes the goals of stakeholders in the system, and how those

\(^1\) [http://www.ftc.gov/opa/2010/12/privacyreport.shtm](http://www.ftc.gov/opa/2010/12/privacyreport.shtm)

\(^2\) [http://www.whitehouse.gov/omb/memoranda_m03-22](http://www.whitehouse.gov/omb/memoranda_m03-22)

\(^3\) [http://www.ftc.gov/reports/privacy3/fairinfo.shtm](http://www.ftc.gov/reports/privacy3/fairinfo.shtm)
goals might be impacted by disclosure and privacy concerns (Jensen, Tullio, Potts, & Mynatt, n.d.; Kalloniatis, Kavakli, & Gritzalis, 2008; Yu & Cysneiros, 2002). The third area “values in design” work that consider how technology can be responsive to ethical, social, and other human concerns (Collins & Miller, 1992; Cranor & Reagle, 1998).

There is a canon of academic research that can be labeled “heuristic” because, like privacy impact assessments, it requires the analyst to answer a set of questions or consider a set of principles with respect to a system handling sensitive data. For example, Belotti et al. ask the analyst to focus questions categorized according to a framework of capture, construction, accessibility, and purposes of information disclosure to identify privacy concerns in ubiquitous systems. Lederer et al. identify five pitfalls for those looking to design privacy-preserving systems, including obscuring potential information flow and inhibiting existing practice.

A different area of academic research focuses on privacy through the lens of requirements engineering, and draws upon the concept of “goals” to guide and justify privacy trade-offs. Yu et al. set out to identify privacy amongst a set of “soft goals” for a system, and derive a set of agent-goal tuples that can be referenced in design. Jensen et al.’s STRAP framework describes a set of actors, each with a goal tree, and identifies obstacles to privacy goals as vulnerabilities. To identify problematic disclosures, STRAP also includes questions distilled from earlier heuristic models.

Several authors have looked at privacy as a value to incorporate in the design process. Designers of the Privacy Preferences Project (P3P) have discussed the difficulty of creating a “social protocol” that takes into account values of different stakeholders (Cranor & Reagle, 1998). One work that inspires the privacy design analysis methodology is Paramedic Ethics for Computer Professionals (Collins & Miller, 1992). The paramedics ethics framework is a useful tool for quickly understanding who the different stakeholders are in a particular system and
how the different alternatives of a decision affect their inter-dependencies but its use is limited to contexts in which an individual is attempting to make an ethical decision among alternatives. It is not focused on privacy, and it is not designed to elicit privacy issues in a system.

This methodology is also inspired by techniques from the field of security threat modeling (Swiderski & Snyder, 2004). Threat modeling is a structured method of identifying a set of security threats to a system by decomposing the application and analyzing each component against categories of security breaches. Whereas threat modeling discusses attackers and vulnerabilities, the methodology described in this paper focuses on actors and disclosures. There have been attempts to use threat modeling in the context of privacy assessments (Deng, Wuyts, Scandariato, Preneel, & Joosen, 2010), but these approaches haven’t combined other insights, such as goal analysis and norms-based evaluation.

The underlying theory of privacy emphasized here is Nissenbaum’s privacy as “contextual integrity”, where privacy is defined by both distribution and appropriateness norms. Privacy threats occur when personal information is disclosed in a context that violates norms or expectations. In order to identify these threats systematically, we need to first identify the set of information disclosure contexts and then identify contexts that potentially transgress norms.

**Methodology**

The first step of the methodology is to identify the key actors in the system. This should include everyone who can interact with the system, either by contributing or receiving data. User models, interviews, ethnography, or other user research may help define who the actors are and establish more contextual information that will be useful in the following steps.
After the set of actors has been defined, the next step is to determine the data that they disclose to the system. For a social networking site, this could include explicit data disclosure such as user-generated content and profile information as well as implicit data disclosure such as browsing history, transaction history, or social ties. Examining the technical architecture of the system, including the database schema, logs, and network traffic patterns, may help collect this set of data. It may also be helpful to review system documentation. Additionally, interviews with technical personnel may elicit information about data processed by the system.

Each piece of data from an actor may be available to different actors, and so the next step is to list the entire set of possible disclosures. For example, in a social networking site a user may share photos through a photo upload feature. Other users may be able to view the photos in a photo gallery. An administrator of the site may be able to review the photos. Those photos may be available to the public with other data, such as the geographic location where the photo was taken or comments about the photo by the person who took it.

Once the set of disclosures has been established, each information flow must be examined for its underlying threat to privacy. Determining if an information path represents a privacy threat at first seems difficult as users’ expectations of privacy vary. However, there are several techniques from existing literature on privacy analysis that suggest a set of heuristics that aid identifying privacy threats. The analyst should ask questions like:

- Are there existing technologies that have similar transactions?
- Is this transaction occurring with the actor’s knowledge?
- Did the user consent to this transaction?
- Can the user prevent this transaction from occurring?
- Is this transaction something that most users would find acceptable?
- How could a vindictive or malicious actor abuse this transaction?

Answering these questions requires an understanding of the context of the information flow in question, so interviewing users, reviewing forums, and
participating in mailing lists may help identify privacy concerns that have been expressed in the past in this system or similar systems.

If the result includes any questionable uses, the transaction should be flagged for later review. The result of this process should be a list of disclosures (defined as data given by one actor and received by another) that may be objectionable. This is the critical step of the process because it is where privacy issues are spotted, and because all possible disclosures are examined it should give some confidence that the list of issues is comprehensive.

Once the set of potentially problematic disclosures is identified, the analyst can consider a set of mitigations, which are design alternatives that reduce the privacy risk of a disclosure. For instance, the user could be given notice about the disclosure, or they could have the choice to opt-in or opt-out. The data could be anonymized before it is disclosed, or aggregated by combining individual records. These mitigations may not completely prevent a problematic disclosure, and they are also not mutually exclusive.

In order to identify the best alternatives, the next step is to determine which alternatives support and harm actor goals. In the case where estimation is exceptionally difficult, analysts could conduct a survey or interview of potential users to attempt to determine and quantify their reaction to the transaction. The priority of these goals may be based on values held by the actors in the system, or focused on preserving the rights of the most vulnerable actors in the system. The results of this goal-based analysis highlights which design alternatives best protect privacy while preserving other user goals.

The privacy design analysis methodology that the tool is based on can be operationalized in a highly structured process, which makes it amenable to capturing in a code-driven tool. At a high level, the methodology consists of the
following steps:

- Determine actors of system
  - For each actor in the system with relevant privacy concerns:
    - Determine potentially sensitive data that actor discloses to other actors
    - For each piece of data:
      - Determine which actors and context (purpose) that data flows to:
      - For each context, flag for potential privacy violations (using norms)
  - Determine goals and sub-goals of actors (individuals + organizations) in the system
  - For each potentially violating context:
    - For each privacy-preserving design option:
      - For each goal:
        - Determine whether option supports or harms goal
        - If option supports important goals and harm is minimized, flag

- Based on important of goals, rank design options by most and least feasible
- Decide on which design options are most appropriate

The methodology is also shown in a simplified flowchart form in the appendix.

In Practice

The purpose of the methodology described above is to surface privacy issues, but the design and development of an information system takes place in a broader context. How would this methodology fit into a product development life cycle, and who would be responsible for using it?

While many organizations have employees dedicated to privacy, their roles are often distant from the technologists who create products that affect privacy concerns. Many times, the responsibility for privacy is tied closely with the legal department, since the economic damages from privacy violations are often born as legal costs. However, this isolation is problematic since the design decisions made by product managers and developers have serious ramifications, and purely legal
privacy analysis may protect the organization from legal threats but is unlikely to consider the contextual integrity theory of privacy and how it may be threatened by the implementation of the system. This methodology is meant to "bridge the gap" between privacy-aware analysts, who understand the nuanced nature of privacy and prevailing norms, and technical system designers who understand the capabilities of the system and the information flows it supports. This is analogous to the role that threat modeling plays in a security context, where security expertise and technical comprehension are both key ingredients.

