NOTE

DEMANDING TRUST IN THE PRIVATE GENETIC DATA MARKET

Benjamin T. Van Meter†

The aspirations of private genetic testing to aid self-discovery, improve medicine, or advance research often obscure a multimillion-dollar market for customers' genetic data. This market's legality and operation is founded on testing companies' promise that in selling individual genetic information, they will still guard their customers' privacy. These companies rely on the techniques of de-identifying and aggregating data to create massive genetic databases that they can sell to both public and private researchers. By selling de-identified, aggregated data, these companies avoid nearly all

† J.D., Cornell Law School, 2020; B.A., Tufts University, 2013; Articles Editor, *Cornell Law Review*, Volume 105. I am very grateful to Professor Robert Hockett and Rohan Grey for their help and advice in preparing this Note. Many thanks as well to the members of the *Cornell Law Review* for their tireless editing and insightful recommendations. Finally, thanks to my eternal editors, my parents Elena Sansalone and Jan Van Meter, who taught me to write and have been correcting me ever since.

regulations limiting the collection and disclosure of medical information. Instead, customers are promised that internal measures can assure "privacy by design." At least 26 million people have already been tested by private companies and their data is sold on this promise: their privacy will be preserved.

However, a growing body of research indicates that genetic information is exceedingly difficult to "de-identify" because an individual's DNA sequence and other genetic data are some of the most identifying pieces of information about them. Multiple researchers have been able to use publicly available, supposedly "de-identified" genetic data and trace it back to the individuals who donated their DNA. Moreover, private companies like AncestryDNA (Ancestry) and 23andMe have based their business models on being able to sell customer genetic data, relying on de-identification and aggregation to justify their lack of regulation and the continued operation of a private genetic data market. As this market grows in size, it is also attracting more and more interest from insurance carriers, employers, law enforcement, and a host of other groups who see new potential in acquiring individuals' genetic data.

This Note argues that to prevent the most damaging consequences of the trade in genetic data, U.S. law should impose tailored fiduciary duties on private genetic testing companies to ensure that their business practices do not harm their own customers. These testing companies rely on their customers' genetic information to turn a profit, while all of the risk of this information's exposure or misuse falls on customers. This Note will proceed as follows: Part I will describe the fundamental difficulties of de-identifying and aggregating genetic data to the point that it cannot be re-identified; Part II discusses how de-identification and aggregation serves to obscure customers' rights in their own genetic data while allowing testing companies to evade federal privacy laws; Part III argues that Ancestry and 23andMe, genetic testing's two largest companies, use privacy agreements that largely deprive customers of any rights in their genetic data in order to keep the data marketable to as many buyers as possible, and Part IV argues that the concept of an information fiduciary should be applied to private genetic testing companies to counter these companies' massive informational advantage over their customers and to guard against genetic data's potential for abuse.

I
THE PARADOX OF "ANONYMOUS" GENETIC DATA

A. De-identified and Aggregated Data

The ability of private companies to trade in medical data rests on their ability to de-identify[1] and aggregate the data they store. De-identification is particularly central to the medical data market, since U.S. law generally allows medical information to be shared or sold only if informed consent is obtained or if the data is de-identified.[2] De-identification has been essential for the creation and functioning of a market in patient data because of the flexibility it offers and the problems inherent in obtaining consent. As one academic observed: "[T]he market for patient data is virtually all for anonymized data."[3]

Informed consent's limitations, at least to its critics, seriously undercut resulting data's value and encourage de-identified data as a more attractive asset. Obtaining consent to use personal medical information for research has been criticized by many academics and medical practitioners because it skews population representation in a given data set[4] and can hinder medical research.[5] It can also be impossible to predict valuable secondary uses for data at the time it is collected, but going back to obtain consent for a new use is often unfeasible.[6] Moreover, "informed" consent can be manipulated or manufactured through suggestive wording, patient emotions, or other tactics at the outset.[7] By comparison, de-identified data gener-

---

1 "De-identification" is more commonly used in North America whereas "anonymization" is more often used in the European Union. For consistency, I will refer to the technique as the former throughout this Note.

2 *See* Khaled El Emam et al., *Anonymizing and Sharing Individual Patient Data*, 350 BRITISH MED. J. 1 (2015), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4707567 [https://perma.cc/3LPH-J4YG].

3 Marc A. Rodwin, *Patient Data: Property, Privacy & the Public Interest*, 36 AM. J.L. & MED. 586, 609 (2010).

4 Jorge L. Contreras, *Genetic Property*, 105 GEO. L.J. 1, 30 (2016); *see also* Michelle Kho et al., *Written Informed Consent and Selection Bias in Observational Studies Using Medical Records: Systematic Review*, 338 BRITISH MED. J. (2009), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2769263/ [https://perma.cc/U6SL-FFUQ] (concluding that differences between participants and non-participants may threaten the validity of test results that require consent).

5 Contreras, *supra* note 4, at 6–7.

6 *See id.* Contreras, *supra* note 4, at 31; El Emam et al., *supra* note 2, at 1.

7 *See, e.g.*, Contreras, *supra* note 4, at 29 (discussing how misleading wording or questions can manipulate informed consent, leading physicians to view it as an "empty charade") (quoting Howard Brody, *Transparency: Informed Consent in Primary Care*, 19 HASTINGS CTR. REP. 5, 5 (1989)); Jennifer A. Drobac & Oliver R. Goodenough, *Exposing the Myth of Consent*, 12 IND. HEALTH L. REV. 471, 489, 514 (2015) (discussing studies showing how patient empathy and emotion affect patients' willingness to consent); George P. Smith, II, *The Vagaries of Informed*

ally is subject to much lighter regulation in both the United States and Europe than data with personal identifiers. For instance, de-identified data is not designated in the United States as personal information at all, so it can largely be shared for research purposes without consent.[8] Accordingly, custodians of medical data rely on de-identification to share or sell data collected without consent or data collected consensually but for other purposes.

Unfortunately, the actual definition of de-identified data is mildly elusive, leading to a variety of sometimes inconsistent anonymization practices.[9] At the federal level in the United States, the Health Insurance Portability and Accountability Act (HIPAA) defines de-identified health data as health information "with respect to which there is no reasonable basis to believe that the information can be used to identify an individual."[10] This definition has been generously described as ambiguous, but research institutions bound by federal ethics guidelines have developed some technical procedures in common.[11] Generally accepted standards and guidelines in biomedical research divide identifying information into direct and quasi-identifiers.[12] Direct identifiers are information that permit "direct recognition [of] or communication with" the individual, such as names, email addresses, phone numbers, and social security numbers.[13] Quasi-identifiers are features of the individual that can only be used to "indirectly identify individuals," such as their date of birth, death, ZIP code, or ethnicity.[14] Given the relative utility of both direct and quasi-identifiers in identifying a person, data custodians aim to ensure that on a technical level, the probability of identifying a particular record is small, though the possibility can never be zero.[15] This process will normally involve removing direct identifiers and sufficient quasi-identifiers that the dataset is "reasonably" anonymous. Even given its ambiguities, research institutions widely consider using de-identified data for secondary uses or

---

Consent, 1 IND. HEALTH L. REV. 111, 112–13 (2004) (discussing the difficulty in setting limits for morally permissible manipulation of patients' consent).

8 *See* El Emam et al., *supra* note 2, at 1.

9 *Id.*

10 *Id.*

11 *See id.*

12 *Id.* at 2.

