FAMILIAL SEARCHES, THE FOURTH AMENDMENT, AND GENOMIC CONTROL

JACOB S. SHERKOW,* NATALIE RAM† & CARL A. GUNTER‡

ABSTRACT

In recent years, police have increasingly made use of consumer genomic databases to solve a variety of crimes, from long-cold serial killings to assaults. They do so frequently without judicial oversight per the Fourth Amendment's warrant requirement by using consumer genomic platforms, which store hundreds of thousands or millions of user genomic profiles and enable law enforcement to infer the identity of distant genomic relatives who may be criminal suspects. This Essay puts this practice into context given recent legal and technological developments. As for the law, the Supreme Court in *United States v. Carpenter* has suggested that technologically driven and expansive datasets may be entitled to the full suite of Fourth Amendment protections. As for technology, we describe here the development of a novel technology that allows users to engage in genomic analysis in a secured environment without making such information available to a third party. Taken together, we present a possible technological solution to ensuring Fourth Amendment protections for direct-to-consumer genomic data.

* Professor of Law, University of Illinois College of Law, Professor of Medicine, Carle Illinois College of Medicine, Professor, European Union Center, Affiliate, Carl. R. Woese Institute for Genomic Biology at the University of Illinois at Urbana-Champaign; Permanent Visiting Professor, Center for Advanced Study in Biomedical Innovation Law, University of Copenhagen, Faculty of Law.

† Professor of Law, University of Maryland Carey School of Law, Adjunct Faculty, Johns Hopkins University Berman Institute of Bioethics.

‡ George and Ann Fisher Distinguished Professor in Engineering, Grainger College of Engineering, Faculty, Institute for Genomic Biology at the University of Illinois at Urbana-Champaign.

This work was funded, in part, by a National Human Genome Research Institute grant, no. R01HG012249-01, received by JSS and CG. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

INTRODUCTION

Police are increasingly making use of consumer genomic databases to investigate a variety of cases, from long-cold serial killings, to abandoned Baby Does, to more recent assaults—and they are doing so nearly always without judicial oversight.[1] This practice—and the attendant lack of search warrants—is largely mediated by consumers storing their genome sequences in databases held and controlled by third-party genomic service providers, such as GEDmatch and FamilyTreeDNA.[2] With a sample from an unknown suspect in hand, police can use data from these services to accomplish long-range familial matching against every database participant—that is, to find the closest genomic relative and, often, surveil them.[3] Under a classic reading of the Fourth Amendment's "third-party doctrine," none of this would be a "search" for which a warrant is required.[4]

But the law is beginning to change—and so is genomic technology. As for the law, the propriety of searches of vast electronic databases under the Fourth Amendment is now contested ground.[5] This would seem to include large-scale genomic databases, some of which include data from tens of millions of individuals that could be extrapolated to hundreds of millions more.[6] In line with a recent Supreme Court opinion, *Carpenter v. United States*, the nature of genomic information is so expansive—like a "detailed chronicle of a person's physical presence"—that it may be entitled to the full suite of Fourth Amendment protections.[7] In addition, unlike papers or personal effects—or really, perhaps, unlike anything else—genomic data

1. *See, e.g.*, Natalie Ram, *Genetic Privacy After* Carpenter, 105 VA. L. REV. 1357, 1359–63 (2019) [hereinafter *Genetic Privacy*] (describing law enforcement use of consumer genomic data); April Rubin, *Woman is Charged in Death of Newborn Abandoned in 1985*, N.Y. TIMES (June 26, 2022), https://www.nytimes.com/2022/06/26/us/mother-arrested-cold-case-baby.html [https://perma.cc/2YPK-WDDF] (describing this use to identify the remains of infants and to arrest and charge their mothers).

2. *See Genetic Privacy*, *supra* note 1, at 1366 (considering "what, if any, authority these platforms have, as a constitutional matter, to facilitate—or prevent—law enforcement access to the millions of genetic profiles that they maintain").