The privacy design analysis methodology is meant to surface privacy concerns and design alternatives for a system as a snapshot in time. Therefore, it should represent a living document that may change over time. For instance, if the architecture of the application changes, if new features are added, then the analysis should be updated. Similarly, as disclosure norms evolve, what does and what doesn’t constitute a privacy violation may change. Thus, the analysis from this methodology should be visited iteratively during a product development life cycle, so that changes can be addressed.

**Tool**

The methodology described above is designed to be algorithmic in its approach, and repeatable across a wide variety of contexts. To this end, the Privacy Design Analyzer tool was created as an embodiment of the methodology to support the analysis process and guide the analyst. At each step, this tool prompts the user to input the data required to fulfill the next step, and then uses it as part of the next prompt. The tool stores the data in a persistent format and allows the analyst to edit and revise the data entered in the process.

**Requirements**

The Privacy Design Analyzer tool has a small set of requirements, since it is only a prototype “proof on concept” project. However, these requirements may be
expanded to create a more robust tool capable of supporting multiple users and more complicated forms of analysis.

_System should support multiple applications._
Initial versions of the tool did not support multiple applications. Since the tool is useful for a variety of applications, functionality that allows user to create, update, and delete applications, each of which has their own set of actors, data, and goals, is a design requirement.

_System should allow users to enter actors, data, and goals into the system._
The requirement is that the user of the tool should be able to create, update, and delete actors, data, and goals into the system. These pieces of data will be used subsequently for analysis.

_System should allow users to select a set of disclosures._
The system allows users to select harmful disclosures from the list of all possible disclosures (defined by the cross-product of all actors, their data, and other actors). This list is used to flag issues for later analysis.

_System should allow users to select possible mitigations (design alternatives) for each disclosure._
The system supports selecting from a list of mitigations. While the list is hard-coded, it may be wise to allow custom options or additional information.

_System should allow users to select which goals impact each mitigation._
For some mitigations, goals will be supported, and for others, they will be harmed. The system needs to collect information from the user about which alternatives support and harm each goal.

_Display list of requirements, prioritized by goal._
The system needs to show the user a report of the design alternatives that best
support user goals.

*Allow user to export data.*

At the end of the analysis, the user may want to export the data for uses in other systems. This tool supports a JSON export format.

*Save data persistently.*

In initial versions of the tool, the data was stored in the users session cookie. This was not a persistent storage mechanism, and was scrapped for a database backed data store.

**Implementation**

The application is web-based, and is built using the Python language, the Flask web micro-framework, and the SQLAlchemy Object Relational Mapping framework, with a JQuery-based front-end. It employs a standard MVC style architecture, and has a fairly simple model with only a few database tables that have simple relationships. The code takes advantage of the Flask templating syntax to create chunks of reusable presentation logic. Each template (view) contains the presentation format, and is rendered at runtime with data from the model. The code for the Privacy Design Analyzer is available on github⁴ and a running version is available on the author’s personal site⁵. According to github, the code is 79% Python and 21% JavaScript.

The tool is generally a multi-step “wizard”. At the end of the run of the tool, the user has a set of design alternatives that support the goals of identified actors in the system. These design alternatives represent starting points for prototyping and implementing interfaces, which upon completion can be further refined by

---

⁴ https://github.com/alsmola/privacydesignanalyzer
⁵ http://alexsmolen.com/pda/
additional runs through the tool.

Future versions of the tool will include functionality to support multiple users – this will require login and registration functionality. This is important so that the tool can be released to a broader audience online, and users can create applications without displaying their information to everyone else. The current suggestion is for a user to clone the source code and run the tool locally – something that may be difficult for novice users.

**Conclusion**

Identifying privacy problems inherent in a design is a difficult issue for system designers in the real world. There are few tools for rigorously and comprehensively finding what aspects of a technology may violate users’ expectations about the disclosure of their personal information, and how these issues can be addressed. The framework presented here differs from other approaches by providing a highly structured methodology for looking at privacy as a consequence of design. The end goal is to elicit the entire set of information disclosure instances that may be problematic from a privacy standpoint, find the best privacy-preserving designs, and encourage system designers to design features in ways that are more sensitive to users’ privacy concerns.

The Privacy Design Analyzer is a helpful proof-of-concept to demonstrate how technology can be used to help operationalize a methodology that otherwise must be manually performed. The technological itself isn’t novel, but it is novel to embody a methodology for examining privacy into a computer-assisted format. It is intended to be useful for technologists designing systems with potential privacy issues to surface these issues and identify alternatives that are privacy-preserving while still incorporating other goals and values of users of the system.
Appendix

Screenshots

Figure 10: Screenshot showing actors entered into the system
Figure 11: Screenshot showing mitigations (privacy preserving design alternatives) for each disclosure
### Privacy Design Analyzer - P2PU

<table>
<thead>
<tr>
<th>Start</th>
<th>Actors</th>
<th>Goals</th>
<th>Data</th>
<th>Disclosures</th>
<th>Mitigations</th>
<th>Impacts</th>
<th>Result</th>
</tr>
</thead>
</table>

#### The following mitigations support goals:

- Use offer choice to protect when users discloses profile information to public
- Use provide notice to protect when users discloses profile information to public
- Use provide notice to protect when users discloses social connections to public
- Use allow pseudonyms to protect when users discloses forum posts to public
- Use offer choice to protect when users discloses clickstream activity to research community

#### The following mitigations do not affect or have mixed effects on goals:

- Use offer choice to protect when users discloses social connections to public
- Use don't disclose to protect when users discloses clickstream activity to research community
- Use anonymize data to protect when users discloses clickstream activity to research community
- Use aggregate data to protect when users discloses clickstream activity to research community
- Use offer choice to protect when users discloses forum posts to public

*Figure 12: Screenshot showing results*
Figure 13: Flowchart of methodology
The Spaces Between: Towards Private Spaces for Peer Learning

Abstract I analyze the spaces of a graduate school in the United States to understand their relevance to learning and privacy. I conclude that the “spaces between” - the classroom before and after class, the student lounge and computer laboratory - play an important role in the learning experience because it is here where students can construct knowledge with their peers and practice the performance of their new identities. The fact that these spaces are located outside the purview of those in authority and that they enable students to choose who they can be intimate with is consequently central to the success of these spaces. In contrast, private digital spaces are unavailable to students, with the result that students attempted to use spaces like Facebook to engage with one another resulting in harms including exclusion, identity crises and self-censorship. Online-only educational spaces seem to replicate only the classroom space but without the protective walls available in conventional learning environments. I conclude with three design implications, calling for online educational practitioners to build for intimacy, to avoid "invisible audiences" and to make sure that students have an online space of their own that is outside the purview of authority.
Introduction

Just before the dot-com bust of 2001, John Seely-Brown and Paul Duguid published their seminal critique of the hype surrounding the information age. While pundits predicted that information technology would lead to the end of the press, of politics and of the university amongst others, Brown and Duguid believed that this "tunnel vision" was stopping us from looking at where technology is really taking us. In their final chapter, entitled “Re-education”, the authors lamented the design of online educational systems, declaring that even though the conventional system was flawed, it at least provided space for authentic learning to "leak" through:

From their old to their new guise, the conventional schools continue to regard learning as a matter of delivery. As we have argued, it is not. The old system, however, with its campuses and classrooms, nonetheless provided plenty of opportunities for students to "steal" knowledge from teachers and construct knowledge with peers, even though that was not necessarily the intention of the design. The new system, however, designs these opportunities out. The secure, narrow channels of digital technology make theft and collective construction much harder. Putting the pedagogy of the conventional classroom on line, that is, cuts out the hidden resourcefulness of the conventional classroom. (Brown & Duguid, 2000, p. 233)

The conventional classroom, I soon learned, was useful not only because of what happened within it during class time but because it was flanked by supportive spaces in which peers were using the in-class experience as fodder for learning from one another. Inspired by Helen Nissenbaum's theory of "contextual integrity" (Nissenbaum, 2004), I set about investigating the norms of information flow that governed the spoken and unspoken in these spaces.

