Your privacy online: Health information at serious risk of abuse

February 23 2015

There is a significant risk to your privacy whenever you visit a health-related web page. An analysis of over 80,000 such web pages shows that nine out of ten visits result in personal health information being leaked to third parties, including online advertisers and data brokers.

This puts users are risk for two significant reasons: first, people's health interests may be publicly identified along with their names. This could happen because criminals get ahold of the information, it is accidentally leaked, or data brokers collect and sell the information. Second, many online marketers use algorithmic tools which automatically cluster people into groups with names like "target" and "waste". Predictably, those in the "target" category are extended favorable discounts at retailers and advance notice of sales. Given that 62 percent of bankruptcies are the result of medical expenses, it is possible anyone visiting medical websites may be grouped into the "waste" category and denied favorable offers.

For individuals, this means profiles are built based on web page visits, potentially resulting in someone being labeled a commercial risk due to the fact that they have used a site like WebMD.com or CDC.gov to look up health information for themselves, a family member, or a friend. Given that data brokers are free to sell any information they collect regarding visits to health websites, those visiting such sites are potentially at risk of being discriminated against by potential employers, retailers, or anybody else with the money to buy the data.

These findings are reported in the article "Privacy Implications of Health Information Seeking on the Web," appearing in the March 2015 issue of *Communication of the ACM*.

Timothy Libert, a doctoral student at the University of Pennsylvania's Annenberg School for Communication wrote the article. He authored a software tool that investigates Hypertext Transfer Protocol (HTTP) requests initiated to third party advertisers and data brokers. He found that 91 percent of health-related web pages initiate HTTP requests to third-parties. Seventy percent of these requests include information about specific symptoms, treatment, or diseases (AIDS, Cancer, etc.). The vast majority of these requests go to a handful of online advertisers: Google collects user information from 78 percent of pages, comScore 38 percent, and Facebook 31 percent. Two data brokers, Experian and Acxiom, were also found on thousands of pages.

"Google offers a number of services which collect detailed personal information such as a user's persona email (Gmail), work email (Apps for Business), and physical location (Google Maps)," Libert writes. "For those who use Google's social media offering, Google+, a real name is forcefully encouraged. By combining the many types of information held by Google services, it would be fairly trivial for the company to match real identities to "anonymous" web browsing data." Indeed, in 2014, the The Office of the Privacy Commissioner of Canada found Google to be violating privacy Canadian laws.

"Advertisers promise their methods are wholly anonymous and therefore benign," Libert writes. "Yet identification is now always required for discriminatory behavior to occur." He cites a 2013 study where individuals' names were associated with web searches of a criminal record, simply based on whether someone had a "black name."

"Personal health information - historically protected by the Hippocratic

Oath - has suddenly become the property of private corporations who may sell it to the highest bidder or accidentally misuse it to discriminate against the ill," Libert said. "As health information seeking has moved online, the privacy of a doctor's office has been traded in for the silent intrusion of behavioral tracking."

Online privacy has for some time been a concern. Studies conducted by Annenberg dating back to 1999 indicate wariness among Americans about how their personal information may be used. And slightly more than one in every three Americans even knows that private third-parties can track their visits to health-related websites.

Libert points out that the Federal Health Insurance Portability and Accountability Act (HIPPA) is not meant to police business practices by third party commercial entities or data brokers. The field of regulation is widely nonexistent in the U.S., meaning that individuals looking up health information online are left exposed and vulnerable.

According to Libert, "Proving privacy harms is always a difficult task. However, this study demonstrates that data on online health information seeking is being collected by entities not subject to regulation oversight. This information can be inadvertently misused, sold, or even stolen. Clearly there is a need for discussion with respect to legislation, policies, and oversight to address health privacy in the age of the internet".

Provided by University of Pennsylvania