The Privacy Preferences of Americans

Lee Rainie

The promises of "big data" seem nearly boundless. Among them: new efficiencies and conveniences in daily life and economic activities; predictive modeling that will helpfully meet human needs; deeper self-awareness and beneficial behavior change as people "quantify" their lives; fuller understanding of people's social interactions; extensive analytics that can make research of all kinds more insightful (especially in the domains of health and wellness); vastly more inputs that give richer pictures about the quality of the environment; "smarter" communities, homes, and workplaces that are safer, cleaner, and cheaper; and more "transparent" institutions that are responsive to their stakeholders.

At the same time, the perils of big data are also clarifying. The collection and analysis of all this data pose threats to people's privacy; potentially increase social and political divisions as some groups find it easy to find ways to exploit the advantages of big data and others struggle to navigate a data-saturated world; possibly cut deeply into fundamental human agency as people find their choices proscribed by algorithms that are applied to the datasets; and conceivably overwhelm public institutions and community social systems in their capacity to set rules and form norms around how the data are used.

The promise and peril of big data frame the dilemma of the networked age, as people function in loose, far-flung personal networks. There are considerable incentives in such a world to disclose and share a great deal of information. Doing so helps people deepen friendships, form communities, become more successful economic agents, and accomplish their goals. Sharing information about oneself and soliciting material from others potentially helps networked individuals realize who can help them when they have problems to solve or decisions to make. In a world of networked individuals, the balance sheet of calculations people make

L. Rainie (✉)
Internet, Science, and Technology Research, Pew Research Center, 1615 L Street, NW, Suite 800, 20026 Washington, DC, USA
e-mail: Lrainie@pewresearch.org

J. Collmann and S.A. Matei (eds.), *Ethical Reasoning in Big Data*,
Computational Social Sciences, DOI 10.1007/978-3-319-28422-4_3

about disclosure has changed from prior eras that were more characterized by close, tight-knit social units. They know there are benefits to personal disclosure. And it takes more effort and calculation to remain masked and hidden.

While most civilians do not know they are living at the dawn of the big data era, they surely have expressed their views over the years about the value they place on personal privacy and the ways in which they act when they are asked to share personal information. The long-standing research on Americans and privacy illuminate the degree to which people feel there should be limits on the collection of personal information and the ways in which companies and the government should behave once their personal data have been collected. Moreover, this privacy research can help the architects of big data research create ethical rules and methodological schemes for using big data in ways that people will accept and might willingly embrace in their lives. The alternative is that the research community can ignore Americans attitudes and behaviors around privacy. That would surely deepen public cynicism and distrust, and, possibly, people's willingness to share their information in ways that could produce societal benefits.

The Pew Research Center has conducted surveys and extensive focus groups and interviews around privacy and disclosure issues since 2000 (Pew Research 1 2000). Over that period, it has gained six fundamental insights about the attitudes and behaviors of American adults that could be applied to the collection and analysis of big data.

1. **The balance of forces has shifted in the networked age. People are now "public by default and private by effort," in the words of communications scholar danah boyd.**[1]

Americans have a strong sense that many entities are gathering information about them. In the internet era, data gathering is a persistent and pervasive practice. People have long been familiar with the idea that they under *sur-veillance* regimes where important and powerful organizations are monitoring them. That process continues with new fervor today as corporations, law enforcement agencies, and government intelligence analysis marches on. At the same time, new forces are unfolding in digital times. After the emergence of social media like email, blogs, Facebook, and Twitter, ordinary citizens are increasingly aware that they themselves have the capacity to monitor the activities of those more powerful. This could be called *sous-veillance* and it underlies many of the efforts to make major organizations more open and accountable. Finally, people know peer-to-peer *co-veillance* allows them to chronicle the environment around them, including those in their vicinity or whose activities are posted in social media newsfeeds.

This systemic monitoring and documentation feels like a "part of everyday life—neither sinister, nor benign. It's the way things are and most of the time it doesn't occur to me to think about it," said one online focus group participant in a Pew Research privacy study (Pew Research 7 2015). Indeed, large numbers of internet

[1]Boyd (2014).

users (at this point, 87 % of the adult U.S. population) know that key pieces of their personal information are available about them online, ranging from photos of them to their political views and affiliations. At the same time, growing numbers of internet users (50 %) say they are worried about the amount of personal information about them that is online—a figure that has jumped from 33 % who expressed such worry in 2009 (Chart 1).