What I learned was that these “spaces between” - the classroom before and after class, the school corridors, the student lounge and computer laboratory - revealed attributes that were specifically related to privacy. Although they were located within the remit of the institution, these spaces were outside the direct purview of authority, they were self-governing and were flexible enough for students to be able to choose with whom they wanted to share the space and how they wanted to
engage with others in the space. The private features of these "spaces between" became the vehicle for the learning that was taking place inside them. Having control over these spaces so that students could control who knew what about them enabled students specifically to join new "communities of practice" (Lave & Etienne Wenger, 1991), to experiment with new identities and to practice performances that would take place during dedicated public events.

In contrast, no such space private existed in the digital realm. Student mailing lists were surveilled by staff who felt it their duty to watch the lists in order to protect students, and the class blogging systems were largely open and visible to the public Internet. In the absence of the student-lounge equivalent in the digital realm, students tried to squash uniquely-shaped educational relationship pegs into the round- holed platforms like Facebook, leading to harms relating to exclusion, identity crises and self censorship.

Two trends in educational technology add to these problems. The first is the unbridled enthusiasm that surrounds learning analytics that track a learner's every interaction in the name of “personalized learning”. The second is the increased use of social networking services in blended and online educational platforms that encourage students to share more and more personal information in the name of 'social constructivism'. The result is a lack of online spaces in which peers can learn with one another outside of the control of those who could exert power over them.

I begin by discussing my methodology for this study, I go on to outlining theories of learning and social constructivism and how they are being adopted selectively by educational technologists. This is followed by a day-in-the-life story of learning at the I School, after which I analyze learning through the lens of Brown and Duguid’s "old" and the "new" systems. I conclude with an analysis of space, identity and privacy and the implications for the design of networked learning systems.
Methodology

In order to study the social life of information at a school, I wanted to get a context-rich understanding of information sharing in a traditional (face-to-face) educational experience. I had a hunch that people were self-censoring in online educational environments and so I chose a face-to-face environment in which I had direct access to the social, informal interactions of its participants. This required me to hone my observation skills in an environment in which I was also a participant.

Studies like Sharon Traweek's *Beamtimes and Lifetimes: The World of High Energy Physicists* (Traweek, 1992), were particularly instructive in this regard. Traweek, who conducted fieldwork at three national high particle physics laboratories over a period of five years, first worked at one of the labs explaining its activities to visitors. Said Traweek, “[a]s Public Information Officer I stood outside social divisions and yet was a familiar part of the lab” (Traweek, 1992, p. 12). She indicates that the fieldworker needs to remain marginal - that its socio-cultural assumptions would no longer stand out in the foreground of her attention if she were to become a fully integrated participant in the community (Traweek, 1992, p. 9). On the other hand, says Traweek, the very nature of participant observation requires the fieldworker to “live her days and weeks and months within the patterns of the community's life, moving in spaces shaped by the community taking part in its activities on its own terms” (Traweek, 1992, p. 10).

I chose to study students at my own school because only then would I be able, with limited time, to gain insight into the lived experiences of graduate students, the ways in which they share information with one another, as well as the ways in which information-sharing revealed deeper understandings of identity and privacy. In order to maintain some distance, I chose to study three masters students in the class adjacent to mine. Each class is very much their own “community” and as a

---

6 I use Traweek’s definition of ‘community’ as ‘a group of people who have a shared past, hope to have a
member of the I School community but separate from the 2010 class, I believe that I was able to retain my marginal status.

Grounded theory proved a useful methodology for this study. An iterative methodology that emphasizes the generation of theory from data in the process of conducting research, grounded theory enables one to stay close to the lived experience of the community, building levels of abstraction directly from the data, developing theories iteratively and then gathering further data to check and refine the emerging analytic categories (Charmaz, 2006). It is important to note here that my initial observations came out of a class in which we were experimenting with social media technologies in the traditional classroom. A particular discussion initiated by one student about “how open should this site be?” seemed illustrative of so many of the questions that we had been asking, and I started theorizing from there about what our direct experience with the artifacts meant for larger questions facing the educational technology and open education movements.

I went on to conduct four hour-long open-ended interviews with graduate students about their interaction with social media tools, their online identities and information-sharing habits. I chose students who seemed either very thoughtful about their presentation of self on the school website, on the one hand, or who were known to be very open in class and outside of class on the other. I also interviewed two staff members who were most influential in mediating the identity of students within and outside of the school and noted my observations in classes that I am taking or acting as a student instructor in, as well as observations that I have done in other areas of the school (for example, in the student lounge and workshop/co-lab space). I also interviewed a co-founder of Peer-to-Peer University (P2PU) and an organizer of one of the P2PU classes. I then coded those interviews and

shared future, have some means of acquiring new members, and have some means of recognizing and maintaining differences between themselves and other communities' (p6).
observations, and then conducted a second round of interviews with the students to check the analytic categories that I had developed. Because my sample is small, this should be seen as exploratory research. My question sets appear in the appendix.

**The rise of social constructivism and the privacy challenge**

The School of Information is housed in the oldest building on the UC Berkeley campus. Built in 1873, this stately red-bricked building, rumored to have been the location for the rooftop scene of Mary Poppins\(^7\), stands in the center of campus opposite Sather Tower. A photograph taken in 1898 shows one of the lecture rooms with about a hundred smartly dressed students in seating that is raked upwards and backwards towards the end of the room. Women in pillbox hats sit on the right of the room and men on the left, with the professor, a fuzzily bearded man in a black suit and a pocket handkerchief, standing between the blackboard and the class.

This room hasn’t changed drastically in the past hundred years. It is no longer raked upwards; chairs are probably more comfortable these days and the blackboard has been replaced with a whiteboard flanked by two screens. Classes are also much smaller -- never exceeding about forty students, and male students no longer wear suit and a pocket handkerchief, standing between the blackboard and the class.

---

\(^7\) The rumor is, unfortunately, false.
suits to class. Female students, at least in this classroom, have decreased since 1898 to about 40% of the student population.

This room may not have changed much, but all the other rooms in the school are equipped with reconfigurable tables and chairs to support greater flexibility in class configuration and learning styles. The CoLab on the second floor is a flexible space with a lounge area as well as movable tables, chairs, and whiteboards, to support informal meetings, group work, and individual study and the other classrooms can be configured in U-shapes or in-the-round to support "seminar-style" classes.

The move from fixed, lecture-style architecture to a more configurable one reflects the rise of social constructivist theories of learning in the time between these two photographs. Rejecting earlier theories of learning as a process of individual, passive assimilation, John Dewey pointed out that education is entirely a social enterprise.

Education, in its broadest sense, is the means of this social continuity of life. Every one of the constituent elements of a social group, in a modern city as in a savage tribe, is born immature, helpless, without language, beliefs, ideas, or social standards. Each individual, each unit who is the carrier of the life-experience of his group, in time passes away. Yet the life of the group goes on. The primary ineluctable facts of the birth and death of each one of the constituent members in a social group determine the necessity of education (Dewey, 1930, p. 3).

Dewey declared that the school was a community and that it made no sense to have students isolated at their own desks or forbidden from talking with one another. He

---

8 When the doors opened in 1873, 167 men and 222 women students enrolled

http://sunsite.berkeley.edu/uchistory/general_history/campuses/ucb/overview.html
advocated for a focus on purposeful activity in social settings and said that the best way to learn a new idea was in "normal communication with others".

Almost a hundred years later, we're still struggling to understand how to put the social constructivist ideal into practice. The devil, it seems, is in the details. According to Etienne Wenger, the structure of education today assumes that learning “has a beginning and an end; that it is best separated from the rest of our activities; and that it is the result of teaching” (Etienne Wenger, 2007, p. 3). Working with anthropologist, Jean Lave who investigated learning communities in countries like Brazil, Liberia and the United States Brown and Duguid found that learning is what we do in the practice of everyday life and that it occurs through a process of social participation within a community of practice.

Communities of practice are formed by people who engage in a process of collective learning in a shared domain of human endeavor: a tribe learning to survive, a band of artists seeking new forms of expression, a group of engineers working on similar problems, a clique of pupils defining their identity in the school, a network of surgeons exploring novel techniques, a gathering of first-time managers helping each other cope. In a nutshell, communities of practice are groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly (Etienne Wenger, 2007).

Brown and Duguid, working at Xerox PARC in the 90s and tasked with “crack(ing) the learning problem” (Kumar & Ramsey, 2008), drew from the work of Lave and Wenger on learning as participation in "communities of practice" to declare that successful learning happens with and through other people and that what we choose to learn depends on who we are, who we want to become, what we care about, and which communities we want to join.