3. *See* Natalie Ram, *Investigative Genetic Genealogy and the Problem of Familial Forensic Identification*, *in* CONSUMER GENETIC TECHNOLOGIES: ETHICAL AND LEGAL CONSIDERATIONS 211, 211–12 (I. Glenn Cohen, Nita A. Farahany, Henry T. Greely & Carmel Shachar eds., Cambridge Univ. Press 2021) [hereinafter *Investigative Genetic Genealogy*].

4. *See* Natalie Ram, Christi J. Guerrini & Amy L. McGuire, *Genealogy Databases and the Future of Criminal Investigation*, 360 SCI. 1078, 1078 (2018) (noting then-current doctrine as applied to consumer genomic platforms).

5. *See infra* Section I.A.

6. *See infra* Section I.B.

7. Carpenter v. United States, 138 S. Ct. 2206, 2220 (2018). *See also Genetic Privacy*, *supra* note 1, at 1375–91 (arguing that genomic data satisfies *Carpenter*'s "test").

need not be *shared* with anyone to be uncovered; genomic information about one person's genomic information can readily be implicated to their relatives, consenting or not.[8] This "entireties" nature of DNA raises difficult Fourth Amendment questions where genomic familial searches are at issue.[9]

As for genomic technology, it is now possible to make use of genomic services for a variety of applications without sharing underlying sequence information or storing genomic data on a third-party server. By using a trusted setting model, we report, here, on our development of technology that would simultaneously give users local control over their genomic information and would limit the information genomic service providers can see when analyzing user data. This gives users of the technology unprecedented control over their genomic data—who it is shared with, who can see its contents, and what information can be derived from its use. If such technology becomes adopted widely—an area we are also currently studying—this would also allow individuals to use many genomic services without entrusting their data to third-party services. Combined with the shifting legal landscape of the third-party doctrine under the Fourth Amendment, this new technology may serve as a bulwark against what others have referred to as "DNA dystopia."[10]

This Essay begins by explaining some basics about the Fourth Amendment, including the connection between warrantless searches, reasonable expectations of privacy, and the third-party doctrine. It then discusses this in the context of genomic data, specifically. Part II describes a model for new, trusted setting genomic technology currently being developed and how it obviates the need for consumers to rely on third-party genomic services. Part III then discusses how this new technology fits in the current—and likely future—Fourth Amendment framework. We conclude with the insight that one solution to removing genomic data from the Fourth Amendment's third-party doctrine—regardless of its continued application —may be technological rather than legal.

8. *See* Natalie Ram, *DNA by the Entirety*, 115 COLUM. L. REV. 873, 877–90 (2015) [hereinafter *DNA by the Entirety*] (identifying ways in which genomic data about one individual may be used to learn about or identify her genomic relatives).

9. *See id.* at 919–29.

10. Elias Rios III, Note, *DNA Dystopia: How the National Security Apparatus Could Map the Entire Genome of America Without Violating the Fourth Amendment or the Constitutional Right to Privacy*, 87 BROOK. L. REV. 1387, 1387 (2022).

I. FAMILIAL SEARCHING AND THE FOURTH AMENDMENT

A. FOURTH AMENDMENT BASICS

The Fourth Amendment enshrines "[t]he right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures."[11] This protection generally requires that the government obtain a warrant, supported by probable cause, before performing a search that is intended to discover criminal conduct.[12] But the Fourth Amendment, by its own terms, only regulates the government's ability to conduct "searches" and "seizures."[13] If the conduct at issue does not constitute a search or seizure, Fourth Amendment scrutiny simply does not apply. To determine whether a search has occurred, courts use the "reasonable expectation of privacy" standard originated in *Katz v. United States*. Under this standard, a search occurs when the government seeks to examine a place, thing, or information in which an individual has an "expectation of privacy . . . that society is prepared to recognize as 'reasonable.' "[14]

In a pair of cases in the 1970s, however, the Supreme Court appeared to carve out from the reasonable expectation of privacy standard any information an individual shares with a third party from whom law enforcement subsequently obtains that information.[15] In *United States v. Miller* and *Smith v. Maryland*, the Court reasoned that the data at issue—bank records and telephone numbers, respectively—was not really private or confidential at all.[16] As the Court explained, the defendants had "voluntarily conveyed" the information at issue to a third party—a bank and the telephone company—and so "assumed the risk" that those records "would be divulged to police."[17] In the decades that followed, lower courts often interpreted *Miller* and *Smith* as establishing a near categorical third-party doctrine, under which "if you share information, you do not have an expectation of privacy

