

These transcriptions may contain errors, especially in spelling of names. These are unfortunate, and we regret that we do not have the resources to fix these errors. Still we believe these transcripts will be valuable to many users.

Participatory Personal Data

Katie Shilton (@KatieShilton), University of Maryland

>> I'm Katie Shilton. I'm an assistant professor here at the iSchool at the University of Maryland, and I'm going to be talking today about Participatory Data, data that people collect about themselves using mobile technologies. But before I start that, Mark can ask us each to give just a little spill about on how we got here. And I got here through sort of an odd path. I trained as an archivist. I thought I want to be an archivist, and so I went to library school. And I was really interested in archives because I'm really interested in who gets to make decisions about our historical record, right? So, who gets to make decisions about what's in the archive, and those are really interesting question. But when I got to library school, I realized that question is changing a lot right now. There's a lot of content out in the world that didn't use to exist. There's a lot of participatory media, participatory content which is really, really cool. And the archivist in me sort of said, "Great, now we'll have a complete historical record" which is not true, right? We all know it's not true. But there are all these things that push back against that as well, right? When we start to record everything, do we really want to keep all of that around? And I got really interested in those questions. And it was less about archives now and more about people's interaction with technology and their practices and why do they record this data and whoa, the technology affordances, what do they allow, what do they not allow, how much of the stuff is going to be kept, and then how much of that is built in to the infrastructure? And so, I say this just sort of demonstrate that as you're on this Ph.D path things shift. They change and there's sort of new areas opening up for those of us who thought we were really traditionalist. And I wanted to do this really old fashion thing. So, that's how I got sort of here. That's why I care about this kind of participatory data. I'm interested in it as a record, but also in the social values that go into both its practice and recording it and also the design that enables it. And so, I'm going to talk today about what this data is, this participatory data I'm talking about. What its relationship is to technologically-mediated social participation, and some of the challenges and concerns that it raises. So, participation with data. What is this data? I'm broadly interested in emergent types of personal data as well as the new data collection practices and technologies that go along with this data practice. So, this could be familiar tools like Facebook and Twitter. We're all creating records about ourselves all the time. We know that the Library of Congress has an archive Twitter postings like that. So, there's a new kinds of records about our selves. But the ones that I've been looking at are data applications that are being built for mobile phone networks. So, I'm really interested in the data we collect about ourselves, or the data others collect about us using our mobile phones. And the reason I'm interested in mobile phones is that they are one of the most pervasive digital--they are the most pervasive digital technology in the world. They have billions of users all over the world. They're quite familiar and accessible, and they ride along with us, right? We take them where we go. And even the simplest phones, non-smartphones can record text data and upload it or send it via SMS. We also--all cellphones record or can record location data. Now, smartphones can do that with GPS, but a non-smartphone is using cell towers to locate itself. And your phone company has a record of that, right? So, location data is built into all of this. And then, smartphones can record photos, videos, increasingly accelerometer data, so how much you're moving. And this kind of data can be used for all kinds of cool new sorts of data tracking. So participatory research, you know, on--by one person if you want to record data about yourself or by thousands of people, potentially. So, many other couple of examples of what you might do with this data. So, these examples are for context. These are from two projects that I worked with during my dissertation research and I'll talk a little bit more about that center that ran that, but these are to give you an example of what this data might do. The first is the Boyle Heights Community Development Project. And this was a development project run in a mixed-income neighborhood in Los Angeles, in East LA. And it