Although learning is social, remind Brown and Duguid, doesn’t mean that networked technology is always highly compatible with constructivist theories of
learning. In an important article entitled “Organizational learning and communities-of-practice: Toward a unified view of working, learning, and innovation” (Brown & Duguid, 1991) Brown and Duguid argue against the assumption that, given the “right” medium people will exchange information freely because they say this overlooks the way in which certain socio-economic groups, organizations, and in particular, corporations, implicitly treat information as a commodity to be hoarded and exchanged.

The constructivist ideal has been embraced nearly wholesale by many educational technology companies who often claim – without any apparent opposition – that their application of social media technology is in keeping with the constructivist tradition. The past few years has seen a mushrooming of platforms providing social media tools (blogs, forums, wikis and profiles) to both formal and information educational initiatives that purport to embrace social constructivist theories. Providing little more than social-media enhanced content management systems that were built for contexts outside of the education one, such platforms purport to enable the building of “vibrant online communities”, “social collaboration across campuses” with results that include “increased student engagement, improved faculty productivity, enhanced online learning and broader marketing outreach”.

One company, that bills itself as “the community platform for education”, declares that they “are pioneering new ways of enabling social learning pedagogies in an online environment” and that “the concept of social learning is rooted in the notion that knowledge is socially constructed and that learning is best achieved through dialogue, collaboration, and transparent sharing of information across a community. A tool need only carry the term “social”, it seems, to be included in the latest set of educational technology or learning management systems. Learning may very well

---

9 http://www.goingon.com/
10 Ibid.
be social, and social media tools may very well enable people to engage socially, but does it necessarily follow that all social media tools enable social learning, and if this is the case, what kind learning results?

The use of social network tools in the educational environment introduces new audiences and new contexts that reconfigure the social environment in complex ways. Anthropologist, danah boyd (2010), writing about dynamics that shape “networked publics” discovers that such spaces are different because of three primary dynamics. Firstly, networked publics introduce invisible audiences that are not always visible or co-present when a person is contributing online, secondly, they result in collapsed contexts where the lack of spatial, social and temporal boundaries make it difficult to maintain distinct social contexts and thirdly, there is a blurring of public and private where “[w]ithout control over context, public and private become meaningless binaries, are scaled in new ways, and are difficult to maintain as distinct” (boyd in Papacharissi, 2010).

Social learning environments highlight this problem. Peer to Peer University (P2PU), for example, is an online-only tertiary education platform that attempts to use the affordances of the Internet to connect learners with one another. Co-founder, Phillip Schmidt, said that they started P2PU because “we were frustrated that there was all this amazing content, you could connect to people online but there wasn’t that social aspect to education that the universities were able to provide and there wasn’t that certification aspect that the universities are providing. So that’s what we set out to do.” Inspired by Richard Light’s Harvard Assessment Study in which he found that the number one indicator for academic success of Harvard students is their ability to form or join study groups, P2PU set about linking free and open educational materials with the students who wanted to learn about particular subjects together.

By using P2PU, users must license all their interactions under a Creative Commons Attribution Share-Alike license that allows anyone to adapt anything on the site and
publish it elsewhere. In a recent discussion on the P2PU research community mailing list, one of the co-founders made a proposal to make a data dump of the entire P2PU system, including clickstream data, available to the public under the same Creative Commons license. When one rather alarmed community member responded that no one “in government, business, academia, non-profits do anything like this” and that they should “proceed cautiously”, the person proposing the project retorted that “nobody else does this, just like almost nobody makes all student interactions open to the public, or offer free classes, or share material using open licenses”.

Discussions about new online educational initiatives often take this path. We think that we’re doing something completely different and so we have to chart new territories and make new rules. But Helen Nissenbaum reminds us that, “although contexts are tuned specifically to the task of framing a response to socio-technical systems and practices that have radically altered information flows, most of their key characteristics can be roughly matched with corresponding characteristics of their more formally developed counterparts” (Nissenbaum, 2009, p. 132). Although innovative online systems like P2PU are new and change the power dynamics of learning in fundamental ways, they still fall within the context of education. Nissenbaum’s theory of “contextual integrity” proves useful in understanding the role of contexts – what Nissenbaum calls “structured social settings” – in designing for parallel social settings on the Internet. Nissenbaum believes that we need to look to formally established spheres (in this case, the tertiary educational institution) in order to understand the key characteristics of its corresponding socio-technical systems and practices (in this case, the P2PU system).

11 View the discussion at https://groups.google.com/group/p2pu-researches/browse_thread/thread/32c8eadda3fc9488?pli=1
P2PU wants peers to be social with one another but has built an entirely new architecture to support this. If students set off to find their new school, they may be surprised to find a video-recorded disco instead. The disco is great for a while: there are fancy suits, no hierarchy and you can leave whenever you want, but this becomes its key problem: there is no one to impress, no markers to highlight who are the access points to communities of practice and no private spaces in which to practice with your peers before you perform in front of those who will give you such access. If in a conventional school information flows like a classical symphony where some parts of soft, some parts are loud, some parts are fast, other parts are slow, the P2PU disco makes information flow at the same tempo and at forte throughout.

In boyd’s language, contexts are collapsed that make spatial, social and temporal boundaries difficult to maintain. You seem to be my “buddy” now but might you be in control of a course I’m taking in the future? There are always "invisible audiences" present as we're interacting (future employers may be watching, now or in the future), with the result that there is a blurring of the public and private. In effect, there are no private spaces – no student lounge that students can return to in order to vent their frustrations, console one another or just “practice” before performing for the crowds that matter later on.

Defining these private spaces, then, becomes the key challenge for online social learning initiatives. Privacy in this sense becomes an important vehicle of the properties of these “spaces between”, enabling students control over information that could be used to make decisions against them in the future and for setting up protective boundaries necessary for certain interactions to take place. Privacy, we learn here, is not just important in order to protect students from harmful decisions, it is also critical to engaged learning among peers, whether offline or online.

I now turn to a real-life example of a day in the life of a first year masters student as she navigates the mediated social spaces of the school.
A story of learning

Sharon tells me about her day.

I got here and went to [class] which was at 11. I sat through [class] and we presented our projects. And then afterwards, Anna asked if I was going to a talk at the law school at noon and I said, no, I have a fire to put out, the fire being my 276 assignment. So I went to the lounge and worked on my paper. So Mary was sitting here and we like to chat with one another when we see each other down there, so I sat next to her and I was working next to her and then Anna came and sat next to me. And then I feeling kind of stressed over the outline and so the two of them helped me go over how to outline.

Tell me more about that...

I had Mary look over what I had so far and she looked over my intro and said: oh, it needs more bullet points and so then Anna came and looked at it and then gave me her document to look at to see how she outlined it and then Mary found a website showing how I would outline. Because my tendency is to just write paragraphs instead of points and so they were trying to show me how to shorten it up and put a dot into it. And then Tracy [second year masters student], who was sitting (opposite) was empathizing and saying yeah, she doesn't remember learning how to outline and that she finds it very difficult to. So it was mostly talking about how I can go about coding.

And so Tracy was talking about having done it last year and...

Yeah... and that it was difficult but that it has actually helped her since she has kind of gotten better over the years... or I guess months, since she's been here.

And was that helpful?

Yeah. I thought that someday I to o could learn to code like a normal person. So yeah at some point I figured out how to do it by just looking at the examples and looking at how short Anna's sentences were.

And then I went to the [class] and I got there four minutes late because I think Stuart [guest lecturer] started at exactly 2.10 and I had to run to the bathroom... so I was slightly embarrassed but the class was half filled so I didn't feel too bad.

...
After class ended, Anna and I went up and introduced ourselves to him because Anna and I are working on a project together and [lecturer], a couple of weeks ago, came up to us and asked us what we were working on and then introduced us to Stuart. So we then talked to him.

And I ended up talking to Tristan after class, and asking him how I could get a job with the [company name redacted] after I was done and, I don't know, how do I become him pretty much.

And was he helpful?

He was really helpful and I was just asking him what he does now and I guess small chat.

