A Reflection on the History of Corporate and Government Data Collection

March 20, 2019

Privacy has been a concern since the early days of computing. Over 50 years ago the political scientist Ithiel de Sola Pool predicted "By 2018 it will be cheaper to store information in a computer bank than on paper; files will be computer-stored and fantastically manipulative. We have already seen the beginning of this trend reflected in public alarm about a national data center. The alarm is over the issue of privacy, ... By 2018 the researcher sitting at his console will be able to compile a cross-tabulation of consumer purchases (from store records) by people of low IQ (from school records) who have an unemployed member of the family (from social security records)." [1] 2018 has come and gone and the dream of online privacy no longer exists given the striking normalcy of modern data collection. The history of privacy and computing in the United States goes back to the beginning and stems from many important events relating to both corporations and the government. Slowly, the amount of data collected has increased and the expectation of privacy has degraded, often through shocking revelations that seem unsurprising years later. Understanding these events sheds light on how the current state of privacy came to be and where it may go in the coming years. Here, I will reflect on the history of this topic and its major events.

Government surveillance has always occurred but gradually expanded in the last two decades to what it is today. Unsurprisingly, it has maintained massive amounts of electronic data on United States citizens since as early as the 1960s. Early critics include Vance Packard who remarked in *The Naked Society*, "There are banks of giant memory machines that conceivably could recall in a few seconds every pertinent action — including failures, embarrassments or possibly incriminating acts — from the lifetime of each citizen," [2]. Around this time the government announced plans to create a National Data Bank from hundreds of individual federal databases. Outrage stopped this and because of privacy concerns congress passed the Fair Credit Reporting Act and Privacy Act in the 1970s. Both increased transparency but did nothing to stop data collection by the government or corporations [3]. Government data collection greatly increased after the September 11th terrorist attacks on the World Trade Center. As a direct consequence of the attacks the Partiot Act expanded the FBI's ability to surveil domestic communications without probable cause [4].

Warrantless wiretapping by the NSA was uncovered and reported on by *The New York Times* in 2005 [5]. In 2008 FISA Amendments Act was passed further expanding government surveillance powers and granting immunity to telecom companies assisting in warrantless surveillance [6]. This is what allowed many of the programs uncovered in leaks by Edward Snowden in 2013. These revelations included massive NSA spying and surveillance programs such as PRISM, which collects internet communications from large companies such as Google. Many early fears of the 1960s and 1970s eventually came to silently pass. These days it is relatively accepted that the government collects data about its citizens, including calls and metadata, as most would would happily trade privacy for security. Gone is the outrage at the thought of creating a national database of the country's citizens. If this were ever to be reversed it would require massive effort on the part of the whole population, which seems increasingly less likely to manifest as the years go on.

On the other hand of modern surveillance in the United States is the data collected by major corporations such as Google and Facebook. Amid early fears of government data gathering, the federal government did little to curtail collection in industry [2]. For example, the Lotus Development Corporation in 1990 attempted to release for sale a database containing the names, addresses, demographics, and purchasing behavior of over 120 million Americans, and was only stopped by significant public backlash [7]. Over time, several technologies allowed the tracking of users through their web activity. In 1994 Netscape invented browser cookies, which can store identifying information about a user. One of the main motivators behind corporate data collection was targeted advertising. There was a time when ads shown on webpages were purely chosen based on the content of the website. In 1999 the large ad company DoubleClick attempted to target ads based on real names using data it would have acquired by merging with a major data broker. They faced backlash from privacy groups that petitioned the FTC, so they sold the data broker at the time. They were, however, eventually bought by Google in 2008 which eventually amended its privacy policy to allow exactly this type of tracking [3]. Google's announcement that it would integrate data from all of its services was a big deal, as it became obvious how much data the company was collecting which it had previously kept separate. These weren't the only privacy issues the company faced. It was revealed that in 2007 the company collected sensitive personal information from Wi-Fi networks using its Street View cars [8]. Several policy changes by Facebook signalled to its users just how little control they had over their own data. In 2006 the company launched the News Feed feature which broadcasted updates about changes to a user's profile to their friends, and in 2009 they changed their terms of service to allow Facebook to use anything uploaded to the site for anything [6]. Although these were highly controversial at the time, they seem relatively normal by today's standards. As with government surveillance, huge scandals have come to be accepted, and sentiment against corporate data collection has gradually gone down. The rise of social media starting in the 2000s has normalized making one's life completely public, and the idea that privacy is desirable may seem absurd in that light.

Modern privacy on the internet only tenuously exists. This has been a long time in the making. Government surveillance has been spurred on by the War on Terror and in the corporate case it has traditionally been advertising. However, new questions and motivations arise constantly. Encryption is now a sore point for the government, and Apple vs the FBI raised questions about the reach of government power. The Cambridge Analytica scandal brought to light just how easily data can be harvested and how it can be used for political gain. For companies, collecting data offers huge opportunities for machine learning and artificial intelligence. The balance between privacy and progress is something that we should always consider, but the rise of modern surveillance capitalism and all of history points towards sliding away from privacy.

References

[1] Ithiel de Sola Pool. *Toward the Year 2018.* Cowles Education Corporation, 1968.

[2] Margaret O'Mara. "The End of Privacy Began in the 1960s". In: *The New York Times* (2018).

[3] Louise Matsakis. "The WIRED Guide to Your Personal Data (and Who Is Using It)". In: *Wired* (2019).

[4] *The Patriot Act: What Is the Proper Balance Between National Security and Individual Rights?* Constitutional Rights Foundation. URL: `http://www.crf-usa.org/america-responds-to-terrorism/the-patriot-act.html`. (accessed: 03.20.2019).

[5] James Risen and Eric LichtBlau. "Bush Lets U.S. Spy on Callers Without Courts". In: *The New York Times* (2005).

[6] Charles Mahtesian. *Privacy In Retreat, A Timeline.* NPR. 2013. URL: `https://www.npr.org/2013/06/11/190721205/privacy-in-retreat-a-timeline`.

[7] Mary J. Culnan. "An Issue of Consumer Privacy..." In: *The New York Times* (1991).

[8] *Notice of Apparent Liability for Forfeiture.* Federal Communications Commission. 2012. URL: `https://transition.fcc.gov/DA-12-592A1.pdf`.