13 *Id.*

14 *Id.*

15 *See id.* at 1–2.

sharing with third parties as of minimal risk to individual privacy.[16]

By comparison, data aggregation as a technique is both distinct from de-identification and its mirror image. While de-identification aims to purge a dataset of information singling out individuals, aggregation aims to collect and standardize data identifying patterns within a group. Both have the purpose of providing a large set of practically useful data while preserving individuals' privacy.[17] Indeed, both select certain kinds of data either to include or exclude from a dataset before it is transferred to a third-party. In some ways, their difference is one of perspective. For example, a cellphone provider's dataset of all its users' locations for the last year could be de-identified by removing all data regarding users' names, phone numbers, and billing addresses while it can be aggregated by only publishing users' timestamped locations.[18] Especially with sensitive information like medical data however, both techniques are employed to the same effect with the same concerns in mind.

These two techniques are also essential to private genetic testing companies like 23andMe and Ancestry for sidestepping regulations and creating a market for their customers' data. For instance, 23andMe promises that customers' "personal" or "individual-level" information will only be given to third parties with customers' explicit consent, but that "de-identified" or "aggregate" information may be shared more freely.[19] According to

---

16 *See* Aaron J. Goldenberg et al., *IRB Practices and Policies Regarding the Secondary Research Use of Biospecimens*, 16 BRITISH MED. J. 3 (2015), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4426182/ [https://perma.cc/M8A9-ZLJ3] (finding that 98% of responding institutions "consider studies using anonymized biospecimens to be no greater than minimal risk").

17 *See* David W. Craig et al., *Assessing and Managing Risk When Sharing Aggregate Genetic Variant Data*, 12 NATURE REVIEWS OF GENETICS 730, 731 (2011), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3349221/ [https://perma.cc/V94F-B3UX] (stating that aggregated genetic data provides some level of privacy protection but that "some degree of residual identifying information remains" in many genetic datasets).

18 *See* Fengli Xu et al., *Trajectory Recovery from Ash: User Privacy is NOT Preserved in Aggregated Mobility Data*, 2017 PROC. INT'L CONF. WORLD WIDE WEB 1241, 1241, https://arxiv.org/pdf/1702.06270.pdf [https://perma.cc/5G43-AW2Z].

19 According to 23andMe's website, "personal" information includes direct identifiers like an individual's name, address, and genotypes (i.e., their DNA sequence) while "individual-level" information includes "information about a single individual's genotypes, diseases or other traits/characteristics . . . ." *Privacy Statement*, 23ANDME § 1(3)–(4), https://www.23andme.com/about/privacy/#full-privacy-statement [https://perma.cc/GQ93-JYVF] (last updated Jan. 1, 2020) [hereinafter *23andMe Privacy Statement*]. Under these definitions, data identify-

their corporate website, "de-identified information" is information that has been "stripped" of identifying data such that a person "cannot reasonably be identified as an individual" from the dataset.[20] Similarly, "aggregate information" is genetic data that has been combined with other users' data "such that no specific individual may be reasonably identified."[21] While a person's individual-level and de-identified information cannot be shared with third-parties without that person's explicit consent,[22] 23andMe can share aggregate information without restriction.[23] Moreover, if a customer does consent to sharing de-identified data with third-parties, but later revokes consent, their information cannot be removed from ongoing or completed research.[24]

Ancestry similarly uses de-identification and aggregation to increase its control over customers' data while limiting the role of their consent. Its privacy policy defines "personal information" as any information, including genetic information, that can "reasonably be linked back to" the individual.[25] While Ancestry will not sell your genetic data to marketers, insurance companies, or employers without explicit consent, it does use customers' genetic information to improve its services, create new products, and conduct "scientific, statistical, and historical research."[26] Indeed, Ancestry allows itself broad use of "de-identified" personal information,[27] but its current privacy policy does not actually provide a standard for de-identifying data.[28] Similarly, Ancestry treats aggregated information as no longer personal information, and so may disclose it in marketing or in scientific publications.[29] By ostensibly breaking the link between an individual and their genetic data, de-identifica-

---

ing an individual's DNA sequence in whole or part would seem to qualify as both "personal" and "individual-level" information. *See id.*

20 *Id.* § 1(2).

21 *Id.* § 1(1).

22 *Id.* § 3(d)–(e).

23 *Id.* § 4(c).

24 *See id.* § 3(d).

25 *Privacy Statement*, ANCESTRY § 1, https://www.ancestry.com/cs/legal/privacystatement [https://perma.cc/2KHQ-C9EY] (last updated Dec. 23, 2019) [hereinafter *Ancestry Privacy Statement*].

26 *Id.* §§ 6–7.

27 *See Ancestry Privacy Statement*, *supra* note 25, § 2.

28 A prior privacy statement from Ancestry that was effective until 2016 indicated that de-identified data is "anonymized" and does not personally identify its source individual. *See Ancestry Privacy Statement*, *supra* note 25, § 7(i). The best working definition for Ancestry's de-identification standard might be the opposite of its definition for personal information, i.e., de-identified data is information which cannot be reasonably linked back to an individual.

29 *See Ancestry Privacy Statement*, *supra* note 25, §§ 7, 9.

tion and aggregation allow testing companies like 23andMe and Ancestry to control, and eventually monetize, their customers' genetic information.

B. Re-identification

Studies are increasingly showing that supposedly de-identified and aggregated datasets are more amenable to re-identification than previously thought.[30] The essential problem is that it is hard to predict every variable that will identify an individual. For instance, a 2015 study looked at de-identified, aggregated datasets of cell phone users' timestamped locations, finding that between 73% and 91% of individual users could be re-identified.[31] The problem was that even among the undifferentiated data of location and timestamps, individual mobility habits were unique enough to find which data points belonged to a single individual.[32]

In some studies of medical data, researchers and journalists have conducted re-identification attacks on many de-identified datasets (usually to identify systemic weaknesses), only to find that the data has not been sufficiently scrubbed to keep patients anonymous.[33] While many anonymization techniques have focused on wiping direct identifiers from data, some re-identification attacks have shown that a sufficient number of quasi-identifiers can be just as effective in identifying individuals.[34]

Truly anonymizing data is difficult enough with most consumer or medical data, but genetic data presents unique, and arguably insurmountable challenges. One essential reason is that an individual's genome is inherently identifying.[35] Moreover, genetic data's de-identification must be somewhat limited, or else it will threaten the data's utility—for a dataset about an individual to be of value to research, it must retain sufficient distinctiveness and connection to an individual to be compared with other datasets.[36] Indeed, some argue that the great prom-

---

30 *See, e.g.*, Xu et al., *supra* note 18 (finding that some individuals could be identified from aggregated cellphone data).

31 *Id.* at 1242.

32 *Id.* at 1243.

33 *See, e.g.*, El Emam et al., *supra* note 2, at 2 (describing two examples of successful re-identification by reporters and academics).

34 *See id.* (discussing the importance of protecting quasi-identifiers in addition to direct identifiers).

35 *See* Contreras, *supra* note 4, at 34 (noting that complete de-identification of genetic information may be impossible).

36 *See id.* at 35.

ise of genetic data is in its ability to provide hyper-personalized medical treatment, so de-identification can render data useless for certain kinds of research.[37]

Given the vague standards for de-identification, the inherent risks of re-identification, and the concerns for de-identified data's utility (and/or marketability), it is no surprise that the genetic data being sold is not exactly anonymous. Researchers have demonstrated that individuals can be identified using only a small snippet of their DNA found on a supposedly de-identified public genetic database.[38] Aggregated genetic information has also proven susceptible to re-identification.[39] Moreover, re-identification may also expose close relatives of individuals included in the dataset, or at least traits about those individuals such as diseases, health conditions, or unknown familial relations.[40] One study projects that around 60% of individuals of European descent in the United States could be identified through familial matching in genetic databases.[41] Linda Avey, a cofounder of 23andMe, admitted that "it's a fallacy to think that genomic data can be fully anonymized."[42] A researcher on one of the re-identification studies was more blunt: "I think the bottom line is now every-

---

37 *See* Judit Sándor, *Genetic Testing Between Private and Public Interests: Some Legal and Ethical Reflections*, 6 FRONTIERS PUB. HEALTH 4 (2018), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5797761/ [https://perma.cc/7RMW-9W2X] ("[F]ully anonymized genetic data cannot be compared with the health data of the specific patient and, consequently, the data are not very useful for scientific research.").

