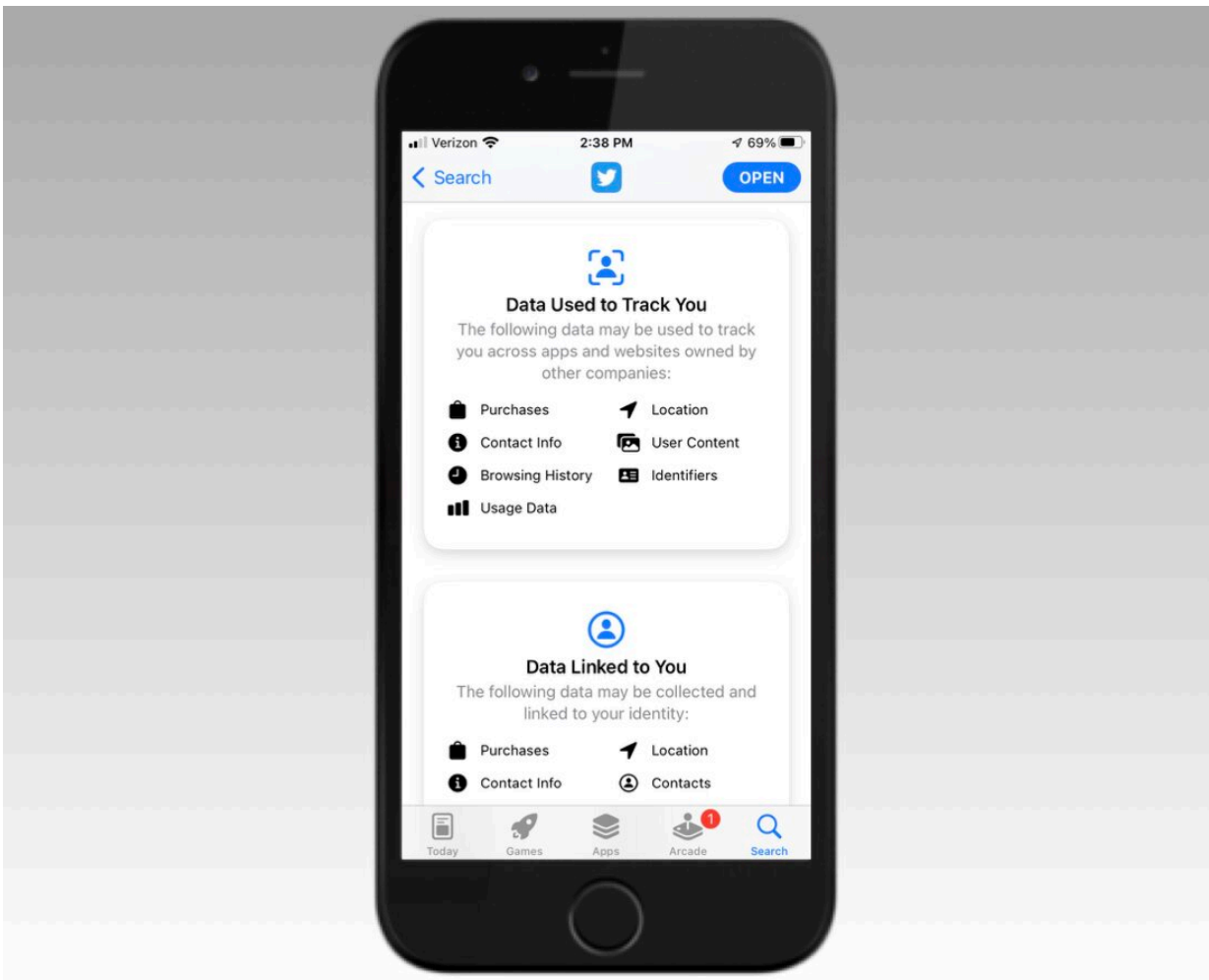# Researchers investigate Apple's privacy labels

May 3 2022, by Daniel Tkacik

A portion of the privacy label for Twitter in the Apple app store. Credit: CyLab

CyLab researchers have been studying privacy nutrition labels for over a decade, so when Apple introduced privacy labels in their app store a little over a year ago, the researchers were eager to investigate them.