Still, even as they recognize that a lot of information is collected about them, Americans are anxious in basic ways about what this means about their privacy. First, they lack confidence that they have control over their personal information: 91 % of adults in the 2014 Pew Research Center survey "agreed" or "strongly agreed" that consumers had lost control over how personal information was collected and used by companies (Pew Research 6 2015). Second, they express a consistent lack of confidence about the security of everyday communications channels—particularly when it comes to the use of online tools (Pew Research 3 2014). For example:

- 68 % feel insecure using chat or instant messages to share private information.
- 58 % feel insecure sending private info via text messages.
- 57 % feel insecure sending private information via email.

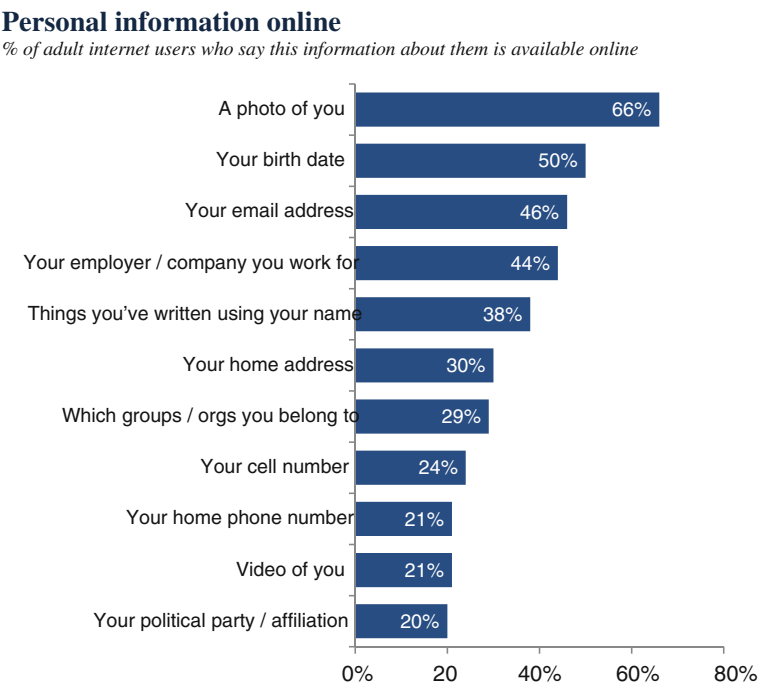

Chart 1 Personal information online. *Source* Pew Research Center survey July 11–14, 2013

- 46 % feel "not very" or "not at all secure" calling on their cell phone when they want to share private information.
- 31 % feel "not very" or "not at all secure" using a landline phone when they want to share private information.

Third, they exhibited a deep lack of faith in organizations of all kinds (public or private) in protecting the personal information they collect. Only tiny minorities say they are "very confident" that the records maintained by these organizations will remain private and secure (Pew Research 6 2015).

- Just 6 % of adults say they are "very confident" that **government agencies** can keep their records private and secure, while another 25 % say they are "somewhat confident."
- Only 6 % of respondents say they are "very confident" that **landline telephone companies** will be able to protect their data and 25 % say they are "somewhat confident" that the records of their activities will remain private and secure.
- **Credit card companies** appear to instill a marginally higher level of confidence; 9 % say they are "very confident" and 29 % say they are "somewhat confident" that their data will stay private and secure.

Online service providers are among the least trusted entities when it comes to keeping information private and secure (Pew Research 6 2015):

- 76 % of adults say they are "not too confident" or "not at all confident" that records of their activity maintained by the **online advertisers** who place ads on the websites they visit will remain private and secure.
- 69 % of adults say they are not confident that records of their activity maintained by the **social media sites** they use will remain private and secure.
- 66 % of adults say they are not confident that records of their activity maintained by **search engine providers** will remain private and secure.

