



The New York Times

https://nyti.ms/2FoCoh9

U.S.

# AP Explains: A Look at DNA-Sharing Services and Privacy

By THE ASSOCIATED PRESS APRIL 27, 2018, 8:31 P.M. E.D.T.

NEW YORK — The use of a genealogy website to track down a suspected California serial killer illustrates both the extraordinary power of DNA-sharing services and the broad privacy concerns that surround the fast-growing commercial market for genetic analysis.

TV commercials for companies such as 23andMe and Ancestry.com pitch their services as simple and fun ways of learning about family heritage and health. And while those companies on Friday sought to distance themselves from the free GEDmatch website used by police, the California case exposed broader questions about what happens after consumers mail their saliva away for DNA analysis and upload the results to the internet.

"For those of us who were skeptical about turning over our genetic information to corporations, this case proved all of those fears true," said Daniel De Simone, a New Jersey researcher whose relatives have used DNA services.

The co-founder of GEDmatch said Friday that he is concerned about privacy after learning that law enforcement used the site and insists that his company does

4

ARTICLES REMAINING

SEE MY OPTIONS

Subscriber login

"This was done without our knowledge, and it's been overwhelming," Curtis Rogers told The Associated Press.

Authorities have not publicly described details of the methods that led to the arrest of 72-year-old Joseph James DeAngelo, a former police officer. But some other researchers who use GEDmatch for a similar purpose told The Associated Press how it might have been done.

Colleen Fitzpatrick and Margaret Press run the DNA Doe Project, a California-based nonprofit that uses DNA from unidentified bodies to look for relatives and learn the names of the dead. They suspect the authorities used a method similar to their own.

Investigators probably started with the complete DNA code from the killer. Then, after putting the data into a format that GEDmatch can read, they plugged it into the Florida-based site and asked it to look for matches, they said.

The site compares particular segments of genetic material, looking for similarities to other samples in its database. The degree of similarity can indicate how related two people are, finding ties as distant as fifth cousins, Fitzpatrick and Press said.

Once the site has returned a list of matches and degree of similarity, more sleuthing begins.

People in the database may have listed their names or just their emails, which in turn might identify them if they've used the same one on other sites. More information can come from searches of public records, Facebook and especially obituaries, which list parents and other relatives.

Then a researcher can turn to online collections of family trees, like those on Ancestry.com. There, one might uncover many trees that include the apparent relatives found on GEDmatch. That allows the construction of speculative trees that include the mystery person, plus those apparent relatives, in an effort to find overlaps that indicate common ancestors.

If common ancestors appear, lineage can be worked out in detail, up to the present day. Researchers look for a spot that contains someone of the right geographic location and age to be the person under investigation.

"It's part art and part science," Press said.

Court records obtained by The Associated Press on Friday showed that investigators had used information from genetic websites a year ago and misidentified an elderly Oregon man as a possible suspect.

A judge signed an order to compel a DNA sample from the 73-year-old man after detectives used a genetic profile based off DNA from crime scenes linked to the serial killer and compared it to information from YSearch.org, a free service that's provided by FamilyTreeDNA.com.

Investigators cited a rare genetic marker, which the Oregon man shared with the killer, to get the judge to issue the order. They also created a family tree and used public records to identify the Oregon man.

A spokeswoman for FamilyTreeDNA.com did not immediately respond to a request for comment.

To some, scouring this publicly shared data to track down the so-called Golden State Killer seems like a worthwhile cause. But for others it raises alarms.

De Simone said he has never used a DNA ancestry service.

"What's especially troubling to me is that neither had DeAngelo," he said. He compared the situation to Facebook's data-protection scandal involving Cambridge Analytica because "it's not only users that are caught up in this net, it's also those with relationships to users. In this case, though, it's not just networked relationships, it's actual genetic relationships."

The big commercial databases insist they have much stricter customer privacy practices than websites such as GEDmatch and don't hand over data without a court order.

"As a private platform, we do not allow the comparison of genetic data processed by any third party to genetic profiles within our database. Further, we do not share customer data with any public databases or with entities that may increase the risk of law enforcement access," 23andMe spokesman Andy Kill said in an email.

It's unclear whether the California case will affect customers' trust in DNA services overall.

"These companies are saying that they're different," said Tiffany Li, a technology attorney and Yale Law School fellow. "I think what's key is this open-source database is made up of data profiles that people mostly got from those private companies."

Li said the demand for personal genetic information that helps uncover long-lost relatives and family backgrounds is high enough that this privacy "dust-up" will likely blow over. But, she said, it should serve as a warning for stewards of DNA databases to be more careful and more transparent about how data is used.

"They should at least try to do more to make people aware," Li said. "The terms could be clearer. The companies could also decide to self-regulate before Congress gets to them and create data standards about the DNA they store."

—

O'Brien reported from Providence, Rhode Island. AP Technology Writer Mae Anderson in San Francisco contributed to this report.