38 Melissa Gymrek et al., *Identifying Personal Genomes by Surname Inference*, 339 SCI. 321, 321 (2013); *see* Amy L. McGuire & Richard A. Gibbs, *No Longer De-Identified*, 312 SCIENCE 370, 370 (2006), https://science.sciencemag.org/content/312/5772/370 [https://perma.cc/T9BK-GL5P].

39 *See* Nils Homer et al., *Resolving Individuals Contributing Trace Amounts of DNA to Highly Complex Mixtures Using High-Density SNP Genotyping Microarrays*, 4 PLOS GENETICS 1, 2–6 (2008), https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1000167 [https://perma.cc/KC7N-3MR7] (suggesting that statistics based on certain genetic traits, such as allele frequency or genotype counts, may still reveal the identities of the individuals being studied).

40 Natalie Ram, *DNA by the Entirety*, 115 COLUM. L. REV. 873, 899 (2015).

41 *See* Yaniv Erlich et al., *Identity Inference of Genomic Data Using Long-Range Familial Searches*, 362 SCIENCE 690, 690 (2018), https://science.sciencemag.org/content/362/6415/690/tab-pdf [https://perma.cc/8UUJ-B6RN]; *see also* Michael D. Edge et al., *Linkage Disequilibrium Matches Forensic Genetic Records to Disjoint Genomic Marker Sets*, 114 PROC. NAT'L ACAD. SCI. U.S., 5671, 5671 (2017), https://www.pnas.org/content/114/22/5671 [https://perma.cc/YQR2-NK3L] (demonstrating the feasibility of long-range familial searches on genetic data stored in criminal offender databases).

42 Peter Pitts, *The Privacy Delusions of Genetic Testing*, FORBES (Feb. 15, 2017, 1:26 PM), https://www.forbes.com/sites/realspin/2017/02/15/the-privacy-delusions-of-genetic-testing/#2dc634791bba [https://perma.cc/4RXL-ZRPV].

body is about to be under genetic surveillance one way or another . . . ."[43]

II
USING DE-IDENTIFICATION AND AGGREGATION TO AVOID GENETIC DATA REGULATION

A. Individual Property and Privacy Interests in Genetic Data

The legal debate over properly protecting genetic data, and data overall, is animated by two conceptions of data: one implicating an individual's property rights, the other implicating their privacy rights. Both property and privacy are imperfect approaches because both grant individuals only limited ability to assert control over their data, but the appeal is evident, especially the economic appeal behind granting data the status of property. An individual with property rights over their genetic information can exercise property law's traditional "bundle of rights," including an individual right to regulate the possession, transfer, or destruction of their genetic data.[44] Also appealing is that property rights in one's own genetic data makes one's interest in that data cognizable under the U.S. Constitution's 5th and 14th Amendments, offering a legal remedy to someone whose genetic data is taken or used without permission (at least by the government).[45]

However, treating an individual's medical information as a form of private property can also perversely jeopardize individual control over genetic data. Medical records are treated by many state laws almost as a form of intellectual property in which healthcare providers own the records themselves but not the underlying health data.[46] Rather, patients have a property interest in their health data so they (and their insurers) have a right of access to the records. Some have compared the arrangement to owning a book but not the intellectual content in it.[47] The comparison is limited, however, because U.S. federal law does not explicitly assign ownership of medical records.[48]

---

43 Megan Molteni, *Genome Hackers Show No One's DNA is Anonymous Anymore*, WIRED (Oct. 11, 2018, 2:04 PM), https://www.wired.com/story/genome-hackers-show-no-ones-dna-is-anonymous-anymore/ [https://perma.cc/NG7J-UZQH].

44 Anya E.R. Prince, *Comprehensive Protection of Genetic Information: One Size Privacy or Property Models May Not Fit All*, 79 BROOK. L. REV. 175, 183 (2013).

45 *Id.* at 183–84.

46 *See* Rodwin, *supra* note 3, at 587–88.

47 *Id.* at 588.

48 *See id.* at 588–89.

Rather, a de facto and state-dependent intellectual property framework, largely benefiting healthcare providers and medical researchers, sits in tension with the interests of patients and medical subjects.

Indeed, U.S. courts' use of an intellectual property framework has permitted ownership of genetic information but has generally favored the ownership rights of research institutions and corporations over those of individuals. In 1980, the Supreme Court's landmark ruling in *Diamond v. Chakrabarty* sanctified an intellectual property interest in a genetically modified organism because it was not naturally occurring but rather the product of human ingenuity.[49] In 1990, the California Supreme Court held in *Moore v. Regents of the University of California* that to the extent an individual has a right over their own cells, it is a privacy right, and granting a property right was deemed inappropriate.[50] The cells at issue had been taken from a leukemia patient in the course of treatment but later patented by his physician and UCLA, who sold it to a biotech company for use in research and product development.[51] California's Supreme Court held that since these cells perform a function shared by all humans, they were no more unique to the patient from whom they were derived "than the number of vertebrae in the spine or the chemical formula of hemoglobin."[52] Moreover, the qualities that made this cell line unique were not its connection to the patient, but the result of researchers' work, thus qualifying it as their property under *Chakrabarty* rather than the patient's.[53]

Privileging the labor of researchers over the personal claims of individuals has been the norm, but courts have not been uniformly hostile to individual claims of property in genetic information.[54] As of 2018, five states even recognized an

---

49 Diamond v. Chakrabarty, 447 U.S. 303, 309–10 (1980).

50 Moore v. Regents of the Univ. of Cal., 793 P.2d 479, 489–90, 494 (Cal. 1990), *cert. denied*, 499 U.S. 936 (1991).

51 *Id.* at 480–82.

52 *Id.* at 490.

53 *Id.* at 491–92; *see also* Greenberg v. Miami Children's Hosp. Research Inst., Inc., 264 F.Supp.2d 1064, 1074 (S.D. Fla. 2003) (holding that individuals do not have a property interest in tissue samples voluntarily donated for research).

54 It is telling however, that many of the successful claims for property-like rights (such as the right to control, exclude, or destroy) over genetic material filed since *Chakrabarty* and *Moore* have not been articulated as property claims. *See* Jessica L. Roberts, *Progressive Genetic Ownership*, 93 NOTRE DAME L. REV. 1105, 1125–26 (2018); *see also* Robyn L. Sterling, *Genetic Research Among the Havasupai: A Cautionary Tale*, 13 AM. MED. ASS'N J. ETHICS 113, 115–16 (2011) (describing the Havasupai tribe's suit against Arizona State University over donated blood samples which sought to secure the tribe's right to control and

individual property interest in one's own genetic material.[55] In practice, however, individuals have few property rights over their DNA itself, even if no explicit rule forbids individual genetic ownership, especially once they've turned over DNA or genetic information to someone else. Physicians are at least bound by fiduciary duties to their patients and must disclose enough information for the patient to give informed consent in the release of their biological materials. In contrast, where an individual donates biological tissue or genetic material to researchers, that individual cannot expect the same level of care.[56] Outside of the clinical context, donors of biological material are not owed fiduciary duties.[57] And once the donation of blood, tissue, or DNA has been made, the recipient is relatively free to claim an intellectual property interest in any product derived from the material.[58] In the erratic patchwork of U.S. laws governing ownership in genetic data, individuals have a much more tenuous property right in their own DNA than the organizations who use it for research and commercial products. Unsurprisingly, then, companies holding medical records and genetic information have largely treated patient data as their private asset.

