Request for Information (RFI) on Public and Private Sector Uses of Biometric Technologies: Responses

DISCLAIMER: Please note that the RFI public responses received and posted do not represent the views or opinions of the U.S. Government, the Office of Science and Technology Policy (OSTP), or any other Federal agencies or government entities. We bear no responsibility for the accuracy, legality, or content of these responses and the external links included in this document. Additionally, OSTP requested that submissions be limited to 10 pages or less. For submissions that exceeded that length, the posted responses include the components of the response that began before the 10-page limit.
Comments of the Electronic Frontier Foundation Regarding

Notice of Request for Information (RFI) on
Public and Private Sector Uses of Biometric Technologies

Office of Science and Technology Policy (OSTP)

Document No. 2021-21975

86 Fed. Reg. 56300

Submitted on January 15, 2022

The Electronic Frontier Foundation (EFF) submits the following comments in response to the Office of Science and Technology Policy (OSTP)’s Request for Information (RFI), published at Document Number 2021-21975. These comments will focus on current and proposed uses of DNA technology in immigration and law enforcement. They will respond primarily to Topic 4, “Exhibited and potential harms of a particular biometric technology,” although they will also touch on Topic 1, “Descriptions of Use;” Topic 3, “Security Considerations;” and Topic 6, “Governance Programs.”

EFF is a non-profit organization that has worked for 30 years to protect civil liberties, privacy, consumer interests, and innovation in new technologies. EFF actively encourages and challenges all branches of government to support privacy and safeguard individual rights as emerging technologies become more prevalent in society. With more than 30,000 contributing members, EFF is a leading voice in the global and national effort to ensure that fundamental liberties are respected in the digital environment.

I. Introduction

DNA contains our most private and personal information. Unlike other biometrics that can (or should) only be used for identification, DNA provides “a massive amount of unique, private information about a person that goes beyond identification of that person.”¹ Every tiny piece of skin or hair or saliva contains a person’s entire genetic code—information that has the capacity to reveal the individual’s race, biological sex, likely eye and hair color, ethnic background, familial relationships, behavioral characteristics, health status, genetic diseases, and predisposition to certain traits. Companies are even able to use genetic data to predict what an unknown person might look like today or in the past, which can then be used to generate a composite image of the person.² DNA data can also be combined with other data from public records and

social media to create a full picture of a person’s life. And unlike a social security or driver’s license number, DNA can never be changed. The depth and breadth of information contained in our DNA justifies strong restrictions on its collection and use.

II. DNA Collection and Use Has Expanded as DNA Technology Has Advanced

DNA has been used in criminal investigations for nearly 40 years, and DNA collection is now mandatory from those convicted of or arrested for most crimes. DNA profiles are stored in various local, state, and national DNA databases, including the FBI’s Combined DNA Index System (CODIS) database, which contains 14.8 million offender profiles and more than 4.5 million arrestee profiles.

DNA technology has advanced significantly in the last few decades. Where once, a useful forensic sample could only be obtained from blood, semen, or other bodily fluids, today, investigators can detect, collect, and analyze DNA from objects merely touched by a person. Because we cannot help but leave DNA behind on nearly everything we touch, this vastly expands investigators’ ability to collect DNA. It also allows investigators to collect DNA from and identify individuals without their knowledge—for example, by collecting DNA from a straw or cigarette they may have used and discarded. The heightened sensitivity in DNA collection technology has also resulted in the collection of trace amounts of DNA from objects, like a door knob or a knife, that may have been touched by more than one person. Probabilistic genotyping software claims to be able to analyze these DNA mixtures and identify unique individuals.

The time it takes to process and sequence DNA has decreased significantly as well. Where once it took labs a month or more to generate a DNA profile, federal and state agencies now have access to Rapid DNA analyzers—self-contained, automated, portable machines that allow non-scientists to process a DNA sample and generate a profile outside a lab in as little as 50 minutes. This allows DNA to be used much like other biometrics—to identify specific individuals on the spot.

Further, the costs of genetic sequencing have decreased so significantly that DNA

---


collection and analysis have expanded far beyond use for health, research, or criminal justice purposes. DNA profiles used for criminal justice purposes have typically contained 13-20 short-tandem repeat (STR) DNA markers, which are specifically chosen from non-coding (and thus less revealing) segments of DNA. But as costs related to DNA processing have decreased, police are beginning to access more information from DNA samples. It is now possible to create a genetic profile made up of more than half a million single nucleotide polymorphisms (SNPs) that span the entirety of the human genome for just $99 or less. This has led to the rise of consumer genetic testing, the portability of genetic data, and the growth of consumer genetic databases that offer to connect people to long lost relatives and provide insights into health, frequently for no cost at all. These sites and databases have already been used by the police in hundreds of investigations to identify people who are genetically related to unknown DNA samples and to identify a sample’s donor, even when that donor has not entered their own genetic data into the consumer database.