Implications for big data: Americans' distrust in the organizations in charge of protecting communications emerges in the same season as several major corporations announced that key personal information about account holders had been breached and several months after Edward Snowden, a contract worker for the National Security Agency (NSA), leaked information to international news media about widespread NSA surveillance of Americans' phone and email records. It is the period in which public attitudes about privacy issues demonstrated a more urgent tone than in previous years. Those hoping to use big data would be wise to make sure that data-sharing arrangements they have with other organizations are secure and that there be mechanisms to disclose clearly the ways in which the data will be used and who will have access to it. Moreover, Americans would be comforted to know if there were data breaches or successful efforts to use the data to re-identify participants. They would also appreciate a process to gain redress from harms caused by data breaches or re-identification efforts.

2. **Privacy is not binary—either on or off—for most Americans. The context and conditions of information transactions matters**.

People's decisions about whether to disclose information about themselves and how to disclose the data are highly context dependent. It depends on what personal information is at issue; who is watching or capturing the data; and what the "value proposition" for personal disclosure. A stark affirmation of this was evident in a 2015 Pew Research Center survey that posed several possible scenarios with Americans and asking whether they would accept the tradeoff of sharing personal information for a good or service (Pew Research 7 2015). The survey covered six possible scenarios and the overwhelming majority of adults—83 %—were open to at least one information-sharing scenario. But only 4 % were open to every scenario; in other words, their answers to whether they liked these information transactions were "it depends." Two examples from the survey illustrate this.

One scenario was posed this way: *Several co-workers of yours have recently had personal belongings stolen from your workplace, and the company is planning to install high-resolution security cameras that use facial recognition technology to help identify the thieves and make the workplace more secure. The footage would stay on file as long as the company wishes to retain it, and could be used to track various measures of employee attendance and performance. Would this be acceptable to you, or not?* Some 54 % said the installation of surveillance cameras would be acceptable under these conditions; 24 % said it would not be acceptable; and 21 % said their views would depend on more details and context for the scenario.

A second scenario drew a very different response: *Your insurance company is offering a discount to you if you agree to place a device in your car that allows monitoring of your driving speed and location. After the company collects data about your driving habits, it may offer you further discounts to reward you for safe driving. Would that scenario be acceptable to you or not?* In this case, only 37 % said the bargain—my driving information in return for possible discounts—was acceptable to them; while 45 % said it was not acceptable; and 16 % said their agreeing to the deal would depend on their learning more details.

In each case, something of potential value was being offered respondents in return for the potential collection of personal information, but different people were comfortable with different deals. The conditions of the offer mattered to them and they weighed the value proposition differently, depending on the circumstances.

Part of the bargain people weigh when they are deciding if they like an information deal or not is how they feel about the party on the other side of the deal.

In the hierarchy of privacy concerns, Americans are not anxious to be known to or surveilled by hackers (the black-hat kind) or advertisers (Pew Research 2 2013). The next most sensitive area of sensitivity for people involves social surveillance. It is a more top-of-mind concern to people than government surveillance. People are more likely to experience or witness reputational privacy breaches within their own networks than they are to be aware of how the government's access to their data might negatively impact their lives (Chart 2).

One last example of how context colors Americans' views on privacy involves government surveillance programs themselves (Pew Research 5 2015). Far from being opposed to surveillance, the public generally believes it is acceptable for the

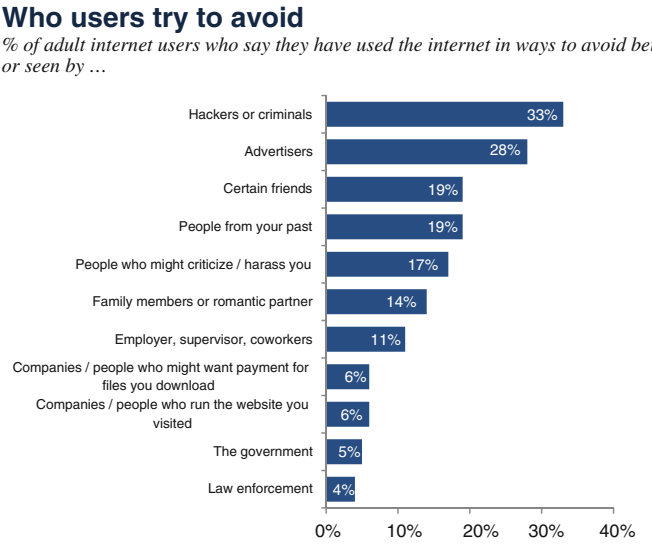

Chart 2 Who users try to avoid. *Source* Pew Research Center survey July 11–14, 2013

government to monitor many others, including foreign citizens, foreign leaders, and American leaders:

- 82 % say it is acceptable to monitor communications of suspected terrorists.
- 60 % believe it is acceptable to monitor the communications of American leaders.
- 60 % think it is okay to monitor the communications of foreign leaders.
- 54 % say it is acceptable to monitor communications from foreign citizens.

Yet, 57 % say it is *unacceptable* for the government to monitor the communications of U.S. citizens. At the same time, majorities support monitoring of those particular individuals who use words like "explosives" and "automatic weapons" in their search engine queries (65 % say that) and those who visit anti-American websites (67 % say that).

Implications for big data: Americans' views on these issues suggest there are ways for analysts of big data to make the case for their work. People are not instinctively opposed to data collection and use. They want to see what the tradeoff is, and under the right circumstances will accept the bargain. This might put some burden on big data analysts to make the case for their work and the benefits that will emerge from it, but it suggests that many are open to sharing information and being tracked if they understand what the upside of research is.

3. **Personal control and agency matter a lot to people**.

If the traditional American view of privacy is the "right to be left alone," the 21st Century refinement of that idea is the right to control their identity and information.

They understand that modern life won't allow them to be "left alone" and untracked, but they do want to have a say in how their personal information is used. There are several pieces of evidence for this. First, 86 % of internet users have tried to be anonymous online at least occasionally and 55 % of internet users have taken steps to avoid observation by specific people, organizations, or the government (Pew Research 2 2013).

At the attitudinal level, Americans say that being in control of who gets information about them. Even though they acknowledge that the boundary line between private and public information has sharply shifted toward "publicness" as the default condition of the modern moment, Americans continue to insist that they care about what happens to their personal information once it has been collected: 74 % say it is "very important" to them that they be *in control of who can get* information about them and 65 % say it is "very important" to them *to control what information is collected about them* (Pew Research 6 2015) (Chart 3).

Implications for big data: This basic American attitude about privacy is difficult to apply to big data. In many cases, the data are collected in ways that are not easy for users to control. All the greater burden falls, then, on researchers to show what they are doing and how they arrive at the insights they do. Perhaps applications of new online trust-building systems might help civilians assess and maybe contribute to the findings—that would including things like Reddit-style up- or down-voting schemes or allowing participant comments on the findings. That might give them a sense of agency and stake in the insights that the data generate.

4. **The young are more focused on networked privacy than their elders**.

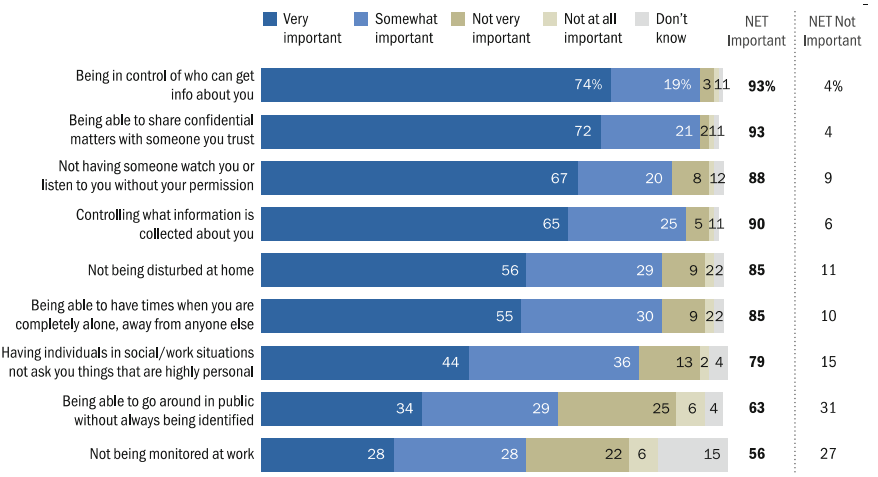

Chart 3 Americans hold strong views about privacy in everyday life. *Source* Pew Research Center survey January 27–February 16, 2015

