DOI:10.1145/3300224

Carl Landwher

Privacy and Security 2018: A Big Year for Privacy

Retracing the pivotal privacy and security-related events and ensuing issues from the past year.

THE YEAR 2018 may in the future be seen as a turning point for privacy incidents and associated privacy-policy concerns. In March, the Cambridge Analytica/Facebook incident opened many eyes to the unanticipated places personal data reaches, and it continues to generate repercussions.[4] Google shut down its struggling Google Plus social networking system in October, after announcing it had exposed the data of approximately 500,000 users,[15] only 1% as many as involved in the Cambridge Analytica case. Facebook revealed another data breach in October, this one affecting a reported 29 million users.[14]

The open GEDmatch genomics database, developed for genealogy research, was used by police and genetics experts to identify alleged murderers in two "cold cases" and several other crimes.[8] The site's founders, at first uncomfortable with its use by law enforcement, seem to now be more comfortable with it. Researchers subsequently estimated that today approximately 60% of Americans of European descent could be identified from their DNA, even if they had never registered their DNA with any site.[6] Further, they forecast the figure will rise to 90% in only two or three years.[9]

The John Hancock Life Insurance Company announced it would sell life insurance only through "interactive" policies that provide financial incentives to track policyholders' fitness and health data through wearable devices and smartphones;[2] and the latest Apple Watch can take your electrocardiogram.

Innovation has its downside and loss of privacy is not easy to remedy.

On the policy front, the long-awaited implementation of the EU's General Data Protection Regulation (GDPR) in late May[12] triggered many reviews of corporate data privacy policies globally. These revisions required untold numbers of clicks by users asked to acknowledge policy changes.

About a month later, under threat from a strong privacy ballot initiative, California passed the California Consumer Privacy Act of 2018.[1] It incorporates some features of the GDPR and gives California consumers the right to know what personal information businesses have about them. Consumers control whom the information is shared with or sold to, and can request that information be deleted. This law begins to require consumer-facing businesses to live up to some of the Fair Information Practice Principles that were mandated for U.S. government systems (but not commercial enterprises) by the Privacy Act of 1974.[13]

"Personal information" in the California law is broadly defined. It includes biometric information, but

IMAGE BY ANDRIJ BORYS ASSOCIATES/SHUTTERSTOCK

also "information that identifies, relates to, describes, is capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household." The law enumerates almost a dozen categories of personal information, but exempts "publicly available" information (also defined in the law). Implementation details must be worked out before the law takes effect in 2020. The law has triggered national discussion and legislative proposals in other states.

Also in June, the U.S. Supreme Court handed down a decision in *Carpenter v. U.S.*[3] This decision represents a notable limitation of the "third-party doctrine" wherein a government request to a third party to produce data an individual has voluntarily surrendered to it does not require a warrant. This doctrine, in place in the U.S. since 1979, is the basis for the idea that once a consumer surrenders data to a company as part of a transaction, the consumer loses any expectation of privacy for that data. As such, it has had major implications for, among other things, Internet-based transactions of all kinds.

The 5-4 decision had four separate dissenting opinions. The majority characterized the decision as "narrow" because it did not overturn the third party doctrine per se. Rather, it recognized the information in this case (cellphone site location information or CSLI records) deserves separate treatment because it is so invasive of "the privacies of life." Further, Justice Gorsuch's dissent argues for overturning the third-party doctrine. He proposes the consumer may well have a property interest in CSLI records held by the telephone company, although that argument was not put forth in this case. Other classes of data routinely collected by third parties could be equally invasive to the privacies of life; more litigation may follow.

In the fall, NIST initiated the development of a privacy framework.[10] Like the cybersecurity framework it released in 2014 and updated in April 2018,[11] the privacy framework is not to be a standard, but a guide to common privacy practices that will help companies assess privacy risk and adopt measures appropriate to the risk. In parallel, the NTIA, also part of the Department of Commerce, released a Request for Comments (RFC) on a two-part approach to consumer privacy: the first part describes desired user-centric privacy outcomes and the second sets high-level goals outlining an ecosystem to achieve those outcomes.[5] The RFC proposes no changes to existing sectoral privacy laws, and, perhaps because it was developed in cooperation with the National Economic Council, the second part on high-level goals emphasizes maintaining "the flexibility to innovate" and proposes to employ a "risk and outcome-based" approach as opposed to one of compliance.

While no one loves red tape, innovation has its downside (remember those innovative collateralized debt obligations?), and loss of privacy is not easy to remedy. Companies already have the option of building in "privacy by design," but relatively few have done so. To me, a requirement

for some baseline of measures seems warranted, even essential.

And Congress, for the first time in years, is showing some interest in drafting comprehensive privacy legislation. This may become a hot topic for the 116th U.S. Congress if public interest continues to be strong.

Returning to the Facebook/Cambridge Analytica incident, this is of immediate importance to those in the computing profession, particularly those conducting research. A researcher with academic connections gained permission from Facebook to put up an app to collect data for research purposes in 2014. This app collected data from some Facebook users who consented to the collection, but also from millions of others without their knowledge or consent. This collection would now violate Facebook's policies, but it was not a violation at the time. The researcher provided this data to Cambridge Analytica, presumably in violation of Facebook's policies. Cambridge Analytica exploited the data for commercial purposes.