was initiated by a coalition of community groups that wanted to do a community needs assessment. So, they were going to develop a redevelopment plan and they wanted to get a feeling for the neighborhood and what people do on a daily basis in that neighborhood. So, the community organizations recruited a bunch of volunteers who run specialized software on their mobile phone and that software popped up location-based surveys. So as people went about their day, it would pop up a survey when they pass the local park or the grocery store or the hospital or the school and ask them about their use of that space, whether they felt safe there, whether there was, you know, access to fresh foods, or--and whether they like their school. All kinds of questions like that. All based on location. So, what you're seeing here are very simple data entry screen and then a map of one of the users, right? So in order to this, it records where you go all the time. A really different example is a class of individual health monitoring tools that are emerging under the name Mobile Health or mHealth. And this is an umbrella of software applications both commercial and research-based that help people track personal metrics about their own health. So, the samples are generally--this sort of involves running software on your phone again and it can initiate--that software will initiate samples. And these are called "experiences of daily living," sometimes are "observations of daily living" or "experience samples." And the idea is your phone knows again where you are and can pop up questions for you throughout the day. So, what's your mood? What's your stress level? How are you feeling? And it can geotype those. It might pop up reminders at meal time, for instance. If we go into a restaurant, it might say, "Oh, what did you have to eat," right? Or it might ask you to input who you're with and who you're talking to, to sort of get the idea from sociality in eating. It might use the alarm on your phone to understand when you woke up and pop a question about your sleep quality, right? So, you could track a whole variety of things using these mobile health technologies. And so you might collect your GPS location trace, your text data, images of food, of people you eat with, and then you might share that, right, with a coach or a clinician, researcher perhaps, friends or variety of people you might share that data with. Yeah?

>> --access to that like partnering with open to mHealth?

>> Yes. Yes, exactly. So-- [Inaudible Remark] I didn't have access to users' data. The user themselves have access to this data. And actually, I'm going to talk a little bit about where this data goes and who has access to it 'cause this is actually a critical question. So--and, you know, I think actually is your concern suggesting as these example suggests, participatory data applications have some pretty useful applications going forward, but they also collect extremely sensitive information. And we got a recent white paper that the group I was working with put out, got a response from somebody at the EPA. And she posts it on a LISTERV [phonetic] and posted "Participatory sensing is either terrifying or electrifying depending on who does it and how the information are processed." And she's right, right? There are two big "ifs" in participatory sensing. Who is collecting this data? Is it somebody you trust? Is it the government? Is it advertisers? Is it your health insurance company? University researchers, maybe we trust them, right? And how is it processed? What can be inferred from this kind of data? So, the question I became really concerned with was what differentiates participation in data collection? Collecting data about your self and your community, perhaps, and surveillance like what is the difference between these two things? When does recording data about your self and your community shift from social participation to surveillance? And I'm going to suggest that there are at least four conditions. There are probably more, but there are at least four that we need to think hard about. And this come out of the surveillance studies [inaudible] primarily, that distinguish participatory data collection from surveillance. The first is the ability to make decisions about when and how to collect the data. So, this is about consent. Do individuals know that they're participating in these projects? Do they forget that they--that software is running on their phone? Do they know what it means to collect this kind of data? And so, there's a real challenge here. What is informed consent on tradition we have in social research mean when these devices are so close to our everyday lives? It's so easy to forget about. A second condition is having access to our own data and controlling its flow. So, this is what Jessica was talking about earlier in terms of privacy, right? This is a classic privacy concern. Who can see data when, and how much