11. U.S. CONST. amend. IV.

12. *See* Carpenter v. United States, 138 S. Ct. 2206, 2221 (2018) ("Although the 'ultimate measure of the constitutionality of a governmental search is "reasonableness," ' our cases establish that warrantless searches are typically unreasonable where 'a search is undertaken by law enforcement officials to discover evidence of criminal wrongdoing.' ") (quoting Vernonia Sch. Dist. 47J v. Acton, 515 U.S. 646, 652–53 (1995)).

13. U.S. CONST. amend. IV.

14. Katz v. United States, 389 U.S. 347, 361 (1967) (Harlan, J., concurring)).

15. *See* Smith v. Maryland, 442 U.S. 735, 745–46 (1979); United States v. Miller, 425 U.S. 435, 440 (1976). *See also Genetic Privacy*, *supra* note 1, at 1369–70 (describing the genesis of the third-party doctrine).

16. *See Smith*, 442 U.S. at 742; *Miller*, 425 U.S. at 442.

17. *Smith*, 442 U.S. at 745. *See also Miller*, 425 U.S. at 443 ("The depositor takes the risk, in revealing his affairs to another, that the information will be conveyed by that person to the Government.").

in it."[18] The third-party doctrine does (and has) posed difficulties in a modern age when much of our information is not tucked away on parchment in writing desks, but rather is on the Internet or in e-mails, DMs, and texts. These forms of communication and inquiry are, as a matter of technological necessity, shared with third parties like ISPs, platforms, and other service providers.

The Supreme Court has only recently begun to grapple with these realities, curtailing the broadest interpretations of the third-party doctrine. Most significantly, in *Carpenter*, the Court held that government access to a week's worth of an individual's historical cell phone location data—data that is compiled and held by cell phone companies—amounts to a search subject to the Fourth Amendment and typically requires a warrant.[19] In other words, cell phone location data is data in which an individual may maintain a reasonable expectation of privacy against warrantless government access and use.[20]

B. GENOMICS AND THE FOURTH AMENDMENT

Consumer genomic platforms pose special challenges under these Fourth Amendment doctrines. Genomic data differs from most other information in that it is nearly always both involuntarily and immutably shared among genomic relatives.[21] That is, we do not choose our genomic relatives, and we cannot sever that genomic tie. When one individual has participated in a consumer genomic platform, many hundreds of her genomic relatives may be implicitly identifiable through their patterns of shared genomic variations.[22]

Police have exploited these features of genomic relatedness to investigate crimes. Most famously, in the case of the Golden State Killer, law enforcement sequenced DNA recovered from victims and crime scenes, uploaded the resulting DNA profile to several consumer genomic platforms,

18. Margot E. Kaminski, Response, Carpenter v. United States: *Big Data Is Different*, GEO. WASH. L. REV. ON THE DOCKET (July 2, 2018), https://www.gwlr.org/carpenter-v-united-states-big-data-is-different [https://perma.cc/TMP6-E43L] (describing this rule as a "central truism of U.S. privacy law" until *Carpenter*). *See also* William Baude & James Y. Stern, *The Positive Law Model of the Fourth Amendment*, 129 HARV. L. REV. 1821, 1871 (2016) ("It is black-letter law under *Katz* that people don't have any Fourth Amendment protection for information given to a third party.").

19. Carpenter v. United States, 138 S. Ct. 2206, 2223 (2018).

20. *Id.* at 2219–20.

21. *Investigative Genetic Genealogy*, *supra* note 3, at 215–16; *DNA by the Entirety*, supra note 8, at 898–906.