Throughout the Pew Research data, those ages 18–29 are more likely than older adults to say they have paid attention to privacy issues, to have taken steps to protect their privacy, and to have suffered some kind of harm because of privacy problems (Pew Research 2 2013):

- They take steps to limit the amount of personal information available about them online—44 % of young adult internet users say this.
- They change privacy settings—71 % of social networking users ages 18–29 have changed privacy settings on their profile to limit what they share with others online.
- They delete unwanted comments—47 % social networking users ages 18–29 have deleted comments that others have made on their profile.
- They remove their name from photos—41 % of social networking users ages 18–29 say they have removed their name from photos that were tagged to identify them.

It is likely that this extra attention to personal reputation emerges from the reality that younger adults are more likely to know that personal information about them is available on line and to have experienced privacy problems. For instance, those 18–29 are more likely than older adults to have had an email or social media account hijacked or had online difficulties place them in physical danger (Chart 4).

The larger point here is that different people have different generational, cultural, or social circumstances which inform their attitudes and behaviors around disclosure or privacy protection.

Implications for big data: There is not a one-size-fits-all set of policies and solutions about how to handle big data. Researchers would help themselves by

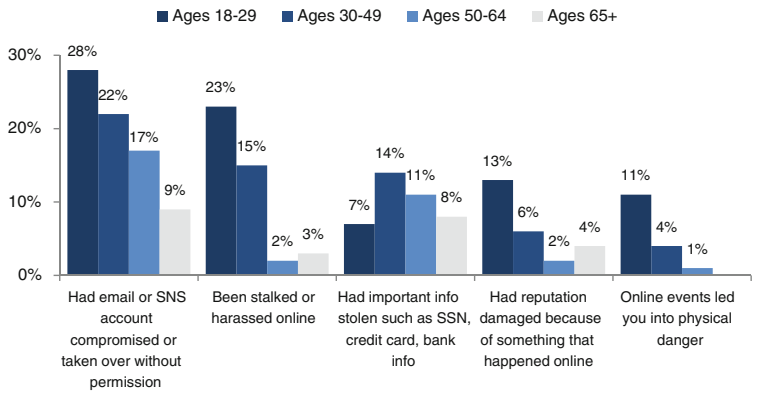

Chart 4 Young adults are the most likely to have had several—but not all—major problems with personal information and identity. *Source* Pew Research Center survey July 11–14, 2013

caring about different demographic, cultural, and generational sensitivities on privacy issues. One way to address this would be outreach to special segments of the population with assurances that the issues that matter most to them are on researchers' minds as they do their analysis and render their findings.

5. **Many know they do not know what is going on when it comes to the nature and scope of data collected about them**.

When it comes to their own role in managing the personal information, most adults are not sure what information is being collected and how it is being used. "I wouldn't know where to begin if I ever wanted to get to the bottom of what kind of profiles exist on me," one middle-aged suburban woman said in a Pew Research online focus group (Pew Research 7 2015). A 63-year-old man added: "Every organization wants at least pieces of me—what I buy, who I vote for, what movies I go to, what music is on my playlist, the medicines I take, how much energy I use, even who my friends are. It is impossible to imagine one part of my life that is not being documented by one company or another."

At the same time, many express a desire to take additional steps to protect their data online. When asked if they feel as though their own efforts to protect the privacy of their personal information online are sufficient, 61 % say they feel they "would like to do more," while 37 % say they "already do enough" (Pew Research 3 2014). When they want to have anonymity online, few feel that is easy to achieve. Just 24 % of adults "agree" or "strongly agree" with the statement: "It is easy for me to be anonymous when I am online."

Implications for big data: These findings underscore how fragile the relationship between big data analysts and the public are. People do not like surprises and will likely be unhappy if their data were used in ways they did not anticipate or that seem "out of the blue."