can participants hide? And Jessica was talking about interpersonal privacy which is certainly important here. You wouldn't want to give your health information to your boss, but I'm also talking about broader surveillance concerns around privacy, right? You wouldn't want to give that health information to your insurance company necessarily or to the government. So, how much can participants hide? Can they lie? Can they fudge data? These are all questions. So, a third concern that differentiates surveillance and participatory data collection is preventing people with power over you. So, that is corporations, law enforcement from using your data against you. So, this is about power and equity. Collecting information about people is a primary source of power in an information economy. So, who benefits from this kind of data collection? And will this data be used in hiring decisions or insurance decisions or even product pricing decisions? We're seeing lots more of that. And the fourth concern is being able to wipe your slate clean, right? If you collect data about your health for a month, does that data always need to be around about you? Should that be there forever, should always be accessible? Do we need total accountability or memory for our actions and routines? So, this is a question about forgetting. Is this something we should be able to do? There's no question that continuously recording your location activities in particular in health behaviors. All of those kinds of things can be pretty invasive. And I want to point out now that there are few or no laws in United States protecting this kind of data. And I'll talk a little bit more about the policy implications later. But right now, this data is completely unprotected legally. So, it becomes a question of social values. What moral and human values are important here? And the conditions I named, privacy, equity, consent, forgetting, are all values that can guide design of these technologies. If there are no laws protecting this data, perhaps, we should design protections into the technologies we are building to the software. And then a larger movement right now, must be going on for 20 years, of people that are trying to figure out what values are built into our technologies or should be built into our technologies. This is broadly called values-in-design or Value Sensitive Design. And my dissertation is part of this larger trajectory trying to figure out these questions of values within technology. And so, these social and ethical challenges particularly raised by this kind of data, let me to want to know how do these challenges affect technology development? Are technologists talking about these challenges as part of their practice? And what makes social values apart of their discussion? So, my dissertation project was an ethnographic study of social values within a design process at the Center for Embedded Networked Sensing which is the mouthful, CENS. And CENS is a--was a large computer science center. It was devoted to developing sensor networks. And they had in the years I was there moved towards developing mobile phone-based sensor networks to sense people. So, they were actively developing software to collect this kind of data. And both the Boyle Heights project and mHealth come out of that CENS. And so I spent two years as a participant observant--participant observer in that lab. I went to meetings. I hang out. I did my work in the lab. And I have to say it was a participant observation with a real emphasis on participant. I was a member of the team. I participated at meetings. I was vocal and active. And that's actually me with much shorter hair, confuses everyone participating in the lab, right? And this is--these are what meetings look like. There are lots of white boards and a beautiful building. This is sort of CENS as ethnographic object. I collected lots of field notes. I did interviews with all 30 or so staff members. They were mix of faculty stuff, full time stuff, and students who work there. And I also did document analysis. I looked at all of things they were writing, their justifications there, you know, their sort of the entirety of their design process. And I began my process by getting oriented to CENS design. I wanted to figure out, okay what are their existing design practices? How do they work? And the answer to that was meetings. It was a very team-driven place. They had at least one full group meeting a week. They had multiple small meetings and then ad hoc meetings throughout the week. So, that is great for an ethnographer because you can see people justifying their decisions, because they have to have to communicate to each other. Much harder to do ethnography and software design when they do it by themselves 'cause you may have to look over-the-shoulder at the code much harder. This was a good situation for studying design. So, I was watching them hash out their problems. And what I was looking for were activities that encourage designers to pay attention to and talk about and design for privacy, consent, equity and forgetting. What made those topics appear? And when I found activities that cause these values, topics to appear, I called those values levers because they pry, open new conversations