22. *See Investigative Genetic Genealogy*, *supra* note 3, at 219.

and investigated the resulting family tree to identify a suspect.[23] Major players in this field, including both GEDmatch and FamilyTreeDNA, permit users to opt out of familial matching for at least some law enforcement purposes. But this is no guarantee of privacy. In several known cases, law enforcement investigators have been able to make matches with the full complement of database participants simply by posing as ordinary users, rather than disclosing their law enforcement or investigative status.[24] By one estimate, law enforcement access to as little as two percent of the U.S. population could make as many as ninety percent of Americans of European descent identifiable.[25]

These investigative efforts thus rely on the ready availability of consumer genomic profiles with which to make familial matches. Both individuals and companies have acted in ways that invite law enforcement scrutiny. For one thing, individual genomic "data dumps" are now routine, with individuals having little alternative if they seek genomic analysis for personal use. For another, many of the platforms in the consumer genomic market have opted to accept uploads of genomic profiles developed elsewhere, usually without robust security features that would prevent sharing impermissibly sourced data.[26] GEDmatch, one of the principal consumer genomic platforms utilized by law enforcement, does no DNA analysis itself. Its database consists solely of DNA data developed elsewhere.[27]

Courts have largely declined to reach the issue of whether law enforcement use of consumer genomic data constitutes a search subject to Fourth Amendment scrutiny, particularly after *Carpenter*. In one case, a court declined to reach this question on the merits, holding that the criminal defendant lacked standing to challenge the search at all.[28] This court did not

23. *See Genetic Privacy*, *supra* note 1, at 1359.

24. *See, e.g.*, State v. Westrom, No. 27-CR-19-3844, at 3 (Minn. Dist. Ct. Oct. 4, 2021) (memorandum opinion and order denying motion to suppress) ("In January 2019, investigators, with help from a genetic genealogist, submitted [a DNA] data file to commercial genealogical websites FamilyTreeDNA and MyHeritage under the pseudonym Steve Bell.").

25. *See* Yaniv Erlich, Tal Shor, Itsik Pe'er & Shai Carmi, *Identity Inference of Genomic Data Using Long-Range Familial Searches*, 362 SCI. 690, 690 (2018).

26. *See, e.g.*, *How It Works*, GEDMATCH, https://www.gedmatch.com/how-it-works [https://perma.cc/XE8J-DD7C] (instructing users how to download their DNA files from one service provider and then upload that data to GEDmatch); *Upload DNA Data*, FAMILYTREEDNA, https://www.familytreedna.com/autosomal-transfer [https://perma.cc/Z3H5-Z8GB] (inviting users to "[u]pload [their] Ancestry DNA™, 23andMe©, or MyHeritage™ autosomal DNA data to FamilyTreeDNA and connect with new relatives for FREE").

27. *See* GEDMATCH, *supra* note 26.

28. *See* State v. Burns, No. FECR129718, at 5–8 (Iowa Dist. Ct. Feb. 6, 2020), *aff'd*, 988 N.W.2d

discuss or even cite *Carpenter*. In another case, a court held that individuals lack any expectation of privacy in their own genomic material when used for identification purposes and therefore cannot possibly assert such an expectation in genomic information shared with others.[29]

But the merits of this question are not so pellucidly clear. The involuntary and immutable nature of genomic relatedness confounds application of the third-party doctrine, which takes voluntariness in data sharing as its foundation.[30] Analysis might—or might not—be more straightforward for the individual whose DNA data is sequenced by or uploaded to a consumer genomic platform. Under the traditional third-party doctrine, no expectation of privacy would be possible in this shared data.[31] After *Carpenter*, the answer is not free from doubt. For now, it may be fair to characterize use of consumer genomic platforms as "voluntary," though one of the authors has argued elsewhere that this may be changing (and that these changes are relevant to the Fourth Amendment inquiry).[32] At least one Justice in *Carpenter* opined that this kind of genomic data ought to receive full Fourth Amendment protection.[33]

In sum, the state of constitutional genomic privacy is unsettled. If individuals seeking to learn about their genomic data—whether to obtain individual information about, for example, health risks or ancestry, or to find currently living relatives—also desire to ensure their privacy, they will need to rely on other tools, if possible, to do so.

