case study

# How privacy tech innovation from Google is helping OpenMined unlock the potential of their data, safely and securely

Learn why the open-source community turned to Google's privacy-protecting technology including Differential Privacy and Federated Learning.

## Executive Summary Module

➜ OpenMined is helping bring data to the masses—safely and securely. The network of coders uses Google's privacy-protecting technology to keep data private.

Andrew Trask started OpenMined with a simple goal in mind: he wanted to help make the world's data more available to researchers and developers while protecting the integrity of the data sets and the privacy of people whose data is in them. He imagines a scenario where data scientists and researchers could work with all the world's cancer data at the click of a button, while every individual's right to their personal privacy remains protected.

In this world, data wouldn't be siloed, and innovation would be easier. "We could dramatically increase the potential of research efforts by expanding the availability of valuable data sets, leading to incredible medical, scientific, and social progress," Trask, a PhD candidate at the University of Oxford, says.

Andrew Trask

OpenMined

**9,500**

Members in the OpenMined coding community

**35**

Core coding teams working with privacy-protecting technology

For OpenMined's network—a group of thousands of volunteer coders—applying privacy protecting AI such as Google's Differential Privacy and Federated Learning is central and core to their mission of unlocking the potential of data for global researchers. They lean on the world's best privacy technologies to make data sets usable, helping researchers gain valuable insights from data while keeping the original privacy promises of each data set. "Instead of an internet that builds answers on public information, we can build one for answering questions with private information that remains protected," Trask says.

At OpenMined, 35 core code teams frequently turn to Google's privacy-protecting technologies to create a more private and secure environment. "Privacy and cryptography are about trust," Trask says. "It's not just trust in the algorithm. It's also trust in the implementation. And I don't think there's a better place you can go to for robust implementation of Differential Privacy than Google. It's been through multiple security audits, and it's really, really robust. In terms of people who know how to deploy DP, there's no one who's done more good than Google." And because Google's implementation of Differential Privacy uses an open source repository, it's widely available for researchers and developers.

When the COVID-19 virus started sweeping the globe, OpenMined's members used Google's open-source Differential Privacy application to protect data that was being collected by public health survey apps while still making use of it. "You want to make it possible for health authorities to know the high-level trend of what's happening in their jurisdiction without them necessarily being able to know anything about any specific person," Trask, who also works with Google's DeepMind team, explains.

One way Differential Privacy accomplishes this goal is to add noise to individual data points before they're sent to the essential server for aggregation. For example, researchers could use Differential Privacy to subtract or add 10 to a person's age, obscuring the actual value to maintain privacy while the overall data set remains statistically viable. "If you average over enough of these, you still get the same means and standard deviations, which are what's useful for statistics," Trask says. Public health authorities can stay informed without compromising the privacy of individuals.

OpenMined's teams are also working with Federated Learning, a technology developed and open sourced by Google. They recently shipped the first web and mobile Federated Learning framework, which will increase use cases for the algorithm. "Now you can build websites, mobile apps, and Javascript apps that train data using Federated Learning," Trask says. "Google invented something and now it's going to take wings and run on its own. It shows their leadership in the privacy space."

But OpenMined's work isn't a one-way relationship with Google. There's a back and forth, building upon each group's privacy-protecting efforts. "When we hop on a call with Google's anonymisation team and we're working back and forth and giving demos to each other, it's just so motivating," Trask says. "The ability to interact with role models at a company like Google is such a compelling and interactive experience for our volunteers. And I've never met anyone at Google who didn't care about privacy."

> "Instead of an internet that builds answers on public information, we can build one for answering questions with private information that remains protected."