about social concerns and they help the designers build consensus around those concerns as part of their work practice. It's something that's important to design for. And so, I started looking for values, levers, and design, and notice their position with the design. And notice that some of these activities I started to see were inherent to design. They were already there. They were part of the design setting while others came from outside influences. I'll talk about sort of both kinds in detail. So, there were many levers that were already part of design at CENS. They were inherent to the setting. And my first sort of "ah-ha" moment when I saw this happening, when I saw what this was, was in meetings with statisticians. There were some statisticians. The team was primarily composed of computer scientist, but there were some statisticians who were members of the team as well because they were dealing with a lot of data and then you have to figure out how to represent it. And in these meetings with statisticians, everybody will have to talk about the data because the data was what statisticians care about. So, they would talk about representation and, yes, they would have methodical discussions, but they would also talk about the sort of ethics of how to represent that data and what that data says and what you could infer from it. And that conversation about the data was really important for getting computer scientists to sort of pay attention to the values issues. So they were really imbedded in that data. The pipes, yeah, the harder to see the values issues. The data, easy to see the values issues. And so, I notice this is not just with statisticians after that, but any interdisciplinary folks who came into the lab, an urban planner, a sociologists. Everybody had to talk about the data. That was the lingua franca of design, right? That was the piece that everybody could talk about. And so, that became a really important values lever bringing interdisciplinary folks in and forcing everyone to talk about the data. Another lever that was related to data was gaining funding. This is sort of the interesting thing to watch. As a project got a grant, the funding team or the design team would grow. They would add a full-time programmer who would actually build the code which was a huge relief to the grad students who didn't want to write that code, right? And so as they added people, they were forced to talk about the data much more as they had to communicate across different people about what they want as designs. So, I started to see more and more values discussions as teams got bigger and having more people worried about the data. And in that final lever I'm going to talk about which was really interesting was experiencing internal testing. When a developers track themselves, they got the values issues right away, right? So, a common practice at CENS was internal pilot. This is just good design practice, you know, they call it "eating your own dog food," right? And it's meant to test for bugs and usability. This was not a specific value-based activity. This was to make sure that the server can handle the load, right, and that there weren't any bugs in the code. But in my interviews, I started to notice a hesitation to test sleep or food studies that we were running when they knew their peers will be looking at that data. And so, the laboratory test, pilot test became a values lever. So, we started seeing quotes like "If I'm going somewhere and I don't want anyone to know, I think I need to turn off this application or leave my cellphone," or another grad student, "I did a data collection for T [phonetic]." Another grad student, "For like a week or something, then I felt like not privacy, but I thought that I wanted to go out more actively. I felt like, oh, they're watching me, I need to be more active", right? So, you started to see a lot of this kind of contemplation or reflection about "Oh yeah, right, this is what users will experience." So, self-testing like interdisciplinary teams foster to focus on the data. And the kinds of data under request made surveillance and privacy concerns really concrete for the designers. There are also a number of introduced values levers in this space. These were not accidental. These were things--interventions by outside people that brought values into the design discussion. So, user feedback was a really interesting wannabe's. CENS leaders made a very explicit decision to do user set the design to it, involve users in the design process to get feedback. And when users had concerns about things like privacy and consent, it did carry a lot of weight in design. People really listened to it. So, I saw this with the grad students and post docs who are working on medical studies or on health studies, users had a lot of privacy concerns and they would take that very seriously. The problem is that user privacy concerns are not consistent across all domains. They're contextual, right? So, we have a lot of privacy concerns in the HealthSpace. Many few were in the campus trash collections space and there was a campus trash collection app that was about recycling and trash. And the users in that study said, "Oh yeah, no, we don't care about privacy. Privacy doesn't matter." And the designers heard privacy doesn't matter to users not privacy

doesn't matter in this context, right? So, user feedback was a very difficult lever to interpret and was very sort of sensitive. So, I think it was a little less effective than it could've been. More effective levers included mandates imposed by UCLA's IRB. And we sort of all hate to hear because we all complain about it. But these were the students who were working with the IRB for the first time. For the most part, computer scientists historically have not had to get their projects approved by the IRB. And it was a really interesting values lever for them to have to fill out the paper work and to decide about what consent look like in the space, and to realize that their institution took it seriously. And finally, there was a lot of value's advocacy and this is unique to this space. What is interesting to think about, I think, is we think about shaping labs going forward. Leader advocacy. The leaders, there were three PI's in this space and they took values very seriously and continually advocated for it, and they sort of helped to normalize these conversations at CENS. So the three major leaders, the three PI's, essentially, they are all faculty, pretty senior folks took privacy, consent, equity, and forgetting very seriously. And they would raise these concerns during meetings. And when we're viewing what projects would go forward, you know, they have a lot of power in terms of deciding what projects went forward, and privacy and consent concerns became part of that. But they also set lab procedures. And I think this is really important. Again, this sounds really boring, paper work, and nobody likes it. But it does do important work in design. They mandated an online form, before you even had to apply to the IRB, there was a pretty simple checklist where you said what kind of data you were collecting? And at one point, I had that on the slide but it's gone now. What kind of data you're collecting? Who? From? Were there special populations kids involved? All those kinds of things. And it sort of forced the students to think through. Okay, right, yeah, I'm collecting data from kids. Maybe some concerns here. So, and it really emphasize that these values were going to be taken seriously by the lab. It also helped those of us who were concerned to the values look out for potential conflicts. So, we could talk to students directly if we saw conflicts arising. Okay, so as I analyze my field notes from these years in the lab, I began to see a process of moving from relatively abstract social values. Those concerns I brought at the beginning, privacy and consent and equity, pretty abstract to design decisions. And it's a complex process to go from an abstract value to a technology design, right? And so in CENS, I notice that this proceeded in roughly three faces. I'm [inaudible] these values into design. First was agreeing on values of importance. So, that was the values lever part of this, right? Values levers helped the team come to consensus like oh yeah, privacy is something we should care about. Consent and equity are things we should care about. Then, we had to operationalize those values as designed principles. Design principles are targeted concrete goals for a system. So, it's very easy to say, "Yes, we care about privacy" but as Jessica demonstrated, a lot harder to build privacy features, right? What does that mean? What is privacy in this system? So, we would come up with things like, okay, local control of the data. The user should be able to have the first point of control over their data. They should own their data and be able to make decisions with it. And legibility, we should make UIs so that people understands what this data says about them, right? Or long-term engagement, people's needs are going to change over time. Let's keep them engaged with the system so they can continue to make decisions about sharing over time, or parsimonious is an important one. Let's record the least amount of data we need to answer whatever the questions is we're trying to get at. Let's not record everything, right? Let's be parsimonious. So, that was something we can operationalize. And then finally, you have to move from those design features to tech features and affordances, to the things you build into the system. So that was, you know, the architecture on how local control is reflected in the very architecture of the system, or the UI design, or other affordances like filters. How are we going to built good filters so that people can share off subsets of their data? Or data retention procedures, that turned out to be really hard to do, right? We have back up all our systems. Data retention is tricky and it involves a lot of--making sure you've actually deleted what you said you've deleted. Some solutions that we came up within a space, that came out of this discussion about values. So, the values levers made this discussion about values part of design and then translating those values into principles and features created some solutions for this space. And these are sort of partial but we're getting there which is exciting. One example solution that came out of these values discussions was what we call the Personal Data Vault. And the Personal Data Vault is a secure cloud repository for all of your data that comes off of