352 (Iowa 2023). The Iowa Supreme Court did not consider the Fourth Amendment claims regarding law enforcement use of consumer genomic data. *See* 988 N.W.2d at 359–69. *See also* State v. Downs, No. 4FA-19-00504CR, at 3–5 (Alaska Super. Ct. Dec. 6, 2021) (denying motion to suppress and concluding both that defendant lacked standing to challenge law enforcement use of consumer geonomic data and that such law enforcement use did not violate the Fourth Amendment).

29. *See* State v. Westrom, No. 27-CR-19-3844, at 8–9.

30. *See Investigative Genetic Genealogy*, *supra* note 3, at 221–23; *DNA by the Entirety*, *supra* note 8, at 904–06.

31. *See* Ram, Guerrini & McGuire, *supra* note 4, at 1078–79 (2018) (noting then-current doctrine in 2018 and anticipating *Carpenter*).

32. *See Genetic Privacy*, *supra* note 1, at 1389.

33. *See* Carpenter v. United States, 138 S. Ct. 2206, 2262 (Gorsuch, J., dissenting) ("Can [the government] secure your DNA from 23andMe without a warrant or probable cause? *Smith* and *Miller* say yes it can—at least without running afoul of *Katz*. But that result strikes most lawyers and judges today—me included—as pretty unlikely.").

II. NEW GENOMIC TECHNOLOGIES AND GENOMIC CONTROL

A. CONTROL OVER DIRECT-TO-CONSUMER GENOMIC DATA

Today, individuals interested in analyzing their genomic data typically use a variety of direct-to-consumer ("DTC") genomic companies, such as 23andMe or Nebula Genomics.[34] Users submit a DNA sample—usually, saliva in a tube—to a DTC genomic company that processes user samples, sequences their DNA, and makes the resulting data available, online and for download, from a password-protected website.[35] Often these sequence services are paired with a variety of analyses, such as health traits based on underlying genomic sequences or information about possible genomic ancestry.[36] Other genomic services, like GEDmatch, allow users to upload their genotype data and conduct other analyses, such as searching for a genomic relative.[37] And still other genomic services, like FamilyTreeDNA and MyHeritage, engage in DNA sequencing and accept uploads of genotype data sequenced elsewhere.[38] In all of these cases, though, the genomic service companies control both the underlying data—that is, users' genomic data—and the analysis tools used to interpret them.

Despite users' lack of control over their genomic data, demand would suggest that millions of consumers seem to be unperturbed.[39] The reasons for this are varied and complex, the subject of its own literature on what individuals feel they learn from knowing about their genomes.[40] But some of this immediate lack of concern stems from technological necessity: the companies providing genomic services *need* ready access to users' genomic data to operate their computational analyses and therefore cannot conduct their analyses from afar. This technological need arises from several practical realities of genomic data. Sequence data—especially whole genome sequences—can make up enormous file sizes. Lower grade whole genome sequences can take up, on average, eighty gigabytes, almost a tenth

34. 23ANDME, https://www.23andme.com [https://perma.cc/9ZRC-BW7K]; NEBULA GENOMICS, https://nebula.org/whole-genome-sequencing-dna-test [https://perma.cc/2EJU-6PE3].

35. *E.g.*, *How It Works*, 23ANDME, https://www.23andme.com/howitworks [https://perma.cc/9LNQ-G7UB].

36. *See id.*

37. *See* GEDMATCH, https://www.gedmatch.com [https://perma.cc/FK76-ZAG7].

38. *See* FAMILYTREE.COM, https://www.familytree.com [https://perma.cc/42GH-8246]; MYHERITAGE, https://www.myheritage.com [https://perma.cc/3EYX-7DL5].

39. NAT'L ACADS. OF SCIS., ENG'G & MED., EXPLORING THE CURRENT LANDSCAPE OF CONSUMER GENOMICS: PROCEEDINGS OF A WORKSHOP 10–11 (2020) (estimating the market to be roughly one hundred million customers worldwide).

