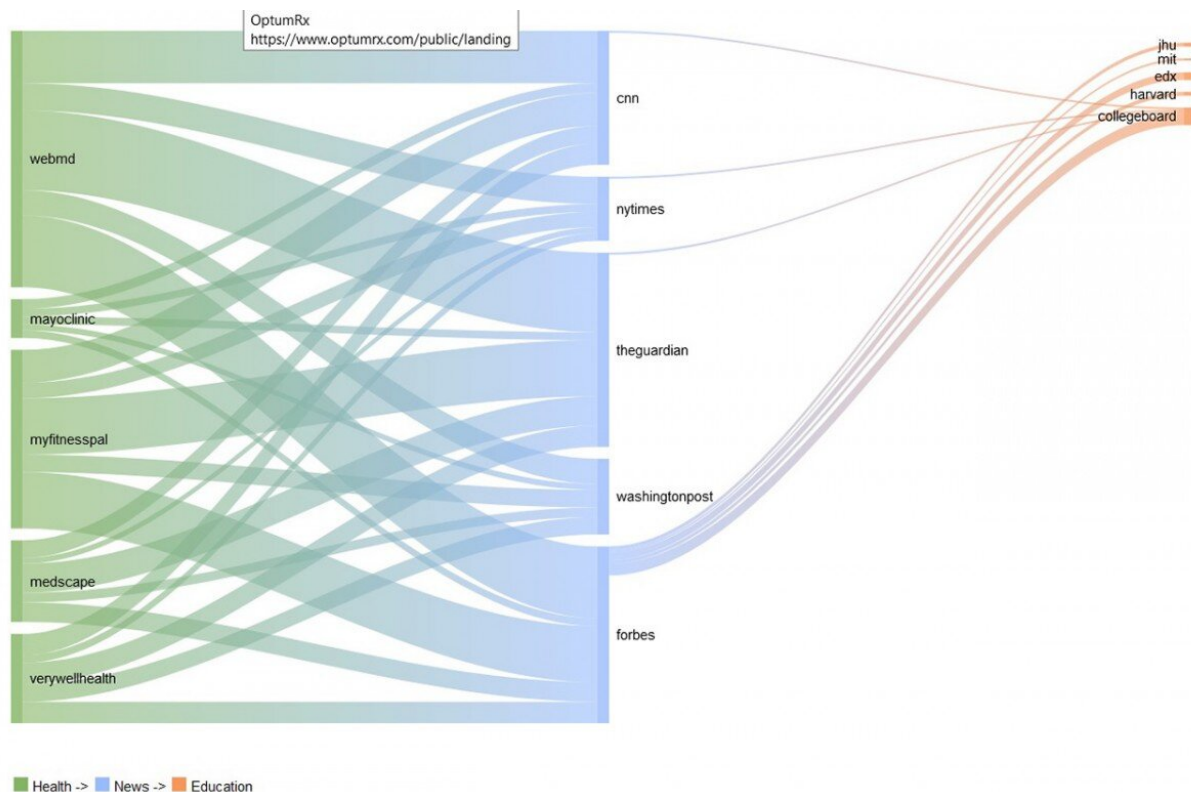


Study: Online trackers follow health site visitors

June 25 2020, by Melanie Lefkowitz



A visualization of one of the researchers' experiments, browsing from health to news to educational contexts, which shows particularly dense connections of user IDs between health care and news websites. Credit: Cornell University

Internet trackers are more likely to follow people who visit popular health sites, such as WebMD.com and mayoclinic.org, to other types of

sites, a Cornell Tech study has found—suggesting that advertisers might be more likely to target people based on sensitive health information than previously understood.

The study examined how the order in which users visit 15 major health, education and [news sites](#) affects the way third-party trackers follow them around the internet. Although the health sites may have fewer trackers than other types of sites, the researchers found, those trackers are more persistent in following page visitors.

"The health care [context](#) is really appealing to advertisers, since it's such sensitive data that allows advertisers to know a lot about you, or even manipulate you to click on an ad that relates to your health problem," said Ido Sivan-Sevilla, a postdoctoral fellow at Cornell Tech's Digital Life Initiative and first author of "Unaccounted Privacy Violation: A Comparative Analysis of Persistent Identification of Users Across Social Contexts."

The paper was co-authored by Helen Nissenbaum, Cornell Tech professor of information science and director of the Digital Life Initiative, and Cornell Tech master's students Wenyi Chu and Xiaoyu Liang. It will be presented July 21 at the Federal Trade Commission's PrivacyCon 2020.

"We wanted to look beyond this one-way mirror of our websites and see what's actually happening among the different social contexts of web browsing—to what extent trackers persistently identify users across different social contexts," Sivan-Sevilla said. "Instead of studying web tracking in bulk, across thousands of websites, we aimed to learn how advertisers take advantage of the fact that the web is comprised of should-be distinct social contexts such as health care, education and news."

Third-party web trackers are entities that collect browsing information about visitors to websites. They are embedded in nearly every [website](#), allowing content publishers to offload website functions such as advertising to other parties. For example, whenever people visit the *New York Times* website—including in the study—dozens of third-party trackers may be collecting data about which articles they read. People's search habits contain valuable information for advertising, site analytics or other uses, which third-party trackers may then use themselves or sell to other companies.

In the study, the researchers sought to empirically investigate whether social contexts—the types of websites people are visiting—matter for trackers. They based their research questions on Nissenbaum's theory of privacy as contextual integrity, which she developed and described in her 2010 book, "Privacy in Context: Technology, Policy and the Integrity of Social Life" (Stanford University Press). According to the theory, privacy demands appropriate flows of information—for example, the information that flows between friends is subject to different rules and norms from the information that flows between an employee and a supervisor.

In the context of this study, tracking people from a health site to a news site is a violation of privacy according to the theory of contextual integrity, Sivan-Sevilla said. "We expect our information from the health care context to be used for health advice, rather than for commercial purposes by advertisers in other websites."

Third-party trackers commonly remember visitors based on unique user identifiers stored via cookies, small pieces of information placed in our internet browsers. The researchers conducted six experiments representing all possible browsing sequences between health, education and news contexts. For each experiment, the researchers determined which user identifiers from the first context were persistently used by

trackers in the following two contexts.

Researchers found that users are followed among all three types of [social contexts](#), between every pair of websites they studied. They also found that health care websites are most likely to link users' identifiers to other types of websites.

Previous studies had found fewer trackers on [health care](#) sites, suggesting these sites were less risky for users' privacy. Looking at the tracking alongside other contexts revealed new patterns, the researchers found.

It's important to examine what third-party trackers are doing, Sivan-Sevilla said, because they're unregulated and little-understood despite the vast volume of [information](#) they collect, use and sell.

"The purpose of our research group is to start building a contextual understanding of tracking practices, adding a distinctive perspective to existing studies," Sivan-Sevilla said. "We want to shed more light on this complex ecosystem of web tracking, hopefully hold the industry more accountable and show regular people what's actually happening here."

More information: Unaccounted Privacy Violation: A Comparative Analysis of Persistent Identification of Users Across Social Contexts. drive.google.com/file/d/196ONO...6bk3bJ_RhC9SAk/view

Provided by Cornell University

Citation: Study: Online trackers follow health site visitors (2020, June 25) retrieved 10 April 2024 from <https://techxplore.com/news/2020-06-online-trackers-health-site-visitors.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.