your phone. And we sort of thought about like a bank account for you data. It's the place you go where you can see it all. You can see what it says about you. You have some nice visualization features. And then from there, you can make sharing decisions. You can send subsets about data out of third parties, right? And this is an idea that's popping up in several places like any good idea. It wasn't desirous, right? It's there. I have been some prototypes built at Stanford in AT&T lab as well as at CENS. And excitingly for us, they're starting to be commercial products emerging. There's a start up here in D.C. called personal.com which is building a data vault and has an interesting model to fund it. So, it's starting to be a way of, you know, I think we're starting to think about controlling our personal data in unique ways. This might be one. So, a value of space design process can inspire technical creativity. And that's great. And we need that. There's a lot more to be done in the space. But we're going to need more than just technical solutions to make sure we can collect personal data and a participatory and not surveillant way. So, there's a need for new legislation to protect some of this personal data that's flowing through this mobile ecosystem right now. The data collected by mobile apps is currently, as I mentioned, not protected under the Electronic Communications Privacy Act. It's not protected under HIPAA even it's medical data because you generate it, and not your medical provider. And so, none of this has any legal protection. So for instance, if data vaults are going to be part of the solution, we're probably going to need to protect the data and the data vault. Otherwise, as Mark mentioned earlier, you've got a real honey pot, right? You've got everybody's personal data in one place. That is right for subpoena and it will be subpoenaed, absolutely. And so, protecting some of this data, perhaps meeting a warrant to access personal data would be a good standard as we start to think about how to encourage this. And I want to point out that there is precedent for protective laws that are--to encourage participation. In the genomics and genetic space, we have the Genetic Information Non Discrimination Act or GINA which was implicitly designed to protect people who wanted to participate in genomics projects by giving their genomic data from discrimination by employers or insurance companies. So, if you make your data public or give it to people, you can't be fired or denied insurance based on your genetic information. Perhaps, we need something like this with this personal data if we think of this as the footprint of our lives, right? And the Federal Trade Commission is currently starting a multi-stakeholder dialogue about best practices and privacy for mobile applications. It's a pretty unregulated space. It will be interesting to see what comes out of that. I will be watching that closely. Okay, so both of these issues, technology solutions, policy solutions emphasize the need to understand an entire data ecosystem. There are dozens of players in this space. You know, my dissertation is about just one, one research lab. But there are start-ups research labs, telecoms that are player in this space because they have all the state as well and all passes through their network, right, so they have it as well. And then regulators, oh and users, right? What are users doing? What are their practices around this data? So, I'm beginning a second study, a new study of start-ups and telecommunications companies and federal regulators and users and specifically, when you use interviews and participant observation to focus on privacy, challenges, and decision making among these groups. And specifically, places where creativity comes into the process where privacy can inspire, I think, you know, we hear a lot about privacy tamping down on innovation, but it also can inspire, values can inspire innovation in interesting ways as well. So, I want to look for both of those things in this ecosystem and see where this is occurring and why. I'm also working on studies of users. We don't know a lot about it, and I was just talking something earlier today. We don't know a lot about the motivations of people for collecting this data, how long they do it? Why and why not, all of these things. So, we're working on--working with my colleague Beth St. Jean who's here and Brian Butler on study of self quantifiers who are sort of enthusiast in this space, and their motivations and practices around this data, and also conducting a survey with another colleague of privacy concerns in context in this space. So, what are people's concerns and their values really in this particular context? But lots more is needed. So if folks are interested in this general area, there's a ton of research to be done in this space. So, for researchers of the sociotechnical, right? For those interested in technologically-mediated social participation, there are a lot of big questions here at the intersection of technological trajectories in mobile data collection as a new and emerging phenomenon. And where intercepts with social challenges is a really important question, so thinking about values and ethics and justice in this space. And there's a lot of room to investigate and