As the costs associated with DNA collection and processing have decreased, federal and state agencies have increased their reliance on DNA. For example, DNA use for immigration-related purposes has expanded, despite a lack of supporting statutory or regulatory authority. In 2019, the U.S. Department of Homeland Security (DHS) began a program to use Rapid DNA on migrant families at the U.S.-Mexico border to identify and prosecute individuals who are not related through a biological parent-child relationship. The program began as a three-day pilot program at two locations and quickly expanded to a 10-month program at seven locations. DHS later extended this contract for up to five years. Although DHS claimed that the testing was voluntary, refusal to submit to testing could factor into a decision of whether to separate parent from

child in immigration detention.\textsuperscript{13} And in a 2020 Notice of Proposed Rulemaking (NPRM), DHS proposed extending its regulatory authority to mandate DNA collection “for any benefit request where [a genetic] relationship must be established.”\textsuperscript{14} DHS would have applied this rule to prospective immigrants and U.S. persons alike, potentially increasing the number of people subject to mandatory DNA collection from zero to an estimated 805,493 each year.\textsuperscript{15} Of those, approximately 336,650 would have been U.S. citizens.\textsuperscript{16} DHS proposed retaining “partial DNA profile” data that would have genetically linked family members in the database.\textsuperscript{17} In the NPRM, DHS left open the possibility that it could, at its own discretion, share DNA test results and DNA profiles with other agencies, including law enforcement agencies.\textsuperscript{18}

III. DNA Collection Presents Unique Threats to Privacy

In the past decade, law enforcement and immigration authorities have been working to normalize biometrics collection and expand the collection of DNA from more and more individuals caught up in the criminal justice and immigration systems with little evidence that doing so solves or prevents crimes or immigration violations.\textsuperscript{19} By vastly expanding the amount of DNA collected and added to national, state, and local DNA databases, however, these efforts are bringing us closer to a regime of DNA collection from the entire population without any public or legislative debate on the serious attendant threats to privacy and civil liberties.

A. DNA Can Reveal More Sensitive and Private Information Than Other Biometrics

DNA collection threatens privacy because each DNA sample contains a person’s


\textsuperscript{15} Id. at 56380. DHS noted that it “currently accepts DNA test results for 11,383 beneficiaries” each year. However, none of these submissions are mandatory.

\textsuperscript{16} Id. at 56380.

\textsuperscript{17} Lynch, \textit{supra} n. 10, at 18–19.

\textsuperscript{18} 85 Fed. Reg. at 56353; \textit{see also} Proposed 8 C.F.R. pt. 103.16(e).

entire genetic code. This not only can reveal a person’s propensity for various diseases like breast cancer or Alzheimer’s and can predict traits like addiction and drug response, but can also identify family members and ancestors, predict a person’s appearance, and may reveal much more information in the future as scientific knowledge advances.

DNA profiles, which contain less data than a full DNA sample, still present privacy threats. One study—conducted when the FBI’s CODIS database relied on just 13 loci—found that the STR profiles in CODIS can identify information about individuals’ ancestry, which may, in turn, be used to reveal information about their phenotypic traits (i.e., physical appearance) based on assumptions about race and ethnicity.\(^{20}\) Another study suggested that the profiles maintained in CODIS can now be matched to SNP profiles in other publicly accessible databases, suggesting that DNA profiles stored in government databases could be used to identify anonymized genomes from health-research databases or other sources.\(^{21}\)

Data aggregation—combining genetic profiles with other government-maintained or publicly available data—increases these privacy risks. In 2012, researchers used genetic genealogy databases and publicly-available information to identify nearly 50 people from just three original anonymized samples.\(^{22}\) More recent research shows that 60% of white Americans can already be identified from a genetic genealogy database representing just 0.5% of the U.S. population.\(^{23}\) While the FBI’s CODIS database does not store any names or personal identifiers with DNA profiles, and the FBI keeps DNA separate from other biometric data,\(^{24}\) DHS proposed storing DNA information in an immigrant’s “A-file,” along with all other biometric and biographic information.\(^{25}\) This


\(^{23}\) Jocelyn Kaiser, *We Will Find You: DNA Search Used to Nab Golden State Killer Can Home in on About 60% of White Americans*, SCIENCE (Oct. 11, 2018, 2:00 PM), https://www.sciencemag.org/news/2018/10/we-will-find-you-dna-search-used-nab-golden-state-killer-can-home-about-60-white (This same research shows that once just 2% of the U.S. population has uploaded DNA, 90% of white Americans would be identifiable.).