The primary issue here is accountability. This was either a violation of the academic's agreement with Facebook, or evidence that the agreements were insufficient to meet Facebook's 2011 consent decree with the Federal Trade Commission (FTC). The privacy of millions of people was violated and the reputation of legitimate academic researchers was tarnished. Facebook apparently had little incentive to hold the researcher and Cambridge Analytica to account. Aware of what happened over a year before the disclosure, Facebook belatedly issued yet another in a long history of privacy apologies.[7]

The FTC and the Securities and Exchange Commission (SEC) are investigating this incident. The SEC could find Facebook liable for failing to inform its shareholders of the incident when discovered. The FTC could find Facebook violated the terms of their 2011 consent agreement by failing to protect their customers' data in accordance with the consent decree. A court could make Facebook pay fines large enough to give it sufficient incentive to enforce the correct privacy policies on researchers and

Congress, for the first time in years, is showing some interest in drafting comprehensive privacy legislation.

those commercial entities that use Facebook data. The U.K. has already levied a fine of £500,000, the largest its legislation allows, but this is unlikely to provide much incentive to a company whose 2017 net income was over $15 billion. The GDPR permits penalties of up to 4% of global revenues, which for Facebook would be well over $1 billion, but the incident occurred before the GDPR took effect. The threat of future fines should give Facebook incentive to prevent recurrence.

Fines levied by the FTC go into the U.S. Treasury. Facebook's users took the risks and are suffering the consequences. Should they be compensated? A penny or dime for each user whose privacy was violated might not be the answer. Perhaps more progress would come from financing investigative journalism or other controls, but might not be within the scope of actions regulatory agencies can take. Imagination might be required to help Facebook hold their clients to account in ways that compensate Facebook users.

Computing professionals involved in "big data" research should pay attention if they wish to gain access to datasets containing or derived from personal information. They must abide by agreements made with dataset providers and remember that exposing data improperly damages public trust in research. Accidental or intentional release of personal data provided for research purposes to anyone else, even if aggregated and anonymized[a] attracts public attention. Researchers who abuse data entrusted to them must expect to be held accountable.

Facebook/Cambridge Analytica was not the first example of abuse, nor will it be the last. The FTC's privacy protection is evidently not working very well. Maybe the time has come for comprehensive privacy legislation focused on aligning corporate incentives so their products provide the privacy people expect and deserve. The California law might be a step in this direction.

A society where individuals are willing to share data for social benefit must make individuals confident that shared data are unlikely to be abused and that abusers can be identified and made accountable. C

a Research into the weaknesses of anonymization or de-identification schemes is needed to understand the limitations of these techniques. Like research that exposes security weaknesses in systems, it must respect the concerns of those whose data is being studied.

References

1. Assembly Bill 375, California Consumer Privacy Act of 2018; https://bit.ly/2z68PCO
2. Barlyn, S. Strap on the Fitbit: John Hancock to sell only interactive life insurance. Reuters (Sept. 19, 2018); https://reut.rs/2DbAq84
3. *Carpenter v. U.S.* 16-402. Decided June 22, 2018; https://bit.ly/2MdFKaE
4. Confessore, N. Audit approved of Facebook policies, even after Cambridge Analytica leak. *The New York Times* (Apr. 19, 2018); https://nyti.ms/2vBniFI
5. Department of Commerce, NTIA, RIN 0660–XC043. Developing the administration's approach to consumer privacy. *Federal Register 83*,187 (Sept. 26, 2018); https://bit.ly/2AErrZP
6. Erlich, Y. et al. Identity inference of genomic data using long-range familial searches. *Science* (Oct. 11, 2018); https://bit.ly/2CadGTP
7. Hempel, J. A short history of Facebook's privacy gaffes. *WIRED* (Mar. 30, 2018); https://bit.ly/2GjTPVD
8. Murphy, H. How an unlikely family history website transformed cold case investigations. *The New York Times* (Oct. 15, 2018); https://nyti.ms/2EnGHhE
9. Murphy, H. Most white Americans' DNA can be identified through genealogy databases. *The New York Times* (Oct. 11, 2018); https://nyti.ms/2pRFhBX
10. NIST Privacy Framework Fact Sheet, Sept. 2018; https://bit.ly/2AcYZOH
11. NIST Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1 (Apr. 16, 2018); https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf
12. Official Journal of the European Union. General Data Protection Regulation. 4.5.2016. (English version); https://bit.ly/2s7bupy
13. Public Law 93-579. Privacy Act of 1974. (Dec. 31, 1974); https://bit.ly/2yKCboa
14. Vengattil, M. and Paresh, D. Facebook now says data breach affected 29 million users, details impact. Reuters (Oct. 12, 2018); https://reut.rs/2CGewZz
15. Wasabayashi, D. Google Plus will be shut down after user information exposed. *The New York Times* (Oct. 8, 2018); https://nyti.ms/2OKoFtH

Carl Landwehr (carl.landwehr@gmail.com) is Lead Research Scientist the Cyber Security Policy and Research Institute (CSPRI) at George Washington University in Washington, D.C., and Visiting McDevitt Professor of Computer Science at LeMoyne College in Syracuse, NY.