40. *See id.* at 13–16.

of a standard hard drive.[41] Higher grade genomic data can be several multiples of that. In surveying the landscape of "big data" applications, one commentator recently noted, "the largest big data around us is genomic data."[42] As a result, users' control over their own genomic data has been impractical because such control would require users to upload vast amounts of data for each application. In addition, it has been difficult for users to gauge which genomic analyses services are trustworthy apart from the companies that conducted their original sequencing. Users who wish to know about what their genomes say about their health, physical traits, ancestry, parentage, or genomic relatives have consequently had little choice but to use services that possess physical control over their genomic data. There has been no other game in town.

B. NEW TECHNOLOGIES FOR GENOMIC CONTROL

This, however, is changing. Recent advances, including those being developed by some of the authors, seek to solve the problems inherent in the classic DTC genomic sequencing model. These technologies give users the potential to exert local control over their genomic data and deploy trusted third-party analyses of that data without necessarily revealing its contents. One easy-to-understand strategy aims to bring the "computation to the data."[43] DTC genomic services have traditionally placed genomic data at the site of computation. For instance, 23andMe performs tests within its own server system.[44] But smaller sizes of genotype data like those generated from "SNP chips," combined with advances in software platforms, make it increasingly feasible to carry out genomic tests elsewhere.[45] For instance, DTC users typically download their data from 23andMe and upload it to GEDmatch.[46] It is increasingly feasible for users to download both the data and the code for a computation and execute the computation within a trusted user environment.

41. *See* Kenneth Joohyun Han, *Big Data Among Big Data: Genome Data*, 3BILLION: BLOG (Feb. 21, 2022), https://3billion.io/blog/big-data-among-big-data-genome-data [https://perma.cc/WV3U-DGTA].

42. *Id.*

43. D. Kalra, B.N. Downs, D.M. Opheim, W. Hale, L. Xi & L.A. Donehower, *A Practical Example of Bringing Computation to Data*, 25 J. BIOMOLECULAR TECHS. S1, S5 (Supp. 2014).

44. *See How It Works*, *supra* note 35.

45. *See* 296 Free Single Nucleotide Polymorphism (SNP) Analysis Tools - Software and Resources, DR MARTTI & DR LLOYD, https://bioinformaticshome.com/tools/SNP-tools.html [https://perma.cc/Z67L-GTK9] (listing hundreds of third-party software tools).

46. *See, e.g.*, Tomohiro Takano, *GEDmatch: Do More with Your 23andMe Raw Data*, GENOMELINK: BLOG (June 29, 2019), https://blog.genomelink.io/posts/gedmatch-do-more-with-your-23andme-raw-data [https://perma.cc/NP7S-VR6B].

The technology largely operates around a trusted platform—akin to that deployed on most smartphones—to allow users local control over their data. Users could, for example, have their genome sequenced by a commodity DNA services company (one that does not permanently store client sequences) and then deposit the resulting genomic data on a trusted device, like a smartphone. Software on that device would then sequester the genomic data in a trusted environment from the device's other applications.

By keeping genomic data within this trusted environment, the software would only allow certain genomic services applications access to the genomic data. The services allowed could be limited by only those that have been verified—like being able to download only those apps that have been verified within an app store. In addition, the trusted platform can limit *which* genomic information is shared with the service, even within the trusted environment. It could do this by, for example, showing the third-party service some genomic information—for example, only a single gene—while denying access to the rest. It could also determine whether the service is able to send information outside of the user's trusted environment. By containing *both* the genomic data *and* the third-party service within this trusted environment, the secured device can limit what information—if any—gets reported back out to the third-party service on the Internet.

The ultimate effect of the technology is something like an at-home pregnancy test. A third-party service provider—here, the marketer of the pregnancy test—provides the test directly to a consumer. The consumer, once provided with the material to conduct the test, then uses the device in a secure environment. The results are not known to the marketer of the test and the consumer can share—or hide—the results as they see fit.