Tell me more about this. Did you actually use those words: 'how I can become you'?

No, no... I asked how did you get involved in working for the [company name redacted]. or maybe how did you get involved with working for the [company name redacted]. And he told me he wrote a paper about [issue] and then I said oh, I'd be interested in working for them. And we kind of talked about that... And he said that if I was interested there was a happy hour tonight with [company names redacted]. And that he would send me the invite if I emailed him.

And then I gathered up my things and was going to leave and go work. But then I realized [guest lecturer] had a talk. And then Anna and I felt conflicted about what to do because we wanted to do work so we sat in the back and worked on our outline. But then [lecturer] sat behind us and so I felt uncomfortable doing work with him behind me. So I kind of paid attention, or half paid attention and dimmed my screen so you couldn't see. But then, when we left, when the talk was over, and Mitch left with us and he asked, 'Oh, how was working on 276?' And I thought great, if you saw, [lecturer] probably saw it. Maybe that's why he didn't say goodbye to me. How rude [of me]. Because I don't like it when other people do work in the presence of something else. And I try not to do it unless I think that its an emergency. And I felt like it was an emergency. I felt really bad the whole time.

Reflecting on Sharon's day at school, we can start to see how and where learning is being practiced and how this learning practice is being shaped by Brown and Duguid's "old" and "new" systems. While educators tend to talk about face-to-face
versus online education as two distinct systems, we find that networked publics are entering and leaving the school as those connected to those networks join and leave, creating an environment mediated by networked technology that draws in new publics, new communities and new ideas.

The "old" system

In discussing elements of the "old" system in Sharon's day, I analyze the key spaces in which Sharon engages directly, with her "whole person", in the learning practice. I'll start by describing the school's spaces and talking about what Sharon's choice of venues reflects. What followed in the opening example was a powerful peer learning experience in which Sharon learned practical ways of expressing her knowledge using the examples of individuals whom she trusted.

There are two main spaces reserved for students at the I School: the downstairs student lounge and the upstairs co-lab. The way in which these rooms are designed and the established norms of information flow that inhabit them make for very different kinds of interactions. The student lounge is on the ground floor of the school and houses student lockers and as such is the place where students generally "check in" to see who is around and where unplanned gatherings happen before a class, especially when students need to print documents in which case there will often be a bunch of students clustered around the printer laughing about whose paper is worse, how they struggled with a particular task or how they feel about the instructor. With a relatively low ceiling and many dividing walls, sound does not carry as well and as such it is used by students working on their own who often also want the company (and sometimes the assistance) of other students. In addition to students meeting at lockers, this room is also where snacks and fridges are located, making it a good place to talk to other students. Although mostly used by students, this small "kitchen" is also used by some staff and faculty members. As such, it is probably the space in the school with the most foot traffic and the most opportunities for interacting with the school community.
The co-lab, on the second floor, has a high ceiling and no dividing walls. Because of its size, and the availability of white boards and moveable furniture, it is generally used for group meetings and work rather than socializing. When students want to separate themselves from the rest of the room, they take the moveable white boards and "wall" themselves off. This is only effective in keeping group members focused and undistracted but also in letting others know that work is being done and that they shouldn't be interrupted. According to one student, this has less to do with them distracting him than him continuously looking up and speaking to people coming into the room. Other students use headphones to show that they are separating themselves audibly from those in the room. When Trudy, another masters student, wants to work, she goes downstairs to the student lounge and chooses the table at the far end of the inter-leading room since she won't get distracted by those entering the room on the other side. For others, a cue that someone wants some degree of quiet and doesn't want to participate in the talk will be that they are looking intently at the screen and not being distracted by others. In these ways, students can be "sociable", surrounded by people, and have them near for support, but still find ways to manipulate the architecture and benefit from the norms that enable students to do undistracted work on their own.

In the opening example, we see Sharon and her peers engaged in a peer learning experience. As Sharon progresses through her I School experience, she is collecting memories like this one that enable her to move further into the communities of practice (Lave and Wenger) she wishes to join. Learning, here, is occurring in what John Dewey termed "situations" in which Sharon and her fellow students are actively engaged. It is not the kind of "armchair" thing assumed by the "assimilation" perspective in which the student sits back to imbibe information. According to Dewey, "Upon this view, thinking, or knowledge-getting, is far from being the armchair thing it is often supposed to be. The reason it is not an armchair thing is that it is not an event going on exclusively within the cortex or cortex and vocal organs... Hands and feet, apparatus and appliances of all kinds are as much a part of
it as changes within the brain" (Dewey, 1930, p. 13). It is this "whole person" learning that we must look for as an example of engaged, successful learning practice.

**Freedom locating**

When Sharon went to the student lounge to work on her assignment, she was offered a great deal of choice of where to situate herself and with whom compared to class time in which she must compete with others for limited space. Sharon could have gone to work in the library where norms dictate quiet and where students go to do solitary work. She could also have gone to the co-lab but feels that it is "too loud" at this time of the semester where group projects are in full swing. Instead, she chose to go to the student lounge where she could still be among her peers (and possibly receive assistance and support from them) in a more intimate environment.

When Sharon saw friends who she trusted in the student lounge, she chose to locate herself near to them, increasing the opportunity for her to engage with them. Sharon sees Anna and Mary as her close friends at the I School. If she can't make a class, it is Anna and Mary (as well as a few others in her trusted network) who she asks for notes or for a summary of what she missed. When I asked her why she thinks she is friends with this particular group, Sharon said spatial and social proximity both play a role.

Maybe because we're women and we have similar personalities... We like to not procrastinate and... I don't know... I'm not sure. We're in serious relationships and have a life outside of the I School. And I'm a lot closer to Anna this semester and we're quite similar in our interests. I think it's a bit of work ethic plus personalities. We're all in the same classes and we're all sitting next to each other.

They sit next to one another in a lot of the same classes, but they all share a similar work ethic that makes collaboration (and encounters like this one) easier. And although they sit next to one another in class, it is these 'spaces between' that extend their friendship into the realm of peer learning. It is here in the student
lounge that Sharon has the opportunity to learn the "language, tools and documents" that Wenger declares are central to moving into communities of practice.

**Freedom to imitate**

Sharon started off insecure about her ability to perform a task that she felt others in the community understood ("coding" research, in this case). The interaction that followed shows how Sharon learns how to imitate a document form, “the outline”, in order for her to gain credence in her community of practice. "Practice", according to Wenger, includes "the language, tools, documents, images, symbols, well-defined roles, specified criteria, codified procedures, regulations, and contracts that various practices make explicit for a variety of purposes" (Wenger, 2007: 47). Sharon says that her “tendency is to just write paragraphs instead of points and so they were trying to show me how to shorten it up and put a dot into it”. She did this by looking at Anna and Mary's outline and imitating the style.

Imitation or what Stanford psychologist, Albert Bandura calls "modeling" is central to social learning.

Learning would be exceedingly laborious, not to mention hazardous, if people had to rely solely on the effects of their own actions to inform them of what to do. Fortunately, most human behavior is learned observationally through modeling: from observing others one forms an idea of how new behaviors are performed, and on later occasions this coded information serves as a guide for action. Because people can learn from example what to do, at least in approximate form, before performing any behavior, they are spared needless errors. (Bandura, 1977)

Although learning theorists like Bandura acknowledge the role of imitation in learning, students in traditional tertiary institutions are still judged on individual performance and imitation in many forms is strongly disapproved of or banned outright. For this reason, students may feel freer practicing this kind of legitimate modeling in spaces that they control. Recall Brown and Duguid discussing the traditional educational system, “[t]he old system... with its campuses and classrooms, nonetheless provided plenty of opportunities for students to ‘steal’
knowledge from teachers and construct knowledge with peers, even though that was not necessarily the intention of the design” (Brown and Duguid, 2000: 233).

**Identifying**

In the opening example, Sharon is offered support from an advanced peer who, with a few words, identified with her and provided her with an example of what she might become. When entering the learning environment, students will immediately begin by trying to understand their identity in relation to others in order to find their place. It turns out that finding one’s place is actually essential to the learning experience. According to Brown, “[w]hat we choose to learn depends on who we are, who we want to become, what we care about, and which communities we wish to join. In this frame, learning is also a matter of changing identity, not just acquiring knowledge” (Kumar & Ramsey, 2008).