understand how these technologies affect users as well as institutions, and how we can best design technologies to reflect fairness and social justice that, you know, if we want to do this, how do we do this right essentially? And I want to close a video that researchers in this space can be part of design. But we bring understandings built on professional ethics from some of our fields, research ethics from many of our fields, and studies of justice, right? Surveillance studies is really all about that. So, we can bring those traditions as a skill set for design. Operationalizing social values can be a skill set for design. So, I'm going to leave it here. Take question from you guys. Leave you with a great technology ethics comic.

[Laughter & Applause]

>> You can ask questions seriously.

>> It's true. [Laughter] Read the comic and then you can ask about the comic. [Laughter]

>> Well, I'm wondering when you talk about the values. Of course, values are the--that you have--vary cultural differences because of our values of shapes like cultures. Are we just talking the way it started?

>> Yes.

>> We love it.

>> Really good question. So, and this gets at this question of values versus ethics and, you know, which of those. So, Value Sensitive Design really struggles with this question because values are contextual and specifically cultural. And so in my study, it was very much US based. It, you know, it's in the US policy context and the US design context. But the challenge is this design is for everyone, right? These things are being designed for everyone, at least, a vision of the designers 'cause this could be used around the world. And so, there's a real challenge about how to do Value Sensitive Design when the context is supposedly everyone because we know that privacy, for instance, is very different across culturally, you know, I think things like equity and consent perhaps are more universal but, you know, we see challenges there as well. So, I don't have a good answer except for that we need a lot of comparative work. And we need--but I think in terms of trying to design for everyone of what global vision for design, we have a huge challenge there that I don't have a good solution for.

>> But do you know if there are other countries who are working [inaudible]?

>> Yes. Yes. And again, Value Sensitive Design for whatever reason has tended to be a very US phenomenon. But they were sort of starting to see, for instance, it comes out of the participatory design movement which is much more global and has many more case studies in a variety of places around the world. And so, I think we're just starting to see Value Sensitive Design or values-in-design also take on cases studies elsewhere in the world. But yeah, I think let's encourage this phenomenon. It really--this is work that really needs to be done.

[Pause]

Emma?

>> How do you know when you yield into Value Sensitive Design, how does that fit with national security issue? For example, if you want to discuss, you know, private data and made some commitment and the other norms to apply that in design, how does [inaudible]--