While the concept for this new technology is exciting, more work will be needed to fully realize its privacy potential. This is especially true when it comes to familial searches that require genomic information from not only a given user but also from others—indeed, all other users against whom one wishes to assess their genomic relations. Nonetheless, additional techniques based on recent advances in cryptography are making it feasible to share data without revealing more than the parties in the exchange expect to share.[47] With the right tools—and enough user buy-in—it is possible to assess

47. *See, e.g.*, Dominic Deuber, Christoph Egger, Katharina Fech, Giulio Malavolta, Dominique Schröder, Sri Aravinda Krishnan Thyagarajan, Florian Battke & Claudia Durand, *My Genome Belongs to Me: Controlling Third Party Computation on Genomic Data*, 2019(1) PROCS. ON PRIV. ENHANCING TECHS. 108, 109 (describing a cryptographic system called METIS "that allows secure computation on encrypted genomic data with the important feature that the data owner controls the type of functions that can be computed").

genomic "matches" without revealing underlying genomic information itself. This would have the effect of blunting the entireties nature of genomic sharing, in which user deposits of genomic sequences can be attributed to relatives.[48] This would contrast sharply with the approach now used by parties like GEDmatch, which must be trusted to hold significant genomic data. Progress in this area suggests that it is possible to have personal genome management, effective sharing of genomic data, and appropriate safeguards for the privacy risks involved in DTC genomics.

III. FAMILIAL SEARCHING WITH GENOMIC CONTROL

What does this trusted setting technology—and for that matter, users' control over their genomic data—mean for Fourth Amendment purposes? We identify five distinct advantages of the trusted setting model over traditional consumer genomics.

First, mechanics. Continuous user control over genomic data obviates application of any permutation of the third-party doctrine. Recall that the third-party doctrine depends, at its core, on some form of voluntary sharing of data with a third-party entity from whom law enforcement subsequently gains access. But under the trusted setting model, no third-party entity would ever gain control over the data. Rather, the genomic data itself would be owned and controlled by users.

To be sure, a commodity sequencer will, at some point, come into contact with genomic sequence. After all, ordinary members of the public do not engage in genome-wide sequencing in at-home laboratories. But if the sequencing entity does not keep a record—as could be arranged by contract—then there will no longer be a third-party record to which law enforcement might gain access. Indeed, as a matter of business, many commodity sequencers do not keep records of the genomic sequences they generate—it is far too expensive to perpetually store such large data files.

Second, the mechanics of system described here means that—unlike, say, bank records—the underlying *information* is not voluntarily conveyed to a third party to control (or otherwise). Indeed, for some applications, the technology is such that the information is hidden *even* to a third party. Using a bank as an analogy—with an admission that all analogies are imperfect—genomic data would, at most, be like a safe deposit box to which the possibility of access means neither knowledge nor control of its contents. And, generally, warrants are required for safe deposit boxes.[49]

48. *See DNA by the Entirety*, *supra* note 8, at 877–90.

49. United States v. Thomas, No. 88-6341, 1989 U.S. App. LEXIS 9628, at *6 (6th Cir. July 5,

Third, the trusted setting model shatters the notion that users who engage in DTC genomics do not have a reasonable expectation of privacy. The use of the privacy enhancing technologies here suggest users have a *strong* expectation of privacy, like those who password protect computers or encrypt hard drives.[50] Doing so, particularly when there are other options out there like 23andMe, indicates that users have an expectation of privacy in their data and that this expectation is reasonable.

The existence of a trusted setting model should not, however, imply that those who use more traditional consumer genomic services lack a reasonable expectation of privacy. Many users may remain unaware of more privacy protective options or may be unable to utilize them for financial, technological, or other reasons.[51] Just as the existence of email encryption has not obviated the reasonable expectation of privacy people have in the contents of their unencrypted emails,[52] the existence of a trusted setting model ought not to undermine whatever expectation of privacy is reasonable in consumer genomic data more broadly.

Nonetheless, the existence and use of a trusted setting model would definitively demonstrate for the consumer who uses it that privacy is expected and reasonable, regardless of the constitutional status of other consumer genomic data.

Fourth, user control over genomic data better comports with traditional Fourth Amendment notions of personal papers and effects—for which reasonable expectations of privacy unquestionably exist. In the trusted setting model, genomic data exists solely in the possession and control of the individual holding the trusted device. No outside entity can gain access to this data except by permission of its owner. Any such access may be restrictive or permissive. In this way, genomic data, not unlike computer hard drive data, is much like the contents of a letter kept in physical custody of

1989) (per curiam) ("[C]itizens have legitimate expectations of privacy in the contents of their safe deposit boxes.").

