Big Data, Big Brother

Author: Currents in Commerce

A version of this article originally appeared in the World Business Academy's monthly publication Currents in Commerce. Join our mailing list to receive Currents in your inbox.

Recent revelations about U.S. and British government surveillance have sparked a heated debate about Big Data. The renewed discussion about the proper balance between privacy and security must include a hard look at companies that track consumers to mine and sell their data—creating a rich database for governments and companies alike.

Silicon Valley’s top tech firms and U.S. spy agencies are “basically in the same business, collecting information on people,” said an ACLU analyst studying technological surveillance. He elaborated: “That creates a weird symbiotic relationship. It’s not that Facebook and Google are trying to build a surveillance system but they effectively have. If they wanted to, Google and Facebook could use technology to tackle the issue, anonymizing and deleting their customers’ information. But that information is how they make their money, so that is never going to happen.”

Companies are making enormous profits by collecting consumer-purchasing data to enhance
their ad targeting. In the outcry about the U.S. government sweep of Verizon phone records, little was said about the fact that Verizon Wireless “sells demographic profiles of customer groups — including ZIP codes for where they ‘live, work, shop and more’ — to marketers.” Yet Verizon and T-Mobile refuse to give customers their own location logs without a subpoena. Amazon’s determination to enhance its ad-targeting capability, partly by gathering more consumer data through expanded product offerings, helped drive its ad revenue up 45% last year over 2011 levels.

The U.S. government argues that the data it collects from its electronic surveillance doesn’t encroach on individual privacy because it is largely “anonymised.” But so-called anonymised data is not really anonymous. A recent study examined mobile phone records of about 1.5 million people in an undisclosed small European country and found that only four different data points on the time and location of a call were enough to identify 95% of the people. A 1997 study of what information is “personally identifiable” showed that a combination of gender, zip code, and birthdate was unique for 87% of the U.S. population.

The futility of relying on the protection of “anonymised data” means that privacy reforms will require limits on companies’ data use and retention, as well as long overdue enforceable “Do not track” preferences.

As Ross Douthout wrote in a recent column: “[E]very looming technological breakthrough, from Google Glass to driverless cars, promises to make our every move and download a little easier to track. Already, Silicon Valley big shots tend to talk about privacy in roughly the same paternalist language favored by government spokesmen. ‘If you have something that you don’t want anyone to know,’ Google’s Eric Schmidt told an interviewer in 2009, ‘maybe you shouldn’t be doing it in the first place.’” Google envisions using its unique access to what you read, write and say to answer “the majority of search queries without you actually asking.”

Not every one minds the loss of privacy that comes with life on the grid, including pop-up ads triggered by personal communications. But many won’t share the awe about the “Programmable World” so evident in the June Wired magazine cover story: “Just as with social networking, the privacy concerns of a sensor-connected world will be fast outweighed by the strange pleasure of residing in it.” The article asserts that “the main existential threat to the Programmable World” is not privacy, but better power sources for all those sensors. Shall we buy this hype?

As the privacy debate unfolds, visit the website WeTheData.com, part of INTEL’s new “data economy” initiative. The website aptly summarizes our present dilemma: “The Arab Spring and Zipcar are part of the same data revolution…. How do we balance anxiety around data with its incredible potential? How do we regain more control over what happens to our data and what is targeted at us as a result? We The Data have the power to topple dictators or empower them.… We The Data can create new forms of social cooperation and exchange, or give us more of the same corporate obsession with better targeted advertising.”

What we don’t need is Big Brother as a public-private partnership.