>> Right. So, you have just identified something which I glossed right over which are values conflicts. How do we set design for conflicting values, right? And when--and you tend to see essentially designers taking a stance. Realizing at some point that their design practice is political and saying-- very explicitly actually, actually, you see this all the time, oh, the police should not have this data, right? So, that's a value statement right there. And because of--so, something I'm sort of working on now is, you know, where are these values coming from in the design community? And there are historical trajectories here, right? There are design practices tended to be libertarian in politics, and that's not always true, and they're definitely collectivist impulses and values within design and it's, of course, not a uniform space. But there are particular political conversations that happened on these settings. They tend to be rather anti-surveillance in the United States, and that's just sort of a history of this community. And so, you now, in terms of do I have an opinion about weighing national security versus participation? Yeah, I come out of surveillance studies. So, you know, I have a political stance that we should push back against surveillance always even at the expense of security. But that's mean, right? That's--I don't get to make that design decision for everybody so I think that this is a tension we're going to continue to see as we see computer science values budding up against the government values, budding up against user values. And I think the best that I can say is, well, let's consult the users on that one, right? That if we think about national security as an interest of the population, but also being able to record our own data without fear of surveillance as an interest in population. Let's get users to weigh those two and decide which one they want. Oh yeah, I have to make the decision. [Inaudible].

>> Something that--well, I just--I thought this was fascinating 'cause I also sometimes talk with the designers and the way they do and don't explicitly think about the values and now, it comes out as a very interesting work. Now, before asking and talking with users, I think that's a great first step. And I have some survey data asking what color students--

>> I'm right.

>> And I think the good audience to ask these questions because they are usually a little better as this are better case scenario. And their imagination on what could be done with the status is quite limited, right? And I don't really blame people. It's hard to foresee how this aggregation of data and its intersections, it's just not something--I mean, there are all sorts of way in which they're aggregated [inaudible] that you don't really expect someone to be able to conceptualize. So, if I ask them who do you think looks at this and who do you think might ever use this data? At most, they're thinking of their peers. Most of the time, it is best for social socially [inaudible] thing, the kind of future visions.

>> Right.

>> It's just not on their radar which brings me to it's like a trash of common situations that cause--well, the users can't be expected to fully grasp the consequences of a very complex system that designers don't have any clue either.

>> Right.

>> So, there's nobody here maybe except like kind of like academics [inaudible] who are thinking the big picture--

>> Sort of "what ifs" and that, yeah, that is possibly--

>> --and call this complex thing. So, I'm for asking the users. I'm just saying even that's going to hit a wall because people are going to look their everyday lives. Their [inaudible] instead of peers, and the data is so big and complex, increasingly. I mean, this is not an academic question.

>> Right. [Inaudible Remark] But it's a good, yeah, I know, practical. I mean on my bad days, I say, "Well, there's--this got--have to be a disaster of some sort," right? And then, everybody will pay attention. But I hate that answer, right? I hope that doesn't have to happen, that we don't have to have some massive image of, you know, I don't know what it would be 'cause we don't [inaudible], right? But I think to get to the question of users and their sort of knowledge of the possibilities, I do--I started to see and I need to do more research to, you know, substantiate this but that there are differences in people's imaginations about what could happen to define our long lines of privilege. So, the graduate students I worked with didn't seem to fear the police and [inaudible], right? But when we worked with student organizers, high school organizers of East LA, they automatically--like immediately, we're like, "Oh, no way, like you're going to track my phone now." [Laughter] You know, like somebody will, the police will ask for that data or immigration will ask for that data. So, there are some, I think, populations that have heightened sense of what bad things might happen or could--because they have been happening, right? And so, I think there are some lines there that might be worth looking into. But, yeah, I think trying to visualize what the eventual harms this data are and understand them all. We probably need some plan of researchers and research and, yeah, all of the above.

>> [Inaudible]. There was a big story that broke this morning about the Romney campaign doing that to identify potential large donors than doing this really advanced analytics and a secret behind the stage in aggregating the social--

>> Right. Fundraising is actually one of the places where this has been happening for a long time, you know, and so, I work as a fundraising person. We did the opposite of what I'm talking about now which was all the surveillance we possibly could 'cause you will have to know where our donors are giving. And if they give it to organizations like yours, and yeah, so I mean, I think we have seen elements of these for a long time and that it's maybe the scope of it now that seems scarier. But, you know, your all kinds of data, not just your financial records but it could also be crossed with your preferences for music and also your health data, all right? So, there are all these cross possibilities. And that's why they are up there. They're really nice privacy and context lid that tries to deal with how do we separate some of the extremes back out? It's a hard problem.