De-identification and aggregation of patient data thus also serves as a central mechanism for extinguishing any individual property claim to the data before selling it to third parties.[59] De-identification and aggregation dilute an individual's claim of ownership over personal data because the data's personal connection to them is made unrecognizable. In its place, the process of de-identifying and aggregating the data can itself constitute the labor necessary for marking the resulting dataset as the property of the data's custodian rather than its source.

---

possess their biological material); Adam Doerr, *Newborn Blood Spot Litigation: 70 Days to Destroy 5+ Million Samples*, PRIVACY REPORT (Feb. 2, 2010), https://theprivacyreport.com/2010/02/02/newborn-blood-spot-litigation-70-days-to-destroy-5-million-samples/ [https://perma.cc/7GUC-AT8Z] (describing Texas parents' lawsuit against the state over infant blood samples collected without their consent, seeking to secure their right to control and destroy their children's biological material).

55 Roberts, *supra* note 54, at 1128.

56 *See* Kristin E. Schleiter, *Donors Retain No Rights to Donated Tissue*, 11 AM. MED. ASS'N J. ETHICS 621, 621 (2009).

57 *See* Jessica L. Roberts, Theories of Genetic Ownership 38, 45–46 (Sept. 9, 2015) (unpublished manuscript) (on file with the Petrie-Flom Center, Harvard University), https://petrieflom.law.harvard.edu/assets/publications/Roberts_Genetic_Ownership_Draft.pdf [https://perma.cc/W9ST-E7HG].

58 *See id.* at 46–47.

59 Rodwin, *supra* note 3, at 588.

Professor Marc Rodwin argued that the deficiencies in an intellectual property model of medical data demonstrate that such data should be considered a public resource, with de-identification and aggregation playing a leading role.[60] To Rodwin, private ownership of data prevents the development of large medical databases, skewing data collection toward solely profitable information and stifling medical advances.[61] At worst, Rodwin worried that "[p]rivate ownership could not ensure a stable source of data," starving medical research of reliable information.[62] Since private ownership provides insufficient or incorrect incentive effects, patient data should be publicly owned and maintained to best realize the data's scientific potential. Though individual patients won't share in the financial value of research derived from their medical data, they will share in the benefits to health and safety from the resulting advances in research.[63] Moreover, patients lack any current legal right to their anonymized data and their data has economic value only because of its utility in the work of others. "[Patients] benefit from what physicians and researchers learn from data from other patients without paying for use of such data. What grounds could they then have to demand compensation for others learning from data routinely collected as part of their medical care?"[64]

Yet, Rodwin's approach exacerbates rather than diminishes the power imbalance in ownership of genetic data. Marking genetic data as a public resource simply codifies individuals' lack of control over their data as soon as it leaves their possession. Private companies and organizations can still reap private profit from the data, but the arrangement is justified by a more efficient utilization of medical data as a resource. Yet no matter whether the data is treated as public or private, de-identification and aggregation are the lynchpin to assuring that individuals' interest in their own data is minimal. By anonymizing and aggregating medical data, those who collect and store data convert it from personal information to a commercial resource. Whatever rights people do have in their genetic data, those rights' primary protection is the technical prowess of companies in keeping the data truly anonymous and unconnected to the individual.

---

60 *See id.* at 589.
61 *Id.* at 600.
62 *Id.*
63 *Id.* at 609.
64 *Id.*

Amassing these collections of "anonymized" data is also essential for undermining an individual's assertion of privacy rights over data. In their landmark article *The Right to Privacy*, Samuel Warren and Louis Brandeis explicitly linked privacy-like rights to the common law concept of intellectual property.[65] In controlling one's intellectual property, one controls the thoughts and feelings communicated to others, regardless of that information's character, value, or specific physical manifestation.[66] Warren and Brandeis, however, ultimately saw privacy as a right inherently independent of property and based in the self, epitomized in the famed "right of the individual to be let alone."[67] "The principle which protects personal writings and all other personal productions, not against theft and physical appropriation, but against publication in any form, is in reality not the principle of private property, but that of an inviolate personality."[68] An individual's privacy interest in specific information is thus fundamentally linked to that individual's personal connection to the information. To the extent that the link can be broken, the information is freed from the claims of the person it is derived from.

In the century since *The Right to Privacy*'s publication, American law has largely separated privacy rights from any roots in property, allowing privacy to encompass both certain rights in criminal investigations and rights central to personal autonomy, such as those related to "marriage, procreation, contraception, family relationships, and child rearing and education."[69] In 1977, the Supreme Court acknowledged in

---

65 Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 198–99 (1890).

66 *Id.* at 198–99, 204; *see also* Jessica Bulman, *Publishing Privacy: Intellectual Property, Self-Expression, and the Victorian Novel*, 26 HASTINGS COMM. & ENT. L.J. 73, 80–81 (2003) (arguing that the European system of *droit moral*, designating a bundle of rights authors have over their creative works, links an individual to their creative works more explicitly than the Anglo-American framework used by Warren and Brandeis—thus basing protection on an individual's personality or expression of self).

67 Warren & Brandeis, *supra* note 65, at 205; *see also* Olmstead v. United States, 277 U.S. 438, 478 (1928) (Brandeis, J., dissenting) (characterizing "the right to be let alone" as "the right most valued by civilized men.").

68 Warren & Brandeis, *supra* note 65, at 205.

69 Paul v. Davis, 424 U.S. 693, 713 (1976); *see also, e.g.*, Lawrence v. Texas, 539 U.S. 558, 578 (2003) (holding that an individual's right to privacy includes consensual same-sex sexual activity without government interference); Roe v. Wade, 410 U.S. 113, 153–54 (1973) (holding that the right of privacy includes women's right to abortion); Eisenstadt v. Baird, 405 U.S. 438, 453 (1972) (holding that the right to privacy's protection of contraception usage extends to unmarried couples); Katz v. United States, 389 U.S. 347, 353 (1967) (holding that the Fourth Amendment protection against unreasonable government searches extends to

Whalen v. Roe that individuals have privacy rights in medical records that comprise two distinct interests—informational privacy, i.e., "the individual interest in avoiding disclosure of personal matters," and decisional privacy, i.e., "the interest in independence in making certain kinds of important decisions."[70] The New York statute at issue in the case required physicians prescribing certain drugs to report patients' names, ages, addresses, drug prescriptions, and dosages to the New York State Department of Health, which would store the information in a computer database.[71] The Court explicitly cited Justice Brandeis' conception of privacy,[72] but ultimately found that the statute did not violate an individual's constitutional privacy interests.[73] They reasoned that a patient's privacy rights were no more compromised by the statute than is already inherent in modern healthcare, where personal medical information is shared with doctors, hospitals, insurers, and public health departments, "even when the disclosure may reflect unfavorably on the character of the patient."[74]

When strongly defended, however, the informational and decisional aspects of privacy can protect an individual's control over their data. Professor Sonia Suter argues that data privacy's concern over data control has led many to conflate privacy and property, but that privacy protects a "fundamentally different" kind of control than property.[75] The control property rights denote is over a commodity, requiring "disaggregation of the object from the property holder" to get protection as property.[76] Property law allows a commodity to be severable from its owner and the property's component parts to be severable from the whole. By comparison, privacy rights strive to keep the object of control, the self, whole, and so the nature of privacy control is inherently bound up with consent and personal choice.[77] She cites the centrality of consent to civil and criminal laws around relationships and sexuality as the law's acknowledgment of "the dignitary interests in controlling access

---

where an individual has a right to privacy, not only a property interest); Griswold v. Connecticut, 381 U.S. 479, 485–86 (1965) (holding that the constitutional right of privacy includes marital privacy regarding the use of contraception).