\(^{25}\) 85 Fed. Reg. at 56347.
would make DNA data and relationship information easily accessible to other users of those databases.

B. Collecting Data on Genetic Relationships Threatens the Privacy Interests of Whole Communities

Consumer genetic databases now make it possible to map generations of family members using the genetic information of a relatively small number of individuals who opt in to sharing their own data. As noted above, police are increasingly accessing this information in criminal investigations. Familial searches through non-criminal data have been hotly debated within the American public, with many people choosing to opt-out of these searches where they can.\(^{26}\)

In 2020, DHS proposed collecting genetic data that would have allowed it to similarly map family relationships.\(^{27}\) This would have made familial DNA searches accessible to any agency or user who had access to DHS’s database and a law enforcement- or immigration-related need. DHS proposed to do this with no public debate and no congressional oversight. But collecting genetic relationship data from immigrants and U.S. persons would have allowed the federal government, in the near future, to map whole generations of family members, and by extension, whole immigrant communities. Agency action like this further compounds threats to individual privacy and autonomy and violates societal norms.

In 2008, the United Nations High Commissioner for Refugees (UNHCR) recognized that DNA testing “can have serious implications for the right to privacy and family unity,” and should be used only as a “last resort.”\(^{28}\) UNHCR noted that, if DNA is collected, it “should not be used for any other purpose than the verification of family relationships” and that DNA associated with the test “should normally be destroyed once a decision has been made.” Proposals such as DHS’s have failed to meet even this bare requirement.

IV. Government-Mandated DNA Collection Exacerbates Racial Disparities and Harms Vulnerable Populations

Expanding the collection of DNA will exacerbate racial disparities that are already present in existing DNA databases and harm vulnerable populations. In 2011, it was estimated that Black individuals made up 40 percent of profiles in CODIS (despite representing only 13.4% of the U.S. population), and that it was possible, even with


\(^{27}\) *See, e.g.*, 85 Fed. Reg. at 56341.

limited CODIS profiles, “to use the database to identify up to 17 percent of the country’s entire African-American population.”\(^{29}\)

Several recent proposals to expand DNA collection have been aimed squarely at immigrants and refugees. In March 2020, the U.S. Department of Justice finalized a regulation that allowed for DNA collected from immigrant detainees to be entered into CODIS—leading to the addition of up to 750,000 new DNA profiles a year.\(^{30}\) This will undoubtedly further skew the racial disparities in CODIS, given that in 2018, 43 percent of immigrant detainees were from Mexico and an additional 46 percent were from Guatemala, El Salvador, or Honduras.\(^{31}\) DHS’s 2020 proposal similarly threatened to disproportionately impact communities of color, both among U.S. persons and immigrants.

V. Government DNA Collection Puts Innocent People at Risk of Being Accused of Crimes They Didn’t Commit

The overcollection of DNA puts individuals at risk of being identified for a crime they did not commit, merely because their DNA already exists in a government database. This is so because we shed DNA constantly and because forensic tools are so sensitive that they can detect DNA on almost any surface, even if it is only a trace amount. DNA may be found not only on items that a person has touched, but also on other items with which the person never came into contact—a phenomenon known as “secondary transfer.”\(^{32}\) This means that crime scene samples can contain DNA from someone who was never at the scene or who was there or touched something transported to the crime scene long before the crime was ever committed. In California, a man spent five months in jail after a database search linked his DNA to DNA found on the fingernails of a murder victim—although the man was in the hospital when the murder occurred.\(^{33}\) Prosecutors believe paramedics may have transferred his DNA to the murder victim when they responded to the crime scene hours after dropping him off at the hospital. He never

---


would have been linked to the crime if his DNA had not already existed in a government database.\textsuperscript{34} Given this, researchers have recognized that “[a] DNA hit does not show that the subject is the offender and there are many reasons why the DNA of an individual may be found at a crime scene.”\textsuperscript{35} Nevertheless, this has not stopped prosecutors from arresting someone solely based on a DNA hit.