>> Do you have any thoughts about like kind of accepting the things happened and then taking in a [inaudible] of education? This is sort of sound like too cynical but--like the forgetting one, it's really difficult. And I think, you know, Mark mentioned this the other day about Facebook saying that, yeah, they can delete things or whatever but I want something that's posted online that becomes very permanent like what are your--the organization you're interacting with, archives, or third party or something like that. And when something gets posted, it becomes very permanent. So, is there any thought to say that's still invaluable to talk about upfront and what can you build to kind of put in that to be more than education thing. And, you know, I've seen a lot of kinds, but it's like talking with people in high school saying "Hey, you know, you really need to think twice before you post everything, every picture you have on Facebook because it's permanent."

>> Right. Yeah, I think there's some balance here between building infrastructures that tend towards forgetfulness, perhaps, or that have more of the ability to delete built in. That's something we can do, right? It doesn't exist right now, but it is something we could do if we put our minds to it and it's actually Europe that's--so right now, in Europe, there is a very controversial piece of legislation that is, I would guarantee you, a right to forget. And that would be precisely be about deleting data. Now, whether or not that could

happen and there's the type of community is up in arms about this. But it's an interesting thing to think about. But I think you're right, that we balance it with literacy and reputation management and sort of the skills that we all do, right? We all do what Jessica was talking about earlier, right? We all present, and so, learning to present online is part of this. But yeah, I think technology is much more likely to--it's harder to present online. We have trouble figuring it out when you have statements like [inaudible] saying, well, they shouldn't be doing that thing. You know, just because we're exposed doesn't mean it is our fault, right? No, we shouldn't. So, I think we need a little balance.

>> Can you talk a little bit more about the CENS IRB interview?

>> Yeah. What do you mean--what specifically would you like to know? Like how they interacted or they actually did it?

>> A little bit of both and sort of talk about what role does IRB play in this project.

>> Oh, such a good question, okay. [Inaudible Remark] Great. Oh, man. So, I could do a whole talk on the role of the IRB because I think the IRB can be critical here as I emphasized, but what I left off for length was the flip side of the IRB which is when they don't understand and the damage that does with designers. And so, these are folks who had to do IRB for the first time. And we're generally receptive to it as much as you can be receptive to adding paper work to anyone's routine which, I mean, these folks like didn't want to document their code. They really didn't want to fill out IRB paper work. However, back to life, you have to do it. So, that became the norm in the lab. It became sort of accepted and [inaudible] help with it. They had an administrator who would help them and, you know, so it became a sort of naturalized process which was nice. And about records of the time, that worked really well. The IRB will come back and say, "Yeah, you maybe need to make the consent form, you know, fit better into somebody who were doing paper consent forms for online surveys." And the IRB will say, "Well, it might actually make more sense if you did it all in context," right? So, sometimes they were really great. But occasionally, something would come back like "Oh, you can't do this invasive tracking thing because you're not keeping your consent forms in a lot filing cabinet," right? And then, that just destroys part of [inaudible] because the designers say and I have quotes in my longer talk of designers thing. Well, they just don't understand. So, why should I have to comply with what they want because clearly, I am a good ethical decision maker and I know what's going in the system. I will do a good job myself, right? And so, this starts to creep into questions to professional ethics. There has not been a strong [inaudible]. BNSF is actually working hard onto building an ethics educational component into computer science. And it's starting to happen but it hasn't been--it's not like law school and business school where you have to take the ethics class library school actually. And so, this kind of sense of professional ethics that you're not the sole decision maker but that, you know, your field might have some standard that would apply here is something I think IRBs could encourage or be part of. That said, I think we need a lot more technical representation IRBs. I would like to see CS folks serving. I think they're going to love me for this, right? [Laughter] Hey, you should be on the IRB. But, you know, I think that that dialogue could happen a lot more effectively. And this is just also UCLA's IRB, right? I don't have experience with others. I haven't done this yet here. So, it will be interesting to see that yeah, I spoke to the IRBs like a national conference every year. And so, I spoke to folks about this like I'm trying to beef up this relationship between technical folks and the IRB because I think we're going to see more and more of these kinds of projects, house and CS departments. And that IRB can be very influential.

[Pause]

>> Well, I think we should thank our speaker.

>> Thank you.

[Applause]