70 Whalen v. Roe, 429 U.S. 589, 599–600 (1977).

71 *Id.* at 593.

72 *Id.* at 599 n.25.

73 *Id.* at 603–04.

74 *Id.* at 602.

75 Sonia M. Suter, *Disentangling Privacy from Property: Toward a Deeper Understanding of Genetic Privacy*, 72 GEO. WASH. L. REV. 737, 767 (2004).

76 *Id.* at 769.

77 *See id.* at 767–69.

to the self."[78] Accordingly, in the context of medical information, the law has recognized that individuals must give informed consent before being subjected to medical care or research.

The inherently personal nature of privacy rights over data makes them vulnerable to data's anonymization and aggregation because these technical processes attenuate an individual's connection to their data. Part of the problem is the nature of data itself. While theories of privacy may conceptualize the self as a whole, stable, integrated subject of legal protection, data by its essence exists in distributed, severable, transferable form. In translating disparate characteristics of the self into binary code, computers create simulacra of individual life with many of the practical characteristics of a commodity. Protecting privacy in the digital domain then requires coping with the property-like features of data. Proponents of informed consent in medical data collection use it to impose concepts of personal integrity and autonomy on a technology that otherwise might promote data as property.

Yet, informed consent's attention from privacy scholars and policymakers may only be encouraging medical data's treatment as private property. Professor Jorge Contreras argued that proponents of informed consent have succeeded in obtaining property-like rights over genetic data, particularly in using informed consent (or lack thereof) to assert individuals' rights to control access to their genetic data, to destroy it, to determine its allowable usage, to maintain control over it after death, and to transfer ownership in it.[79] Contreras argues that the fixation with using informed consent to assert property-like rights in a research or clinical setting has driven others to treat medical data as property in other fields, including in private sector data-collection firms.[80] Moreover, informed consent can be too easily manipulated or manufactured.[81] To Contreras, a procedure that was supposed to protect privacy in patient data has thus actually led to its commodification. By securing property-like rights to protect privacy, informed consent advocates ensured that genetic data would be treated as property by whomever held it, be they individual or corporation.

---

78 *Id.* at 768.

79 Contreras, *supra* note 4, at 21, 23–24, 26–28; *see also* Barbara A. Koenig, *Have We Asked Too Much of Consent?*, 44 HASTINGS CTR. REP. 33, 33 (2014) (arguing that the focus on consent in contemporary biomedical research has become the modern equivalent of a fetish).

80 Contreras, *supra* note 4, at 28–30.

81 *See id.* at 29–30.

B. Genetic Data's Statutory Protection and its Absence

The result of these conflicting views of data has been our current system of privacy regulation in the United States being a complex patchwork of common law, constitutional law, and federal and state regulation.[82] Most data collected commercially in the United States has little or no direct legal protection; rather, Congress and many state legislatures have singled out certain kinds of data for legal and regulatory protection, medical data in particular.[83] Statutory and regulatory protection for genetic data is contingent on whether the law defines the party collecting and holding the data to be a covered entity and whether the data involved is individually identifiable. De-identification and aggregation of data serves to transform legally sensitive data, i.e., personally identifiable health data, into relatively unregulated anonymous data.[84] Both techniques are thus essential for maintaining genetic data's legal limbo and continuing the genetic data market in current form. Three laws are most relevant, even as their requirements are entirely evaded by private genetic testing companies: The Common Rule, the Health Information Portability and Accountability Act (HIPAA), and the Genetic Information Nondiscrimination Act (GINA).[85]

The Common Rule provides the ethical baseline for all government-funded research, requiring human subjects to give informed consent to research and providing for Institutional Review Boards to monitor research and record-keeping practices.[86] However, it does not govern commercial entities unless they are receiving government funding for research.[87] Moreover, regulators have recently clarified that the Common Rule does not apply to "research involving only coded private information" if that information was "not collected specifically for the currently proposed research project" and the information does not allow for the identity of individuals to be "readily as-

---

82 Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477, 482 (2006).

83 Contreras, *supra* note 4, at 17.

84 *Id.* at 33.

85 Genetic Information Nondiscrimination Act of 2008, Pub. L. No. 110-233, 122 Stat. 811; Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936; *see* 45 C.F.R. § 46 (2019).

86 *See* 45 C.F.R § 46.109; Stanley G. Korenman, *Common Rule*, TEACHING RESPONSIBLE CONDUCT RES. HUMANS, https://ori.hhs.gov/education/products/ucla/chapter2/page04b.htm [https://perma.cc/P4CM-245J] (last visited Dec. 19, 2019).

87 *See* 45 C.F.R. § 46.101(a).

certain[ed]."[88] Genetic testing companies avoid the Common Rule and other similar laws because they neither qualify as a covered entity and because de-identification and aggregation of data pushes their data outside legal protection. 23andMe's policies, for instance, are aligned with this interpretation of the rule, as it currently argues that its data collection and analysis operations are not covered by the Common Rule.[89] Researchers working with 23andMe to use customer data in genetic studies have indicated that their research does not qualify as involving "human subjects."[90] In its own public statements, 23andMe has rather coyly claimed that its customers are more than just "human subjects," but rather "partners" in research.[91] What this position indicates is that to the extent 23andMe asks for consent before retaining, using, or selling customer data, it is doing so "voluntarily" and as a courtesy, not to comply with a legal mandate.[92] 23andMe's approach to the Common Rule is symptomatic of how private genetic testing companies avoid legal obligation: their status as a commercial entity helps them avoid direct regulation while de-identification and aggregation helps them avoid regulation of their data.

HIPAA shares the Common Rule's concern for personally identifiable medical information ("protected health information"), elevating it to the most stringent protection under

---

88 *Coded Private Information or Specimens Use in Research, Guidance*, U.S. DEP'T OF HEALTH & HUM. SERVICES (Oct. 16, 2008), https://www.hhs.gov/ohrp/regulations-and-policy/guidance/research-involving-coded-private-information/index.html [https://perma.cc/D8GY-EX42].

89 *See Protecting People in People Powered Research*, 23ANDME BLOG (July 30, 2014), https://blog.23andme.com/23andme-research/protecting-people-in-people-powered-research/ [https://perma.cc/GBC6-KTZQ] [hereinafter *Protecting People*].

90 Nicholas Eriksson et al., *Web-Based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits*, 6 PLOS GENETICS 16 (2010), https://journals.plos.org/plosgenetics/article/file?id=10.1371/journal.pgen.1000993&type=printable [https://perma.cc/TWH9-6U5L].

91 *See, e.g.*, Linda Avey, *It's Your Data . . . Shouldn't You Have Access To It?*, 23ANDME BLOG (June 22, 2009), https://blog.23andme.com/news/its-your-data-shouldnt-you-have-access-to-it/ [https://perma.cc/39W3-M9NB], ("At 23andMe, we believe it's time for a research revolution, where the people involved—let's no longer call them human subjects—can play a more active role and contribute more directly to studies of most interest to them and their families."); *Protecting People*, *supra* note 89 ("We work hard to make our customers true partners in research rather than 'human subjects.'").