VI. Government-Mandated DNA Collection Must Take Account of Reliability, Accuracy, and Security Issues

DNA analysis is far from infallible, and faulty DNA processing has threatened people’s civil liberties. In 2015, for example, a San Francisco crime lab analyst repeatedly made assumptions about poor-quality, incomplete genetic evidence, falsely linking a DNA profile to a defendant and potentially causing errors in as many as 1,400 other cases.\textsuperscript{36} Similarly, the Washington D.C. Crime Lab lost its accreditation several times for its error-prone DNA analyses.\textsuperscript{37} And just this month, the Virginia Attorney General announced the Virginia Beach Police Department used forged DNA reports to get confessions.\textsuperscript{38}

Accuracy and reliability challenges have continued, even as technology has advanced. In 2017, the Swedish National Forensic Centre published a report detailing serious problems with certain Rapid DNA analyzers, finding that “36% of the runs had

\textsuperscript{34} This is not an isolated occurrence. In another case, the main contributor of DNA found on the murder victim’s underwear had been dead for two years before the murder was committed. Erin E. Murphy, \textit{How DNA Evidence Incriminated an Impossible Suspect}, The New Republic (Oct. 26, 2015) https://newrepublic.com/article/123177/how-dna-evidence-incriminated-impossible-suspect.


problems or errors effecting two or more samples.” 39 This resulted in a 23% failure rate, even with higher quantity samples. 40 Notably, many local, state, and federal agencies that use Rapid DNA systems have provided no statistical or peer-reviewed studies as to their accuracy. There have been similar problems with newer high-sensitivity testing of trace amounts of DNA and mixtures. Independent examination of the source code of probabilistic genotyping software has revealed mistakes and flaws that call into question the accuracy of these tools and their suitability for the criminal justice system. 41

VII. Laws and Regulations Have Not Kept Pace as DNA Advances

Laws have not kept up with advances in DNA technology. The most recent Supreme Court case to address DNA collection was decided nearly a decade ago. In Maryland v. King, the Court upheld, by a slim majority, the warrantless collection of DNA from arrested persons, holding that a DNA swab did not violate an arrestee’s expectation of privacy. 42 The Court solely addressed the collection of DNA from a specific class—arrestees—and limited its analysis to the minimal physical intrusion of the cheek swab and data contained in a CODIS profile. However, this hasn’t stopped police and immigration authorities from citing King to justify everything from warrantless collection of DNA from free people without their knowledge and immigrants who are not subject to arrest warrants to searches through consumer DNA databases.

It is possible, given what we now know about DNA and how much it can reveal, that the Court would decide the case differently today. For example, in a more recent case, the Court held the Fourth Amendment did not allow the warrantless collection of a blood sample from an allegedly intoxicated driver because “blood tests are significantly more intrusive [than breath tests].” 43 This was true even though the police used only the limited blood alcohol information contained in the sample.

There have been some efforts to regulate DNA collection and search by statute or agency policy, particularly with respect to familial searches. The FBI has disclaimed association with familial searching as a matter of policy, 44 while Maryland and the District of Columbia have banned them through statute, citing concerns that “genetic

40 Id.
44 See FBI, supra n.25 (noting “familial searching is not currently performed at NDIS.”).
surveillance” would largely target people of color.\textsuperscript{45} Montana now requires a warrant for familial and partial match searches through its state DNA database.\textsuperscript{46} Other states, such as California, that expressly allow for familial searching limit its use to unsolved criminal investigations where the crime is serious and “has critical public safety implications.”\textsuperscript{47}

Few rules have been put in place to govern other police practices such as searches of consumer genetic genealogy databases and the collection of DNA with or without a person’s knowledge or consent. Montana and Maryland are the only two states that require a warrant to search genetic genealogy databases.\textsuperscript{48} Maryland’s law also places limits on law enforcement’s ability to collect DNA from someone without their knowledge. And after reports that a police agency was collecting DNA from juveniles based on dubiously obtained “consent,” California passed a law limiting the practice.\textsuperscript{49}

\textbf{VIII. CONCLUSION}

We hope these comments assist OSTP in understanding the threats that continued government expansion of DNA collection and use pose to privacy and security, including particular harms to communities of color and immigrants.

If you have any questions, please contact Jennifer Lynch at .

Sincerely,

Jennifer Lynch
Saira Hussain
Electronic Frontier Foundation

\textsuperscript{45} Maryland Public Safety Code § 2-506(d); DC ST§ 22-4151(b); James Rainey, \textit{Familial DNA puts elusive killers behind bars. But only 12 states use it} (April 18, 2018) NBC News https://www.nbcnews.com/news/us-news/familial-dna-puts-elusive-killers-behind-bars-only-12-states-n869711.


\textsuperscript{47} See, e.g., California Dep’t of Justice DNA Data Bank Program, Memorandum of Understanding Familial Searching Protocol, https://oag.ca.gov/sites/all/files/agweb/pdfs/bfs/fsc-mou-06142011.pdf?.

\textsuperscript{48} See supra n. 46.