Not all privacy apps are created equal

April 3 2020, by Adam Conner-Simons

Credit: CC0 Public Domain

New privacy laws like Europe's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) have spawned a new industry of companies and platforms advertising that they can anonymize your data and be compliant with the law.

But MIT researcher Aloni Cohen says that he has his doubts about these claims, and his team's latest work shows that there's reason to be skeptical.

Specifically, a new journal article from Cohen and professor Kobbi Nissim argues that an anonymity technique called k-anonymity—which is used by many companies that make such claims—does not prevent a user from being singled out and de-anonymized by looking at the platform's wider data. The researchers study a new type of attack they call "predicate singling out," modeled after a type of GDPR privacy violation called singling out.

"I think it's reasonable to say that many of the claims made by these 'anonymity-as-a-service' companies are suspect," says Cohen, whose article with Nissim was published online today in *PNAS*. "This paper is one step in testing that and showing the holes in their approach."

The team made the case that companies using k-anonymity to anonymize data might instead employ differential privacy, a newer technique that involves precisely controlled randomization to mask the presence or absence of any particular individual in a dataset. The researchers show that differential privacy prevents predicate singling out attacks.

Differential privacy is seeing growing adoption in settings where more traditional approaches to anonymization are deemed inadequate. The US Census Bureau is using differential privacy to provide confidentiality for the 2020 census. The adoption of GDPR also spurred Facebook to use differential privacy to aid social scientists studying disinformation online.

"While we show differential privacy prevents predicate singling out attacks, it's not necessarily full-fledged anonymization under the law," says Cohen. "On the other hand this work shows that, as a general rule, you should be skeptical of any company that tells you that their use of k-anonymity gives you "GDPR compliance.'"

The paper also represents an intriguing new example of how math and

computer code can be used to quantifiably determine whether companies are actually following the law.

"We feel that proving that something is PSO-secure is not just a mathematical concept, but one that can be used to support a legal conclusion, and that should actually have legal consequences," says Cohen.

More information: Aloni Cohen et al. Towards formalizing the GDPR's notion of singling out, *Proceedings of the National Academy of Sciences* (2020). DOI: 10.1073/pnas.1914598117

Provided by Massachusetts Institute of Technology