92 *Protecting People*, *supra* note 89 ("Although technically only federally funded research has to meet that standard, 23andMe voluntarily applies it to our own internal research.").

HIPAA's Privacy Rule.[93] The Privacy Rule sets out extensive regulations covering the collection, use, storage, and disclosure of personally identifiable information by "covered entities," including healthcare providers, insurers, and laboratories, and those entities' "business associates."[94] Under HIPAA, health information includes genetic information, construed broadly to include information about any individual's genetic tests, the genetic tests of that individual's family members, or the "manifestation" of a disease or disorder among that individual's family members.[95] To be governed under HIPAA regulations, individually identifiable health information must have been "created or received by a health care provider, health plan . . . employer . . . or health care clearinghouse."[96] If this data 1) relates to the physical or mental health of an individual, their healthcare, or their payment for healthcare, and 2) identifies an individual or provides a "reasonable basis" for an individual's identification, then it is covered under HIPAA's definition.[97]

Yet private DNA testing services can avoid HIPAA's framework entirely because of their reliance on de-identified and aggregated data—even as recent studies show that it is increasingly possible to re-identify data,[98] there have not been enough of them to demonstrate that the data being sold by private genetic companies still provides a "reasonable basis" for identification.[99] The testing companies themselves are not covered by HIPAA, since they do not offer healthcare or insurance services (at least for now). But by de-identifying and aggregating the data, private companies can sell the data to institutional clients that would qualify for HIPAA or under other regulations. Other federal regulations governing research similarly acknowledge patient ownership of, or privacy in, genetic data contingent on whether that data is individually identifiable.[100] De-identification and aggregation may not serve to truly keep individuals anonymous, but they do ensure that genetic data is marketable under existing legal regulation. Data that is unburdened by HIPAA and other regulations can

---

93 *See* Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (1996); 42 U.S.C. § 1320d-9; 45 C.F.R. §§ 160, 164 (2003).

94 *See* 45 C.F.R. § 160.310.

95 *See* 42 U.S.C. §§ 300gg-91, 1320d-9.

96 42 U.S.C. § 1320(d)(6); 45 C.F.R. § 160.103.

97 *Id.*

98 Xu et al., *supra* note 18.

99 El Emam et al., *supra* note 2.

100 Ram, *supra* note 40, at 894–95.

be freely stored, transferred, and sold in ways barred for immediately identifiable data.[101] Since HIPAA's passage in 1996, this loophole has become a "cash cow" for genetic testing companies.[102]

Another statute that de-identification and anonymization help guard against is the Genetic Information Nondiscrimination Act (GINA), which prohibits some types of genetic discrimination in health insurance and employment.[103] GINA was passed out of Congress's concern that genetic data's potential as a basis for discrimination would discourage genetic testing and further genetic research.[104] In addition to its anti-discrimination provisions, it also provided an individual right of access to genetic data stored by others, matching some of HIPAA's provisions.[105] GINA acknowledged this right of access to genetic data as a "genomic civil right," but the right is selectively applicable and enforced.[106] Since GINA tied its access rule to HIPAA, the right only applied to data stored by HIPAA-covered entities.[107] Moreover, other regulators like the U.S. Food and Drug Administration and the Centers for Medicare and Medicaid Services (which regulates clinical laboratories) have at times argued that individual access to genetic data could increase potential for research data's misuse, leading many organizations covered under the GINA-HIPAA access rule to deny individuals access.[108]

So, as with both the Common Rule and HIPAA, private testing companies and their data are largely exempt from GINA's provisions. The perverse result of these laws has been to carve out a market for genetic data, albeit in de-identified and aggregated form. De-identified genetic data can be released without individual consent under the Common Rule and the HIPAA Privacy Rule, and then once that data reaches a non-

---

101 *See* Pitts, *supra* note 42.

102 *Id.*

103 Genetic Information Nondiscrimination Act of 2008, Pub. L. No. 110-233, 122 Stat. 881, 882 (codified as amended in scattered sections of 26, 29, and 42 U.S.C.).

104 *Id.* § 2.

105 *See* 42 C.F.R. § 164.524 (2019).

106 Barbara J. Evans, *HIPAA's Individual Right of Access to Genomic Data: Reconciling Safety and Civil Rights*, 102 AM. J. HUM. GENETICS 5, 5–6 (2018), https://www.cell.com/action/showPdf?pii=S0002-9297%2817%2930493-7 [https://perma.cc/DD5P-MZ8U].

107 *See* Cell Press, *Accessing Your Own Genomic Data is a Civil Right But Requires Strategies to Manage Safety*, SCIENCEDAILY (Jan. 4, 2018), https://www.sciencedaily.com/releases/2018/01/180104131621.htm [https://perma.cc/K5QW-HPXV].

108 Evans, *supra* note 106, at 6–7.

HIPAA covered organization, there's no statutory protection against that data being re-identified.[109] To the extent that GINA limited this market, it was only in the sale of genetic data to health insurers and employers, but potential customers like pharmaceutical companies and providers of "life, long-term care, or disability insurance plans" were left exempted.[110]

III
OPERATING THE MARKET FOR GENETIC DATA

A. Terms of Service and Privacy Policies

Private genetic testing companies have made individual privacy and control over data important pillars of their pitch to potential customers, but even their advertising contains qualified language sharply limiting customer protections. On its website, 23andMe promises to put "you in control," offering "[m]eaningful choice" and "[p]rivacy by design" to customers.[111] But these statements have an odd echo of legal regulations, including a promise not to "sell, lease or rent your individual-level information" without consent.[112] De-identified, aggregated information that cannot "reasonably" be traced back to an individual, however, will be shared with third-parties regardless.[113] Similarly, Ancestry promises to be "good stewards of your personal information," ensuring "you have CONTROL [sic] of your own data."[114] But, Ancestry also reserves the right to use customer data for nearly any purpose internally and won't share data with third-parties without consent "other than as described in [its] Privacy Statement."[115]

And it is in these privacy statements and terms of service that the status of this data becomes clear. Both companies

---

109 *Id.* at 6.

110 *See* Linnea Laestadius, *Transparency and Direct-to-Consumer Genetic Testing Companies*, THE PETRIE-FLOM CTR: BILL OF HEALTH BLOG, (Nov. 22, 2016), http://blog.petrieflom.law.harvard.edu/2016/11/22/transparency-and-direct-to-consumer-genetic-testing-companies/ [https://perma.cc/UY24-V9MV].

111 *Privacy is in our DNA*, 23ANDME, https://www.23andme.com/privacy/ [https://perma.cc/R5LJ-GJP4] (last visited Feb. 12, 2020).

112 *Id.*

113 The purposes specified for sharing data with third parties are "to perform business development, initiate research, send you marketing emails and improve our services." *Id.*

114 *Privacy*, ANCESTRY https://www.ancestry.com/cs/privacyphilosophy [https://perma.cc/T4BJ-MW4H] (last visited May 13, 2019) (emphasis omitted).