Hearing Tracy (who was a year ahead of Sharon in the master’s program) tell her about her own problems with coding in the past, how she overcame them and how she learned that this practice was useful lent a sense of authority to the task and helped allay Sharon’s fears that this might prevent her from joining the community of those who could outline (“perhaps one day I too will be able to code like a normal person”).

This kind of supportive peer-learning community offers an important buffer to what can be a frightening process of going back to school. In 2005, Stephen Brookfield celebrated his thirty-fifth teaching year by completing the second edition of his book “The Skillful Teacher”. Brookfield’s insight into the minds of students as they encounter the learning experience is particularly helpful in understanding the vulnerable situation in which they find themselves.

Students frequently feel like imposters, believing they don’t deserve to be in the role of learner. They worry about committing cultural suicide as friends and family see them changing because of college. They often feel in limbo, that they are leaving old
ideas and capacities behind as they learn new knowledge, skills, and perspective. Sometimes it feels as if learning is calling on them to leave their own identities in the past. However, if they can find others with whom they can share these fears - a supportive peer-learning community -- many of their anxieties apparently become much less corrosive (Brookfield, 2000, p. 32).

Tracy and Sharon are a year apart and take very few classes and social events together. This space, then, provides an important opportunity for serendipitous peer learning between students who may not be part of the same intake.

**Rules of engagement**

After working on her assignment in the student lounge, Sharon attends a class on the first floor of the building. In class, Sharon engages with the instructor (and the class, by proxy) when she offers a story of her own experience related to the information he is relaying. This experience connects Sharon briefly to the community of practice because she is able to identify with the material and thus find a "way in". The opportunities for such engagements are few and far between in a traditional class like this one where forty students must compete to find their own ways in to the community. Sharon said that the first time she raised her hand in class was extremely difficult.

Oh, that was really hard. I did not like that. I'm a very quiet person and also I was raised to be quiet, and also it's weird to be in grad school where you're expected to do the readings, to be critical about it and to have a point. I don't think I'm at that point where I can do all three for all classes. So, in fact, last semester I didn't really talk in class. But I'm trying to be better this semester. I tend to talk when I really know the answer to something.

Trudy said that she actually has a physical reaction when she raises her hand.

It's something I still struggle with: my heart races, I am questioning whether it even makes sense. And sometimes I have to hear myself to understand what I'm saying. And I sometimes say something and it sounds total rubbish but I have to say it to understand it. And I have to do to learn. I still struggle with it. But once I get started it's
fine. I forget that I'm presenting, or I forget that I'm talking about something. But right before, I get a little shaky and nervous.

In the classroom, as opposed to the student lounge, many students seem to accept the role of "empty vessel". That is not to say that students are passive. As a group they recognize the value of lecturer engagement and chastise those who monopolize the lecturer's time. Here students are seen enforcing normative rules about when they can talk, when they should listen, and how to divide the attention fairly. Peter says that, even though he would prefer to ask a lot of questions, that he sometimes holds back in class because of such norms.

(T)he things that I know about I had to hold back. The things that I don't know about I've had to restrain myself from being too inquisitive and clarifying. In the past, I would dig deeper and deeper and deeper without restraint saying: well, that doesn't make any sense - can you explain it further? And that would drive everyone nuts. So I've learned to keep it to myself and ask somebody after class.

He says that other students have alerted him to the fact that he can monopolize engagements with the instructor sometimes. Peter appreciates that students at the I School have 'social skills' and that in class he receives cues about when he can engage.

They (other students) tell you - visual cues, audio cues: now's the time, now's not the time.

Peter has subsequently learned the rules of the community.

I frequently wait to see if anyone else raises their hand if I know the answer for sure and I've already responded a couple of times I just wait to see if anyone else has it, and if they do, cool, let 'em be - you know, let them have - not the 'glory' - but make sure they get their due because the professors are looking for that. And if no one else raises their hand, I'll slowly raise my hand and like: I know this one too. It's like a time-delayed Watson.

Alternatively, Trudy finds it difficult to speak in class because she feels that it requires an instant opinion and she takes time to digest things.

I don't know I need time to think about things. There are definitely people in our class - everybody knows who they are - who all you have to say is one sentence and they have opinions and they have
In the classroom, where a limited period of time as well as participation points introduce a competitive factor to student engagement, students like Trudy vacillate between giving up on ever participating in class and raising their hand on occasions in order to play by the institutional rules. While this may be instructive in learning to engage in larger groups, it becomes critical - especially for people like Sharon and Trudy - that there are spaces in which they are not being judged or graded for them to engage in less stressful learning practice. Outside of class, students are able to shrug off their "empty vessel" persona and don their apprentice overalls, becoming engaged in their whole persons in learning practice.

**Between classes**

The classroom, after class, forms the backdrop for another serendipitous learning occasion in our opening example. Tristan, an I School alumnus who is in town for an event on campus, has joined the class and offers Sharon a "way in" to the community that he is now at the center of and which she is at the periphery. Sharon identifies with Tristan because of their shared experience of the I School and wants to learn the practices that will enable her "to become him". Although many of the students who I spoke to may not have known exactly what they wanted to do when they left school, many of them held up examples of alumni who were central to the communities that they wanted to join. Initiating the conversation by telling Tristan that she had watched him on video, Sharon focused on the interests and experiences that they shared. In effect, she was telling Tristan "I'm just like you therefore I should naturally be in the same community as you". Tristan then responded positively to Sharon by inviting her to a get-together that evening where members

---

12 Instructors allocate part of the final grade to participation in class (and sometimes on class blogs)
of the "inner circle" of his community would be socializing, thereby enabling Sharon an opening and a step into the community that she wished to join.

**The new system**

**Blurring boundaries**

Engaging with respected peers in networked spaces like Facebook can become complicated - especially in the relationship between student and alumnus. In order to connect with an alumnus, the student may attempt to use social inroads because of their mutual school experience. Students generally friend others from the school on principle and so this may offer a strategic way of gaining credence in communities we wish to join. Peter, another student I spoke with, describes feeling torn between wanting to "friend" a respected peer (who is a recent I School graduate and part-time lecturer) on Facebook and having him see his frequent updates.

I friended Oliver for the first time yesterday. I have uber respect for the work he can do. I'm like: I aspire to be you! So I didn't want to waste his time with my drivel. Because I wanted to put my best foot forward. If I wanted to get a job at (company name redacted) or some other startup that he knows about, I wanna be able to say: hey, Oliver, I'm skilled. I didn't want to have this looming personality through my posts that he would have to sift through and say: well, really… What do you do, Peter? I'd want to have my abilities first and personality second.

Because there is no (simple) way to distinguish between nuanced relationships like this one, Peter had an experience where one of his friends from another sphere reflected badly on him:

On Facebook I post things and a friend of mine responded in a rather crude, childish manner - genitalia stuff and I was like why? And so I deleted it. But whoever saw it, it's a reflection of me because a friend of mine posted something in relation to my status.

To some extent, they are peers because of their common experience, but in other respects, there is an unequal power balance at play here. The alumnus no longer has the role of "student" but the more esteemed "professional" role in society. Peter
knows that Oliver might prove helpful in finding job opportunities for him or help him to join the community he wants to gain entrance to. As a student on the periphery, however, he (theoretically) doesn't have much to offer the alumnus. Because there is no other platform for Peter to use to connect to Oliver in what is a highly contextual relationship, Peter must use the all-purpose, private Facebook platform where he has little control when his role as part of other plays overlaps with this emergent one.

This fear extends to the “invisible audiences” (boyd, 2010) that students often think about when they participate in online spaces. Many are fearful of the impression that they're making on potential employers, respected peers and lecturers. After being a prolific blogger in her teens, Sharon recently deleted her old blogs and restricted her social media profiles.

Another reason why I took everything offline - because it's nice if they search me and see the work I've done with other companies - but Flickr, Facebook - no way. Good reads - I even took that down - they don't need to know what I'm reading.

