Debate

Networked Privacy

danah boyd

Microsoft Research, USA. dmb@microsoft.com

Privacy rhetoric often focuses on the individual (Solove 2008). Computer systems are designed to give individuals control over their “personal” data while legal narratives often speak of individual harm and informed consent by individuals. Models that go beyond the individual often focus on groups (e.g., access-control lists that support bounded entities) or articulated lists of others (e.g., “joint rights” models that focus on multiple defined entities). But what are the implications of privacy in a networked world where boundaries aren’t so coherently defined and when entities aren’t so easily articulated?

Curious about what secrets might be hidden in my DNA, I decided to spit in a tube and turn my DNA over to the genetic testing service 23andMe. What came back was fascinating: hints that my ancestors might have origins that differ from the family narrative, and disease probabilities that suggest that family medical stories are either inaccurate or statistically curious. Through this test, I learned information about myself, but I also learned information about members of my family. Furthermore, by choosing to subject my DNA to this testing process, I didn’t just reveal data about myself; I gave away data that provides insights into my mother, brother, grandparents, and even children that I don’t yet have. I never asked my future grandchildren for permission to offer their data to a scientific database. I made a decision about the privacy of my data that affects numerous people who are implicated but who have no say. And, in doing so, I learned information about them that they may not wish to know, let alone have me know.

Our data—and with it, our privacy—is increasingly networked. What we share about ourselves tells heaps about other people. Sometimes, as with DNA data, we’re linked by immutable factors. In other situations, the connection is social or locational. I can’t even count the number of photos that were taken by strangers with me in the background at the Taj Mahal. And my friends often post photos of me with them without asking my permission. Yet, there’s also a third layer of connection. Our data also provides a probabilistic image of who we are based on comparisons to other people.

In the early days of personalization based on simplistic usage patterns, there were innumerable moments in which personal portraits were wrong in funny ways. Like many others, I had my “My TiVo thinks I’m gay” story. As an ethnographer who studies teen culture, I travel all over the United States where I grab Wi-Fi from truck stops and sleep in Motel 6s and hang out at too many fast food joints. A month into an intensive tour of the Midwest, I couldn’t help but notice that I was getting fascinating advertisements for grill guards, CB radios, and other accessories for my non-existent big rig. Google had pegged me for a long-haul trucker.

Machine learning algorithms have improved tremendously over the last decade and personalization has become so ubiquitous that most people are unaware of how their data is aggregated with others to

boyd, d. 2012. Networked Privacy. *Surveillance & Society* 10(3/4): 348-350.

<http://www.surveillance-and-society.org> | ISSN: 1477-7487

© The author(s), 2012 | Licensed to the Surveillance Studies Network under a [Creative Commons Attribution Non-Commercial No Derivatives license](#).

construct portraits of individuals that predict their interests based on others' habits. Our interpreted selves aren't simply the product of our own actions and tastes; they're constructed by recognizing similar patterns across millions of people. How machines see us depends on how our data connects to others. The tastes and interests of people who don't yet exist within systems can be easily predicted based on the patterns of others. And, when machines have access to a person's social network, the predictions are even stronger. We aren't as unique as those of us in the West might want to believe; we are the product of the people we know and the socio-cultural environment in which we are situated.

Control and Interpretation

Any model of privacy that focuses on the control of information will fail. Even achieving true control is nearly impossible because control presumes many things that are often untenable. Control assumes that people have agency, or the power to assert control within a particular situation. Control assumes that people have the knowledge and skills to truly control information. And control assumes that people understand the situation well enough to make informed decisions about what should be shared to whom and when. Furthermore, in a networked age, a reasonable amount of control is not enough; control has to be absolute control. One slip-up or data leakage and whatever was once protected can easily enter into a networked public where it may enter broader databases, be aggregated with other data, and circulate. In a networked world, data is more persistent, replicable, searchable, and scalable than ever before. Trying to achieve perfect control will only lead to frustration.

If we cannot rely on control to achieve privacy in a networked age, how then can we think about networked privacy? Focusing on articulated lists of relevant actors and trying to obtain rights from affected persons is bound to fail. There's no way that consent from my not-yet-alive grandchildren is possible. This suggests that focusing on permission at the data acquisition level is not going to be viable. We need to understand privacy in context (Nissenbaum 2009).

Many of the teenagers I have interviewed have given up on controlling access to content (Marwick and boyd 2011). Nosy parents and friends who are "in their business" challenge their ability to have agency within social situations and, even when they think they understand the boundaries of a particular online social setting, they feel as though things change so fast that it becomes impossible to actually achieve control. Some try to achieve privacy through technical means or by simply demanding that adults "keep out," but I'm fascinated by the diverse groups of teens who have taken a different tactic. Rather than trying to limit access to content, they work to limit access to meaning. They use pronouns and in-jokes, cultural references and implicit links to unmediated events to share encoded messages that are for all intents and purposes wholly inaccessible to outsiders. I call this practice "social steganography." Only those who are in the know have the necessary information to look for and interpret the information provided.

Needless to say, even if teenagers' efforts to achieve social privacy keep their parents in the dark, they don't stop algorithmic interpretation—and misinterpretation—of their interactions. They are still subject to advertising and personalization based on what they post and they may be rejected by colleges and have limited job opportunities based on the interpretations of machines or people. But their efforts to achieve privacy without relying on the control of information are still an important signal. Not only have the next generation not given up on privacy, but they're actually trying to find ways to achieve privacy in networked publics.

If we assume that the future of data is networked and that we can no longer rely on control of data to achieve privacy, it becomes imperative to look for alternate models for dealing with networked privacy. My guess is that we need to start by shifting to a model that focuses on usage and interpretation. Who has the ability—and the right—to interpret what data they think they see? In which domains is it acceptable to

discriminate based on interpretation of aggregate data? What are the mechanisms by which people can challenge how they've been interpreted?

Looking Backwards, Moving Forward

In Geoff Bowker and Leigh Star's (1999) seminal text, "Sorting Things Out," they detail the classification schemes used to describe people's race under apartheid in South Africa. The population was divided into four racial categories—White, Black, Asian, and Colored—and their lives were configured accordingly. Housing, transport, and jobs were all segregated by race. But imagine what happens to families when a parent is categorized differently from a child or when siblings fall under different categories. Families cannot live together, kids aren't able to go to school together, and parents cannot take their children to the doctor. Momentarily putting aside all of the injustices of apartheid, these categories become untenable as an organizational structure because people's lives are interwoven with others' lives.

The challenges of networked privacy are not new issues, but social media and networked culture magnifies them in significant ways. The data that underpins networked sociality and algorithmic life connects people across numerous axes time and time again. The future is only going to be more networked, more interwoven, more of a gnarly hairball that's impossible to untangle without harsh cleaving. Expecting that people can assert individual control when their lives are so interconnected is farcical. Moreover, it reinforces the power of those who already have enough status and privilege to meaningfully assert control over their own networks.

In order to address networked privacy, we need to let go of our cultural fetishization with the individual as the unit of analysis. We need to develop models that position networks, groups, and communities at the center of our discussion. And we need to find a way to empower people by freeing them to share in ways that don't negatively affect how others' lives are interpreted.

References

- Bowker, Geoffrey and Susan Leigh Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Marwick, Alice and danah boyd. 2011. Social Privacy in Networked Publics: Teens' Attitudes, Practices, and Strategies. Presented at Oxford Internet Institute's A Decade In Internet Time, September.
- Nissenbaum, Helen. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Palo Alto: Stanford Law Books.
- Solove, Daniel. 2008. *Understanding Privacy*. Cambridge, MA: Harvard University Press.