

You Can be Anonymised But You Can't Hide

Article By:

Cameron Abbott

If you think there is safety in numbers when it comes to the privacy of your personal information, think again. A recent study in [Nature Communications](#) found that, given a large enough dataset, anonymised personal information is only an algorithm away from being re-identified.

Anonymised data refers to data that has been stripped of any identifiable information, such as a name or email address. Under many privacy laws, anonymising data allows organisations and public bodies to use and share information without infringing an individual's privacy, or having to obtain necessary authorisations or consents to do so.

But what happens when that anonymised data is combined with other data sets?

Researchers behind the Nature Communications study found that using only 15 demographic attributes can re-identify 99.98% of Americans in any incomplete dataset. While fascinating for data analysts, individuals may be alarmed to hear that their anonymised data can be re-identified so easily and potentially then accessed or disclosed by others in a way they have not envisaged.

Re-identification techniques were recently used by the [New York Times](#). In March this year, they pulled together various public data sources, including an anonymised dataset from the Internal Revenue Service, in order to reveal a decade's worth of Donald Trump's negatively adjusted income tax returns. His tax returns had been the subject of great public speculation.

What does this mean for business? Depending on the circumstances, it could mean that simply removing personal information such as names and email addresses is not enough to anonymise data and may be in breach of many privacy laws.

To address these risks, companies like Google, Uber and Apple use "differential privacy" techniques, which adds "noise" to datasets so that individuals cannot be re-identified, while still allowing access to the information outcomes they need.

It is a surprise for many businesses using data anonymisation as a quick and cost effective way to de-personalise data that more may be needed to protect individuals' personal information.

If you would like to know more about other similar studies, check out our previous blog post ['The Co-Existence of Open Data and Privacy in a Digital World'](#).

National Law Review, Volume IX, Number 219

Source URL: <https://natlawreview.com/article/you-can-be-anonymised-you-can-t-hide>