50. United States v. Zhu, 23 F. Supp. 3d 234, 238 (S.D.N.Y. 2014) ("Zhu's use of passwords and encryption weighs in favor of finding a reasonable expectation of privacy.").

51. *Cf.* Neil Richards & Woodrow Hartzog, *The Pathologies of Digital Consent*, 96 WASH. U. L. REV. 1461, 1478 (2019) (identifying ways in which privacy can be compromised without genuine knowledge or voluntariness and explaining "why consumers might understandably care about their privacy *and* agree to data practices that undermine their privacy and expose them to the risks of informational harms").

52. *See* Carpenter v. United States, 138 S. Ct. 2206, 2230 (2018) (Kennedy, J., dissenting) (approvingly citing case holding that "e-mails held by Internet service provider" are protected by the Fourth Amendment without reference to encryption); *cf.* Riley v. California, 573 U.S. 373, 386–91 (2014) (holding that cell phones may not be subject to warrantless searches incident to an arrest and rejecting the risk of data encryption as a reason to permit such searches).

the owner.

Significantly, under the trusted setting model, genomic information about an individual—or her genomic relatives more broadly—cannot be learned by inference. Law enforcement use of consumer genomic data depends on inference for its successes, as virtually all arrests following this method to date have involved relatives of consumer genomic platform users. Arrested individuals, in other words, did not dump their own genomic data on a consumer platform; rather, those individuals were implicitly findable through the genomic data they share with genomic relatives in predictable patterns. The trusted setting model, by contrast, makes it possible for users to restrict access to their genomic data to a greater degree and thus make information about themselves and their genomic relatives more difficult to infer. Indeed, the trusted setting model enables an individual to exercise control such that an outside entity learns only that a particular user sought out a particular analysis, but not whether that analysis was successful or what result it returned. In this way, the minimum amount of data shared might be akin to the limited information available on the envelope of a letter, while the contents remain hidden and secure from prying eyes (government or otherwise). Employed in this way, users can have control over genomic information and use DTC services with—in our opinion—reasonable assurances that a warrant would be needed to search.

CONCLUSION

Despite the public's general antipathy toward warrantless searches, demand remains strong in the DTC genomic market. As recently as 2019, GEDmatch recorded as many as a thousand new users *every day*.[53] It remains unclear whether this demand reflects people's willingness to risk law enforcement use of their genomic data or simply a lack of knowledge or understanding of these risks. Nonetheless, some of that choice has been one of necessity: in order to participate in a consumer genomic platform, users have been required to share their genomic data with service providers and permit those services providers to maintain a copy of the genomic data in their database. There were few other technological options.

But, as described here, recent technological advances can allow users to both participate in DTC genomic analysis and maintain control of their own genomic data. This technology allows users to keep their genomic data

53. Sarah Esther Lageson, *Privacy Concerns Don't Stop People from Putting Their DNA on the Internet to Help Solve Crimes*, CONVERSATION (June 7, 2019), https://theconversation.com/privacy-concerns-dont-stop-people-from-putting-their-dna-on-the-internet-to-help-solve-crimes-118091 [https://perma.cc/C3YT-TRLZ].

in a computationally trusted setting, only disclosing a portion of it to verified services—or, in some cases, to receive the analyses such services provide without disclosing the underlying sequence at all.

Regardless of the ultimate constitutionality of current investigative searches involving consumer genomic data, the genomic control tool described here would preserve individuals' expectations of privacy in their genomes. It would keep it out of the hands of third parties, some of whom may be inclined to share such data with law enforcement even in the absence of a warrant. And it better comports with classical conceptions of Fourth Amendment maxima, such as an individual's private letters or personal effects.

For a long time, new technologies have significantly challenged interpretations of the Fourth Amendment. Now we have a technology that is as much of a legal answer to the Fourth Amendment as it is a technological one.