"Naturally we had a lot of questions," says CyLab director Lorrie Cranor, a professor in the Institute for Software Research and Engineering and Public Policy, and the principal investigator of early privacy label research at Carnegie Mellon. "Are the labels conveying accurate information? How easy is it for app developers to make the labels?"

At the upcoming ACM *CHI Conference on Human Factors in Computing Systems*, CyLab researchers will be presenting new research aimed at those questions.

A close look at compliance and accuracy

In one paper, titled, "[Understanding iOS Privacy Nutrition Labels: An Exploratory Large-Scale Analysis of App Store Data](#)," CyLab researchers present comprehensive measurements of Apple privacy nutrition labels to learn about the rate of compliance in creating the labels and how accurate existing labels are.

"No one has conducted a large-scale analysis of Apple privacy labels like this before," says Yucheng Li, a student in Heinz College's Master of Information Systems Management program and the lead author on the study.

The researchers crawled the Apple's US app store every week from April to November 2021, capturing information on privacy labels and metadata for over 1.4 million total apps. Apps that were originally published before December 8, 2020 (just under 1.2 million apps) were not required to create privacy labels unless they release app updates but

could voluntarily do so, and apps developed after that date (roughly 275 thousand apps) were forced to do so.

Compliance to create the labels has been so-so.

"Over half of all apps in the app store still do not have a privacy label," says Li. "Although the overall compliance rate seems to be steadily increasing, the speed of compliance on older apps is on a downward trend. We speculate that if old apps don't have a privacy label now, they probably won't create one in the future."

The researchers also found that app updates seem to be an important driver of privacy label creation, as 64 percent of the apps released version updates at the same time as they published their privacy label. Lastly, they found that out of the apps that created a label, 43 percent have made at least one update, but under six percent made an update to the label. This means the current privacy label may not reflect the most up to date information.

Challenges faced by developers

A second paper being presented at the conference, titled, "[Understanding Challenges for Developers to Create Accurate Privacy Nutrition Labels](#)"—which earned an Honorable Mention recognition from the conference organizers—identified some key challenges for developers to create privacy labels, which may explain some of the findings of Li's paper.

"If the labels are not accurate, they will probably do more harm than good to users," says Tianshi Li, a Ph.D. student in the Human-Computer Interaction Institute and lead author of the study. "We currently have little understanding of developers' ability to create privacy nutrition labels."

The researchers observed 12 iOS developers create privacy labels using a replica of Apple's developer tool, and then followed up with an interview, asking them about their apps' data practices and the developers' understanding about the terms in their privacy labels.

"Developers generally felt positive about privacy nutrition labels," says Li. "But despite those positive reactions, errors and misunderstandings were still prevalent."

Nine out of the 12 developers made errors in the label that weren't corrected before being prompted by the interviewer. Among eight apps that already had a privacy label, six were re-created inconsistently.

Other challenges developers faced included limitations of Apple's documentation about creating privacy labels—some developers reported being confused by jargon used in the documentation while others complained the documentation used vague, ambiguous terms or provided ineffective examples—and information overload.

"Developers complained that they had to read a lot of text," says Li. "This information overload could have further implications than just the time spent on the task of creating labels."

One developer, for example, admitted that they might not always update their privacy labels in a timely manner because "upgrading it, or at least reviewing it on every update would be tiresome."

The researchers argue that in order for privacy labels to achieve more widespread adoption, there needs to be better support for developers, including better design and evaluation of developers' tools to create the labels to serve a wider range of developers who span a spectrum of skills.

Faculty and students who authored these studies are part of the CyLab Usable Privacy and Security (CUPS) Lab, the Computer Human Interaction: Mobility Privacy Security (CHIMPS) Lab, and the Systems, Networking, and Energy Efficiency (Synergy) Lab at Carnegie Mellon University.

Provided by Carnegie Mellon University