6. **Americans believe changes in law could make a difference, though their exact policy preferences are not fully clear**.

In the midst of all this uncertainty and angst about privacy, Americans are generally in favor of additional legal protections against abuses of their data. Some 68 % of internet users believe current laws are not good enough in protecting people's privacy online (Pew Research 2 2013); and 64 % believe the government should do more to regulate advertisers, compared with 34 % who think the government should not get more involved (Pew Research 3 2014).

When asked to think about the data the government collects as part of anti-terrorism efforts, 65 % of Americans say there are not adequate limits on "what telephone and internet data the government can collect."[2] Just 31 % say they believe that there are adequate limits on the kinds of data gathered for these

[2]Due to differences in the method of survey administration and questionnaire context, these findings are not directly comparable to previous Pew Research telephone surveys that have included a version of this question.

programs. The majority view that there are not sufficient limits on what data the government gathers is consistent across all demographic groups. Those who are more aware of the government surveillance efforts are considerably more likely to believe there are not adequate safeguards in place.

Relatedly, there is a striking divide among citizens over whether the courts are doing a good job balancing the needs of law enforcement and intelligence agencies with citizens' right to privacy: 48 % say courts and judges are balancing those interests, while 49 % say the courts are not (Pew Research 5 2015).

Implications for big data: It is easy to imagine that analysts of big data would gain public approval by helping people understand what is going on in the world of hyper-data collection and providing strategies and tools to help Americans regain a sense they are more knowledgeable about this environment and more competent to navigate it.

7. Conclusion: The future of privacy

When the Pew Research Center canvassed hundreds of technology experts and pundits about the fate of privacy in the coming decade, there were several themes in their predictions about the future that are relevant to the long-term viability of big data research (Pew Research 4 2014): First, Few individuals will have the energy or resources to protect themselves from "dataveillance." Privacy protection will likely become a luxury good. There will be technology tools and marketplace solutions that will be embraced by higher socioeconomic groups, but the capacity of average citizens to achieve privacy will diminish. Second, the prospect of achieving by-gone notions of privacy will become more remote as the Internet of Things arises and people's homes, workplaces, and the objects around them will "tattle" on them. Third, living a public life will be the new default. People will get used to this, adjust their norms, and accept more sharing and collection of data as a part of life—especially Millennials and the young people who follow them. Problems will persist and some will complain but most will not object or muster the energy to push back against this new reality in their lives.

In a way this is good news for the expansion of big data initiatives and those who would use them. Still, those who take advantage of these new realities bear the risk of pushing things too far and engendering a backlash if they do not accommodate insights from Americans' long history of asserting that privacy matters, there are ways that it is directly connected to social trust which is the bonding agent for any society, and there are parts of life that are best left unmonitored and protected from prying eyes.

References

Boyd, D. (2014) *It's Complicated*. Yale University Press. Downloaded on November 7, 2015 at http://www.danah.org/books/ItsComplicated.pdf

Pew Research 1. (2000, August 20). Trust and privacy online. Retrieved on November 2, 2015, from http://www.pewinternet.org/2000/08/20/trust-and-privacy-online/

Pew Research 2. (2013, September 5). Anonymity, privacy, and security online. Retrieved on November 7, 2015, from http://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online/

Pew Research 3. (2014, November 12). Public perceptions of privacy and security in the Post-Snowden Era. Retrieved on November 7, 2015, from http://www.pewinternet.org/2014/11/12/public-privacy-perceptions/

Pew Research 4. (2014, December 18). The future of privacy. Retrieved on November 7, 2015 at http://www.pewinternet.org/2014/12/18/future-of-privacy/

Pew Research 5. (2015, March 16). Americans' privacy strategies Post-Snowden. Retrieved on November 7, 2015 at http://www.pewinternet.org/2015/03/16/americans-privacy-strategies-post-snowden/

Pew Research 6. (2015, May 20). Americans' attitudes about privacy, security, and surveillance. Retrieved on November 7, 2015 at http://www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/

Pew Research 7. (2015, December). Privacy and information sharing (forthcoming).