

Genetic database that identified Golden State Killer acquired by crime scene DNA company

The crime scene DNA sequencing company Verogen announced yesterday that they've acquired the genomics database and website GEDmatch. The acquisition makes the relationship between the company and law enforcement explicit, but raises uncomfortable questions for users and experts about data privacy and the future direction of the platform.

GEDmatch was primarily used by genealogists until 2018, when police, the FBI, and a forensic genealogist identified the suspected Golden State Killer by tying crime scene DNA to relatives who had uploaded their genetic information to the site. Since then, the platform has helped identify around 70 people accused of violent crimes.

In response to privacy concerns, the company changed its terms and conditions last spring to only allow law enforcement access to data if users actively opted in. But until now, interaction with law enforcement was still a secondary function to the platform.

"Beforehand, it was somewhat of an ad hoc system, and whoever wanted to be there was going to come in and use it," says Brad Malin, co-director of the Center for Genetic Privacy and Identity in Community Settings at Vanderbilt University. The acquisition signals a shift in purpose. "Now, it will be used for law enforcement in a much more systematic manner than was the case."

The announcement took many in the genetics and genealogy community by surprise, and many genealogists are leaving the platform. “There have simply been too many changes, all of them in the direction of making their data the product rather than the website a service,” said lawyer and genealogist Judy Russell in an email to *The Verge*.

GEDmatch users were prompted to accept new terms and conditions indicating the platform’s new ownership, and could either agree and enter the site, or remove their data from the platform. Verogen will still allow users to keep their data from any use by law enforcement, CEO Brett Williams told *BuzzFeed News*, maintaining the opt-in approach. “It will be interesting to see in the future if the new owners will implement policy changes that will increase the number of individuals available for law enforcement searching,” says James Hazel, postdoctoral fellow at the Center for Genetic Privacy and Identity in Community Settings at Vanderbilt University.

However, opt-in is not a foolproof system for data protection: last month, a Florida detective received a warrant to search the entire GEDmatch database, regardless of whether users agreed.

The details of the process of users entering into the new agreement seem to be insufficient, Malin says. “Normally, with informed consent, you try to indicate what all of the risks are going to be and present the benefits. This is not so much consent as a contractual agreement,” Malin says. “With that, it’s buyer beware.” Even if people agree to the new terms, they might not fully understand everything they’re being asked. In addition, users were not notified of the acquisition, and people who aren’t using the platform regularly – and didn’t see the login prompt – wouldn’t have the opportunity to make a decision on whether to remove their data from the database.

The amount of data available to law enforcement on GEDmatch shrunk dramatically once it instituted an opt-in system, but the full database includes around 1.3 million user profiles and grows every day. Those profiles can be used to identify many times that number of people: a study published in the journal *Science* showed that researchers could identify around 60 percent of people with European ancestry in the United States using a data set around that size.

The full database, then, is a powerful tool. However, data collected by a private, commercial company within the disparate system of genetic databases in the US has inherent biases and weaknesses. People with European and Caucasian ancestry are disproportionately represented in most genetic databases, for example, and many people in the US are adopted. "Just trying to characterize people by biological relationships may work in some situations, but not in others," Malin says.

Systems like GEDmatch can also be manipulated, and it's possible for people to create false profiles and fake familial relationships within the database, researchers have found.

Those types of gaps in currently existing databases are part of why Malin and his colleagues, including Hazel, outlined the benefits of a public, universal forensic genetic database in a 2018 paper – a huge amount of genetic information is already contained in databases in the United States, but it's fragmented and inconsistently managed. A universal database would be a more effective and less discriminatory tool, they argued.

"We were calling for an environment with more control," he says. "Now [GEDmatch and similar companies] take data and put it in a private, commercial domain where there is no clear oversight associated. That's a scary prospect."

<https://www.theverge.com/2019/12/10/21005443/golden-state-kill>

er-genetic-database-identity-company-acquisition-crime-scene-
dna-data