As we advance towards our new identity after “having learned” we hold on to some of the remnants of our old selves, trying those out with different people and advancing our new selves with others. In order to ensure that we're not “outed” with such apparent contradictions, we must ensure that we separate what some know about us as distinct from what others know. Privacy in this context, then, is about retaining the integrity of the context in which we share that information. Context, here, is specifically about information related to our different identities (rather than only about who we share that information with). As Goffman (Goffman, 1959) says, we have the ability to choose our props and costume and scenery to display to the audience and we choose these in order to retain coherence.

**Bringing the outside in**

The use of social media tools imported largely wholesale into the institution surfaces important inconsistencies for students. Last year, in a class on participatory media in education at the I School, instructors tested out new learning management
software that made heavy use of social media tools. Because the class was small and students knew each other well, and because it was successfully facilitated as a peer-learning seminar, students and instructors were generally fairly open (in the "whole person" way that Dewey describes) in the classroom, with lively debates and lots of social banter. The online component of the class reflected some of the jocularity of the class, although comments and postings by students were generally formal in relation to in-class conversation. The class was also being offered as online-only for those who couldn't make the class time, and it was from one of the online-only students that the following question came:

I wanted to get others’ thoughts and feedback on how open the site content like blogs should be. I know I felt a bit self-conscious and self-edited when I considered the off chance that content could be viewed or searchable by students or employers. This might be a function of the site – but I wanted to get some other views. Can we make the content accessible only to logged-in members? Does it matter?

This comment was interesting from a number of perspectives. It seemed to be concerned with what we generally call “privacy” but the content that it referred to was not what we generally refer to as “private information” (including your name, contact details, grades and private emails). The student was taking the normative model of what blogs are supposed to be (irreverent, personal, “diary-style”) and applying it both to what was expected of them in the class (i.e. to be conversational, to engage personally) but then recognizing that what they said would be public and thus accessible to those who might take what they said out of context. Using Nissenbaum’s "contextual integrity" framework, there seemed to be a threat that the norms of information flow in the learning environment may be disrupted if the information was moved to a different sphere.

Although blogging in class often aims towards personal reflection, the result of the fact that these posts are more often than not made available to the world persistently and out of control of the student is that students take themselves out of
posts. When I ask Sharon what she feels about blogging in class she said, “I mostly did it to get participation points - it wasn't anything personal”.

This reflects a missed opportunity. Networked technology offers a number of affordances that enable (not necessarily better) but perhaps just more occasions for students to engage socially with one another and thus better enable the kind of learning practice that can result from this trust.

**The digital spaces between**

While the institution provides social spaces for students to work together and collaborate networked spaces are less available.

Perhaps the most sociable list is "noise". Although "noise" is purportedly a space where "anything goes", there are, in fact some invisible rules in place which only become visible when they are broken and Noise is used by faculty, staff and alumni as well as students.

In the absence of such opportunities, students sometimes coordinate their own networked technology outside the reach of the administration. According to Sharon:

(T)he class of 2012 is all playing Words with friends together. It's like Scrabble for your phone. So Jasper - we have all our classes together and we started chatting - and one day he asked me if I play it and I said no and he said I should download it because other people are playing. And from there other people discovered I was playing and so we were all playing each other. So that's a weird bonding situation so I don't see Celia outside of class but we're still playing this game at like midnight...

Students also prefer to use other means to organise "outside" events because they are often unclear about rules for posting "social" events on an “academic” list. The student mailing lists are watched over by school staff who feel that because it is something all students are required to sign up to, they must make sure to exert some protective influence. Students often don’t know what the rules are for posting to the students list because it is unclear who is "watching" and although it would be
much easier in some cases to invite students to social events through the list, more often students use other platforms like Facebook or Evite.

This has caused problems in the past where new, perhaps quieter students are not invited to social events hosted by advanced peers because they were simply forgotten or were not already Facebook friends with the hosts or simply choose not to be on Facebook. In this way students who feel that Facebook violates privacy and that they do not want to join on principal, often become even more socially isolated.

**Conclusion**

It's difficult to imagine this school without its student spaces. Where would they go between classes other than the library (where they must be silent) or to a professor's office hours (where they must fall into the role of "empty vessel")? It would be a dull place, certainly, where students only run into each other in the classroom - a space where they are expected to act in a particular way and where the professor is always the authority figure.

As with the opening example, it is often the spaces between the classroom - in the corridors, before and after class, in the student lounges and outside the school - where we see these engagement of the whole person in learning practice. The key features of such spaces are that they are located within the remit of the institution and therefore carry with them the contextual norms and values shared by the school community; that, although these spaces were situated within the institution, they were located outside of the direct view and control of those in authority, and thirdly, that they were places where students had control over who knew what about them. Within the structure of individual grades and stressful social roles, the student lounge becomes a haven for students where they can conduct conversations unfettered by the need to perform or to display a particular competence and where they can have the opportunity to learn from one another.
In contrast, no such space existed in the digital realm. Student mailing lists were surveilled by staff who felt it their duty to watch the lists in order to protect students, and the class blogging systems were largely open and visible to the public Internet. In the absence of the student-lounge equivalent in the digital realm, students tried to squash uniquely shaped educational relationship pegs into the round- holed platforms like Facebook, leading to identity crises and self-censorship.

The future of learning systems technology seems to be in replicating spaces in which students can be social with one another in the context of learning. And having every interaction recorded and analyzed for educators to understand. But in doing so, we're attempting to bring the outside in, bringing the conversations that happen in the corridors and the student lounge and the classroom into the school hall, thereby disrupting the architecture that sustains peer learning. Using the school analogy, students don’t just continue talking the way they usually talk in the student lounge just because someone put a couch in the school hall.

As Glynda Hull writes in the foreword to Teaching Tech-Savvy Kids, “[i]f social networking sites are used by and large to extend existing friendship networks, what instructional scaffolds might increase their reach and function, and to what pedagogical, epistemological, and ethical ends? What youthful practices should remain youthful practices, allowed to flourish in out-of-school contexts untouched by school-based purposes, and where might sturdy bridging by built?” (Parker, 2010).

**Acknowledgements**

Thank you to Deidre Mulligan for her support and advice throughout this project and for making some of the connections to privacy that were so helpful. Thank you also to Paul Duguid for his comments and suggestions on an earlier draft and for inspiring the sections on learning - in both his teaching and his work. Thanks also to the students and staff who I interviewed for their time and for their openness.
Thanks to Helen Nissenbaum for providing the framework for thinking about this problem and to friends and colleagues at the I School for listening to me drone on about this with at least the facade of interest. I appreciate your solace and support in our spaces between. It is what got me through it.
Appendix

Literature review

I started this project by looking at theories of privacy from legal theorists like Daniel Solove, Julie Cohen and Helen Nissenbaum. Daniel Solove’s work in *Understanding Privacy* (Solove, 2010) is useful for extending our conceptions of privacy beyond “secrecy” (what he calls the “secrecy myth”) to its role in the maintenance of intimate relationships. Solove contends that 'The privacy-as-secrecy conception fails to recognize that individuals want to keep things private from some people but not others.' Although useful in understanding how privacy is being redefined, Solove’s view is limited to understanding privacy in terms of the visible harms, rather than those that may be invisible (i.e. when certain groups stop participating because of higher risks).

Helen Nissenbaum’s theory of “contextual integrity” (Nissenbaum, 2004) is useful to explain why even those activities taking place in public may actually be privacy violations. Nissenbaum posits two types of information norms that govern all areas of life: norms of appropriateness and norms of flow. Norms of appropriateness dictate what information about persons is appropriate to reveal in a particular context. Norms of distribution refers to the transfer of information from one party to another. Nissenbaum argues that personal information revealed in a particular context is always tagged with that context and that norms are relative or non-universal. She states that distributing social goods of one sphere according to criteria of another constitutes injustice. This is a particularly important perspective when looking at the ability to transfer information expressed in the educational sphere and using it to, for example, assess someone’s suitability for a particular job.
danah boyd (2010) says that, while Nissenbaum argues that we need to approach privacy through the lens of contextual integrity, she believes that ‘we need to examine people’s strategies for negotiating control in the face of structural conditions that complicate privacy and rethink our binary conceptions of public and private. boyd introduces the concept of “networked publics” to explain how networked technologies reorganize how information flows and how people interact with information and each other. According to boyd, privacy is not dead, it is merely in a state of transition as people try to make sense of how to negotiate the structural transformations resulting from networked media.