115 Ancestry's list of potential uses seems exhaustive, including using data for: "genealogical or genomic research projects," verifying customer identity, advertising, marketing, product development, detecting fraud or criminal activity, "internal business purposes," and other "research initiatives." *Id.*

offer slightly greater protection to personal data than de-identified data, but both are essentially property of the companies collecting it. Ancestry defines personal information as that which can directly identify an individual or may "reasonably" be linked back to them.[116] 23andMe similarly defines it as information which can be used to identify an individual "alone or in combination with other information," while aggregate information is that which does not permit a specific individual to be "reasonably" identified.[117] Both companies state that they will only share genetic information either with consent or in de-identified/aggregated form.[118] However, if an Ancestry customer does consent to sharing data or does so with other users, the customer largely loses the ability to delete the data.[119] 23andMe has similar restrictions on deletion, though it permits customers to request that their personal information not be used in future research projects.[120] By in large, customers consent to their data's usage anyway. For example, 80% of 23andMe's users have consented to their data's usage by third parties.[121] Moreover, both companies' terms of service require that customers waive any property right to research or commercial products developed as a result of their data's usage.[122] It is against this backdrop that 23andMe warns unironically: "[g]enetic [i]nformation you share with others could be used against your interests."[123]

Possibly the biggest loophole in these privacy statements and the terms and conditions is that companies reserve the right to change the material terms of these agreements at any time. Customers may hand over their genetic data, believing that the companies are bound to abide by these agreements, only to find that all of the provisions for customer control over

116 *Ancestry Privacy Statement*, *supra* note 25, § 1.

117 *Terms of Service*, 23ANDME § 1(e)–(f), https://www.23andme.com/about/tos/ [https://perma.cc/YC8Z-NKSZ] (last updated Sept. 30, 2019) [hereinafter *23andMe Terms of Service*].

118 *Ancestry Privacy Statement*, *supra* note 25, § 3; *23andMe Terms of Service*, *supra* note 117, § 8.

119 *Ancestry Privacy Statement*, *supra* note 25, §§ 8–10.

120 *See 23andMe Terms of Service*, *supra* note 117, §§ 8, 13.

121 *23andMe for Scientists*, 23ANDME, https://research.23andme.com/ [https://perma.cc/4LZE-2QLC] (last visited Apr. 8, 2020).

122 *Ancestry Terms and Conditions*, ANCESTRY § 3, https://www.ancestry.com/cs/legal/termsandconditions [https://perma.cc/45E8-XBJU] (last updated July 25, 2019); *23andMe Terms of Service*, *supra* note 117, §§ 6(k), 13.

123 *23andMe Terms of Service*, *supra* note 117, § 5.

data and customer privacy can be changed unilaterally.[124] Representatives of both Ancestry and 23andMe confirmed to me that customers who disagree with material changes to their terms and conditions or privacy policy cannot object to them or exempt personal data entirely from the changes.[125] However, both companies do allow customers to cancel their accounts and delete at least some of the identifying information in the companies' possession.[126] Given this glaring loophole, it is not clear that customers truly have any guaranteed control or privacy over their genetic data. As Professor Elizabeth Joh remarked, the "first rule of data" is that "once you hand it over, you lose control of it. You have no idea how the terms of service will change for your 'recreational' DNA sample."[127]

B. Corporate Usage of Individual Data

Given the wide discretion private testing companies have reserved for themselves over their customers' personal and genetic data, they have amassed massive databases on their customers including "genomic sequence, name, self-disclosed family history, health conditions, race, ethnicity, sexual orientation, age, social networks, place of employment, as well as a record of every website that [a customer] clicks on, photos, and real-time tracking of [a customer's] geographic location."[128] Much of this data's value resides in its implications for individuals, even when what's being sold or shared with others is de-identified and aggregated. As described above, this data is vulnerable to being re-identified, but even without re-identification it is highly valuable to a wide range of companies that rely on predictive modeling or risk assessment of certain groups or individuals. An individual who gives DNA for an ancestry test

---

124 *Ancestry Privacy Statement*, *supra* note 25, § 13; *Ancestry Terms and Conditions*, *supra* note 122, § 7; *23andMe Privacy Statement*, *supra* note 19, § 13; *23andMe Terms of Service*, *supra* note 117, § 26.

125 The Ancestry representative also reassured me that any changes to these agreements would be "in the best interest of the customer," but admitted that he didn't "think anyone reads those things."

126 23andMe's representative stated that users could cancel their accounts if they objected to material changes. Ancestry's website notes that customers who disagree with changes to the privacy agreement or terms of service may cancel their subscriptions "if applicable." *Ancestry Terms and Conditions*, *supra* note 122, § 7.

127 Elizabeth Joh (@elizabeth_joh) TWITTER (Feb. 4, 2019, 6:58 AM), https://twitter.com/elizabeth_joh/status/1092437273134563329 [https://perma.cc/8S88-V449].

128 Katherine Drabiak, *Caveat Emptor: How the Intersection of Big Data and Consumer Genomics Exponentially Increases Informational Privacy Risks*, 27 HEALTH MATRIX 143, 159 (2017).

and consents to its use in research may end up finding it used in denying her insurance or targeting her for advertisements. As Sara Chodosh of Popular Science remarked, "[t]he product isn't really a kit, then—the product is you."[129]

And for private genetic data, business is booming. There are now over fifty companies offering direct-to-consumer genetic testing and more than twenty six million people's DNA now sits in private commercial databases.[130] 23andMe and Ancestry are the two "superpowers" of the industry, having collected DNA on more than 23 million people combined (or around 88 percent of the commercially-collected DNA on the market).[131] The size of these private genetic databases is essential to their value, also limiting the success of new genetic testing startups.[132] Especially given the complex relationships present in genetics, research using genetic data requires "massive datasets" to be useful.[133] Unsurprisingly then, private genetics' two titans have been quick to monetize all of the DNA they have collected. Both have been selling data to research institutions and private for-profit companies, ranging from P&G Beauty to Pepto-Bismol.[134] PharmaExec.com declared the new availability of genetic data to be a "gold rush."[135] As another researcher remarked, "[p]eople need to realize that

---

129 Sara Chodosh, *Getting Your Genetic Disease Risks from 23andMe is Probably a Terrible Idea*, POPULAR SCIENCE (Apr. 7, 2017), https://www.popsci.com/23andme-is-probably-terrible-idea [https://perma.cc/8HZR-5KXP].

130 Antonio Regalado, *More Than 26 Million People Have Taken an At-Home Ancestry Test*, MIT TECH. REV. (Feb. 11, 2019), https://www.technologyreview.com/s/612880/more-than-26-million-people-have-taken-an-at-home-ancestry-test/ [https://perma.cc/AN3W-66AJ]; Kim Hart, *Genetic Testing Firms Share Your DNA Data More Than You Think*, AXIOS (Feb. 25, 2019), https://www.axios.com/dna-test-results-privacy-genetic-data-sharing-4687b1a0-f527-425c-ac51-b5288b0c0293.html [https://perma.cc/8AUA-3M9Z].

131 Regalado, *supra* note 130 (noting that Ancestry currently has genetic data of at least fourteen million people while 23andMe has data of at least nine million).

132 Ben Hirschler, *Cashing in on DNA: Race on to Unlock Value in Genetic Data*, REUTERS (Aug. 3, 2018, 4:01 AM), https://www.reuters.com/article/us-health-dna/cashing-in-on-dna-race-on-to-unlock-value-in-genetic-data-idUSKBN1KO0XC [https://perma.cc/W6WL-PE2S].

133 *Id.*

134 Nicole Martin, *How DNA Companies Like Ancestry and 23andMe Are Using Your Genetic Data*, FORBES (Dec. 5, 2018, 2:49PM), https://www.forbes.com/sites/nicolemartin1/2018/12/05/how-dna-companies-like-ancestry-and-23andme-are-using-your-genetic-data/#46691f306189 [https://perma.cc/YFS7-SPD5].

