The Washington Post
Democracy Dies in Darkness

Opinion: After a rape survivor's arrest, it's time to rethink genetic databases

By Jennifer King

February 17, 2022 at 1:59 p.m. EST

Jennifer King is a privacy and data policy fellow at the Stanford Institute for Human-Centered Artificial Intelligence.

San Francisco District Attorney Chesa Boudin made a disturbing announcement on Monday: Rape survivors' DNA profiles had been routinely submitted to the city's crime lab for inclusion in databases used to identify criminal suspects. The practice came to light, he said, after a woman's DNA from a rape kit collected several years ago led to her recent arrest for a property crime.

Boudin's office said Tuesday that he had dismissed the case against the woman, and the San Francisco Police Department has said it is investigating the crime lab's practices. But it is not too early to say: Including the DNA of rape survivors in a database used for criminal identification is a violation of privacy, full stop. It doesn't matter whether the rape survivor herself later commits a crime.

The majority of sexual assaults go unreported, often because survivors want to avoid the additional trauma of a physical exam to collect evidence. Boudin was certainly right to raise an alarm over how the potential misuse of their DNA could further discourage survivors from contacting police.

This episode offers a glimpse of the concerns over privacy — as well as matters such as consent to data collection — that will arise as genetic information is stored in ever-greater amounts, and as governments take an ever-greater interest in exploiting it.

According to Boudin, it appears that people reporting sexual assaults were not informed that their own DNA, included in the rape kit, would be used for other purposes beyond the investigation of their assault. Further, he suggested this might be a problem across the state, although several other California departments have denied using DNA from rape kits in unrelated investigations.

But it does *not* appear that these rape survivors were explicitly informed that the information could be used for criminal identification, nor were they given the opportunity to refuse permission for that use. In any case, there is a substantial body of research, including my own, showing how difficult it is for people to understand terms and conditions in the best of circumstances, much less in the aftermath of a sexual attack.

Let's give the San Francisco Police Department some benefit of the doubt — maybe it will turn out that it uses a single genetic database that doesn't allow for differentiating between survivors and perpetrators. Separate databases should be obligatory, because the potential problems are so obvious.

The problem is that law enforcement's appetite for genetic material is insatiable. In California, that hunger was fed by voters, who approved Proposition 69 in 2004, allowing police to collect DNA samples from anyone they arrest, even if that person is ultimately not charged with a crime. The program is decentralized (local jurisdictions collect the bulk of this data, which isn't necessarily shared at the statewide level), and any oversight of these databases occurs at the county level.

Government's eagerness for genetic material has only increased since 2018, with the identification of the Golden State Killer, a serial rapist and murderer in the 1970s and 1980s whose case had gone cold until the Sacramento district attorney teamed up with forensic genealogists to submit the suspect's DNA to a consumer genomic database.

The use of DNA to identify long-sought criminal suspects to bring them to justice is perhaps an ideal example of this technology in action, but it also provides the justification for unchecked DNA collection. After all, the Golden State Killer was living anonymously among us for over three decades — what other serial killers might be hiding in our midst? If you have nothing to hide, the saying goes, then you have nothing to fear from having your DNA added to a database to identify criminal suspects.

As the Golden State Killer case showed, people themselves are doing much of the DNA database construction voluntarily — by using popular services such as 23andMe and Ancestry to trace their own genealogy through genetic testing. The phenomenon has exploded, outstripping regulations overseeing the industry.

The result is that while only a relatively small number of people in the United States have participated in these genetic-testing services, the familial networks they reveal are staggeringly extensive: Analysts predict that soon nearly anyone in the country with European ancestry might be identifiable through these relationships.

Given the rise of authoritarianism around the world, and the expanding power of artificial intelligence, it is not difficult to imagine how these mountains of genetic information could be abused — and already are, by the Chinese government in its oppression of the Uyghurs.

The circumstances under which we allow the government to collect DNA for identification purposes must be exceptional. Otherwise, the generalized collecting and collating and comparing of DNA by law enforcement will only proliferate as officials sift through familial networks past, present and future.

Nobody wants suspected rapists and murderers to evade justice. But widespread DNA collection without cause and without constraint is not the answer.