Moving from theories of privacy to empirical studies, I found the recent literature on youth and privacy to be particularly instructive in setting the stage for the way in which social relationships influence issues of privacy, community and learning. Much of the literature in this field seems to be a direct response to earlier literature and stories in the press indicating that “young people don’t care about privacy” (Nussbaum). In a series of empirical studies of youth in the United States and Europe, boyd (2010), Livingstone (Livingstone, 2009) and Parker (Parker & Hull, 2010) analyse how young people are negotiating the boundaries between their private and public identities using social network sites, indicating that young people have a nuanced understanding and concern for privacy, even though they also sometimes share intimate details of their lives with large networks of “friends”. Research by Hoofnagle et al (Hoofnagle, King, Li, & Turow, 2010) on young American adults (aged 18-24) showed that “with important exceptions, large percentages of young adults are in harmony with older Americans when it comes to sensitivity about online privacy and policy suggestions”.

Such studies should how young people understand privacy but does not explain how community membership is instructive in understanding many of the problems that young people face in open networks like Facebook. In a literature review on “Youth, Privacy and Reputation” Marwick, Murgia-Diaz and Palfrey point out that “Much of the studies of privacy online focus on risk, rather than understanding the necessity
of private spaces for young people where they can socialize away from the watching eyes of parents or teachers. These seeming contradictions demonstrate how understandings of risk, public space, private information, and the role of the Internet in day-today life differ between children, teenagers, parents, teachers, journalists, and scholars" (Marwick, Murgia-Diaz, & Jr, n.d.).

Christina Nippert-Eng’s “Islands of Privacy” (Nippert-Eng, 2010) looks at how people “do privacy” and provides an interesting lens with which to study the effects of power and accessibility on privacy. Nippert-Eng is instructive in how the architecture and artifacts define boundaries in which we want to be private (even if those boundaries encompass a group rather than just the individual). Nippert-Eng’s “boundary work” (Nippert-Eng, 1996) in which she interviewed people at a particular lab about their separation of “home” and “work” lives is particularly useful for its focus on the social dynamics of privacy - especially since it is a study of privacy rather than just informational privacy.

In addition to privacy research, I also looked at literature covering social constructivism and education covered in the body of the paper. In the early part of the twentieth century, John Dewey, for example, saw learning as an active, social process of constructing knowledge rather than acquiring it, and that instruction should be a process that involves supporting that construction rather than of merely communicating knowledge. In their “e-Learning and Social Networking Handbook”, Mason and Rennie (Mason & Rennie, 2008) understand the move towards the use of networked technology in the classroom as highly compatible with constructivist theories of learning.

Brown and Duguid offer an important critique on wholesale adoption of 'social media technologies’ in education. In an important article entitled “Organizational learning and communities-of-practice: Toward a unified view of working, learning, and innovation” (1991) they say that the assumption that, given the "right" medium people will exchange information freely overlook the way in which certain socio-
economic groups, organizations, and in particular, corporations, implicitly treat information as a commodity to be hoarded and exchanged. In their work at Xerox PARC in the 1990s, Brown and Duguid explored the inherently social nature of learning, and how successful learning happens with and through other people and that what we choose to learn depends on who we are, who we want to become, what we care about, and which communities we want to join (Brown 2008).

The structure of education today assumes that learning “has a beginning and an end; that it is best separated from the rest of our activities; and that it is the result of teaching” (E. Wenger, n.d., p. 3). But anthropologist, Jean Lave and computer scientist, Etienne Wenger, investigating learning communities in countries like Brazil, Liberia and the United States found that learning is what we do in the practice of everyday life and that it occurs through a process of social participation within a community of practice.

As Lave writes, rather than “acquiring knowledge”, learning is actually about changing participation in social life. “There is no such thing as ‘learning’ sui generis,” she writes, “but only changing participation in the culturally designed settings of everyday life. Or, to put it the other way around, participation in everyday life may be thought of as a process of changing understanding in practice, that is, as learning” (Jean Lave in Illiris).

As McDermott (in Murphy 1999:17) puts it, “[l]earning traditionally gets measured as on the assumption that it is a possession of individuals that can be found inside their heads... [Here] learning is in the relationships between people. Learning is in the conditions that bring people together and organize a point of contact that allows for particular pieces of information to take on a relevance; without the points of contact, without the system of relevancies, there is not learning, and there is little memory. Learning does not belong to individual persons, but to the various conversations of which they are a part.”
This view of learning is in stark relief to those advocating for “open education.” With a focus on abstracted information itself, the open education movement declares that online educational resources, isolated from the context and communities that gave them life, are in themselves “revolutionary”. This is epitomised in Charles Vest’s 2006 article in EduCause in which he declares the open materials and platforms are enabling the emergence of what he calls a “meta-university” - “a transcendent, accessible, empowering, dynamic, communally constructed framework of open materials and platforms on which much of higher education worldwide can be constructed or enhanced” (Vest, 2006). Discussions about the “power” of “open education” are expressed in terms of the growth in the “accessibility” of educational resources. According to Caswell et al from Utah University’s Center for Open and Sustainable Learning “OpenCourseWare act as enablers to achieving the universal right to education. These technologies, and the associated changes in the cost of providing access to education, change distance education's role from one of classroom alternative to one of social transformer” (Caswell et al, 2008).

In assessing the resources available in answering the question of whether education should be open, there are a number of gaps for pursuing worthwhile research. Social scientists like Paul Dourish have done a good job of starting to bring the world of privacy and security - so long dominated by lawyers and technologists - into the world of sociology and design. But there is very little that assesses the impact of different concepts of privacy (online and offline) on institutional arrangements (such as the 'school'). The goal of this paper, then, is an attempt to bring together legal, sociological, technological, historical and educational literature together to provide an empirically grounded perspective on the 'open education' question.

**Interview questions**

**Question set 1**

*Roles, participation*

How did you come to the I School?
What do you see as your role in the class/classroom?
How did you feel on your first day at the I School?
Can you describe how you felt when you raised your hand/speak up for the first time in class? Do you feel like you understood the culture of the I School when you arrived? Can you think of any instances where that was/was not the case?
Do you see yourself as an active member of the I School? Why/why not?

Sharing
How did you feel about blogging in class? What did you like/not like about it?
Have you presented your work outside of class before? How did that feel?
If you miss a class, how do you catch up?
Have you ever been asked to share/take notes for someone who has missed class?
How do you feel about people recording classes?
Have you stopped saying something recently when you thought someone might steal your idea before you were able to say it first?
How important is the final project presentation to you? How do you want people to see you?

Identity
Have them look at their I School profile/Google themselves
You have a picture and some information about yourself here. Why did you choose to show/not show that particular info? What does this say about you?
Do you feel like people know you here at the I School/p2pu? What don't they know about you?

Privacy
What does privacy mean to you?
When was the first time you realized privacy was important?
Have you ever said something online that you wished you hadn’t? What was the result?
Do you use different usernames? What does your username say about you?
Do you use Twitter, Facebook, LinkedIn differently? How?
How do you feel about students/instructors friending you on social network sites?
Do you have a rule about that? Do you think others have the same rule?
Do you feel restricted by who you are friends with on social network sites?
Do you rely on specific friends at the I School to help you navigate the culture?
How do you feel about the fact that things you’ve blogged about might be persistently available in the long term? Do you know what the I School’s privacy policy is?
Is there anything else that you feel I should have asked you?

Question set 2
Can you draw the classrooms that you sit in for each class and show me where you sit?
Why do you sit in that particular space?
How do you feel when you're not in your seat (e.g. when someone else takes it before you arrive?)
Who do you sit next to? Do you notice what they're doing?
Think back to yesterday. Describe your movements within the school. Is this typical?
When you saw people in the school, what did you talk about? (go through each of the places)
What was the last thing you said in class? Is this typical?
What communities do you see yourself as a part of? For example, you could see yourself as part of the I School community, the 2012 community... what else?
Citations


Hoofnagle, C. J., King, J., Li, S., & Turow, J. (2010). How Different Are Young Adults from Older Adults When It Comes to Information Privacy Attitudes and Policies?