135 Bill Fox, *The Genetic Data Gold Rush: Balancing Privacy and Health Outcomes*, PHARMACEUTICAL EXECUTIVE (Oct. 18, 2018), http://www.pharmexec.com/genetic-data-gold-rush-balancing-privacy-and-health-outcomes [https://perma.cc/M6X5-5KFE].

they are actually paying for companies to monetize their most personal information and they are getting nothing for it."[136]

23andMe in particular has been very open about its ambitions for its customers' data. As a board member stated, "[t]he long game here is not to make money selling [DNA testing] kits, although the kits are essential to get the base level data . . . Once you have the data, [23andMe] does actually become the Google of personalized health care."[137] Data collection has been the whole point since the company's founding, especially in marketing its data to pharmaceutical companies.[138] One researcher estimated that 23andMe sold individual genotypes for an average price of $130 per person,[139] though the company does not make its sales figures public. This estimate didn't include 23andMe's most recent, most valuable data deal yet—$300 million from pharmaceutical giant GlaxoSmithKline in exchange for access to customers' genetic data for the development of new drugs.[140]

Ancestry and other companies may show slightly more concern with the image of selling customers' genetic data, but this concern has not prevented new business partnerships. Ancestry had a deal with U.S. biotech company Calico in 2015 to share customer data for an undisclosed sum which has since ended.[141] Ancestry does however maintain data sharing arrangements with academic institutions. Other private deals for genetic data have caused worry, especially when a genetic testing company is acquired. For instance, drug maker Amgen

---

136 Hirschler, *supra* note 132.

137 Charles Seife, *23andMe Is Terrifying, but Not for the Reasons the FDA Thinks*, SCI. AM. (Nov. 27, 2013), https://www.scientificamerican.com/article/23andme-is-terrifying-but-not-for-the-reasons-the-fda-thinks/ [https://perma.cc/K6PA-A9K2].

138 *See* Thomas Goetz, *23andMe Will Decode Your DNA for $1,000. Welcome to the Age of Genomics*, WIRED (Nov. 17, 2007, 12:00 PM), https://www.wired.com/2007/11/ff-genomics/ [https://perma.cc/TU85-D2PM] (describing 23andMe's founder discussing the potential for consumer genomics data in pharmaceutical research at the so-called "Billionaires' Dinner").

139 Hirschler, *supra* note 132. By comparison, an individual's personal data—such as age, gender, and location—can be sold for only a fraction of a cent or as much as 26 cents for specialized data such as health conditions and prescriptions. Emily Steel et al., *How Much is Your Personal Data Worth?*, FIN. TIMES (June 12, 2013), https://ig.ft.com/how-much-is-your-personal-data-worth/#axzz2z2agBB6R [https://perma.cc/4UDT-VJ8K].

140 Lydia Ramsey, *Why Pharma Giant GSK Just Made a $300 Million Bet on 23andMe's Approach to Finding New Medicines*, BUS. INSIDER (July 25, 2018, 9:21AM), https://www.businessinsider.com/why-gsk-invested-300-million-in-23andme-genetic-drug-discovery-collaboration-2018-7 [https://perma.cc/566M-HXBR].

141 Hirschler, *supra* note 132.

bought DeCODE Genetics in 2012,[142] acquiring the DNA profiles for nearly half of Iceland's adult population in the process.[143] GlaxoSmithKline bought Human Genome Sciences the same year for nearly three billion dollars.[144] And this possibility has crossed the minds of executives at Ancestry and 23andMe: both warn in their terms of service that in the case either company is acquired, their genetic database will likely be an important asset in the transaction.[145]

Insurers have also been deeply interested in using genetic data, and many worry (or hope) that genetic data will be central to modeling customers' risks for products like life insurance. Many insurance companies even seem to feel entitled to their customers' genetic data. The American Council of Life Insurers (ACLI) believes that no state law bars insurers from using existing genetic test results, and that if a customer refuses to provide that information, the life insurer has a "right to void a policy."[146] "Both the applicant and the insurer must 'put their cards on the table,' . . . Life insurers rely on the honesty of applicants," ACLI stated.[147] Critics worry that insurance companies will "seek out people who are genetically pure, creating a ghetto of the uninsured."[148] ACLI for its part argued that individuals withholding their genetic information are trying to "game the system," adding to costs for everyone.[149]

And these are only the well-established commercial uses for genetic data. The unpredictable or non-commercial uses can be just as worrisome, especially as genetic research speeds along. Researchers have investigated the link between an indi-

---

142 Meg Tirrell, *Iceland's Genetic Goldmine, and the Man Behind It*, CNBC (Apr. 6, 2017, 3:31 PM), https://www.cnbc.com/2017/04/06/icelands-genetic-goldmine-and-the-man-behind-it.html [https://perma.cc/8FGY-QLNB].

143 *Science*, DECODE GENETICS https://decode.com/research/ [https://perma.cc/39CM-6SFY] (last visited Feb. 3, 2020).

144 Michael J. de la Merced, *Glaxo to Buy Human Genome Sciences for $3 Billion*, N.Y. TIMES (July 15, 2012, 6:07 PM), https://dealbook.nytimes.com/2012/07/15/glaxosmithkline-in-talks-to-buy-human-genome/ [https://perma.cc/R7E5-683M].

145 *Ancestry Privacy Statement*, *supra* note 25, § 7 (providing that in the event of an acquisition or bankruptcy, "[Ancestry] will share your Personal Information with the acquiring or receiving entity"); *23andMe Privacy Statement*, *supra* note 19, § 4(f), (providing that in the event of a merger, acquisition, or sale of corporate assets "your Personal Information will likely be among the assets transferred.").

146 Kelly Song, *4 Risks Consumers Need to Know About DNA Testing Kit Results and Buying Life Insurance*, CNBC (Aug. 4, 2018, 10:00 AM), https://www.cnbc.com/2018/08/04/4—risks-consumer-face-with-dna-testing-and-buying-life-insurance.html [https://perma.cc/37UR-UM9V].

147 *Id.*

148 *Id.*

149 *Id.*

vidual's genetics and everything from their financial acumen[150] to their driving ability,[151] perhaps proving of interest to banks considering a loan application or a car insurer considering a new policy. Moreover, the law enforcement implications of widely available genetic data are troubling,[152] as the recent identification of the Golden State Killer through a commercial genealogy website revealed.[153] In short, the market for genetic data looks poised to expand exponentially as the pool of available data grows and industries find new applications for the data. For individuals to retain any control over their DNA or any semblance of genetic privacy, private companies' legal obligations must grow along with their revenue streams.

---

150 *See* Megan Molteni, *Researchers Want to Link Your Genes and Income—Should They?*, WIRED (Apr. 12, 2019, 7:00AM), https://www.wired.com/story/researchers-want-to-link-your-genes-and-incomeshould-they/ [https://perma.cc/2MB2-SZ3P].

151 *Bad Driver? Blame Your Genes*, REUTERS (Oct. 29, 2009, 1:35AM), https://www.reuters.com/article/us-genes-driving/bad-driver-blame-your-genes-idUSTRE59S0M720091029 [https://perma.cc/T8PH-KPCV].

152 Law enforcement use of genetic data raises significantly different concerns for individual liberty and privacy, and so is largely beyond the scope of this paper. Suffice to say, all genetic testing companies indicate that they will at least comply with valid court orders to turn over the genetic data of their customers. Worryingly (to the author at least) is that public or commercial access to genetic data may allow law enforcement to warrantlessly collect individuals' genetic data. *See, e.g.*, Kristen V. Brown & Bloomberg, *A Major DNA-Testing Company Is Sharing Some of Its Data With the FBI. Here's Where It Draws the Line*, FORTUNE (Feb. 1, 2019, 7:47 PM), http://fortune.com/2019/02/01/genetic-testing-consumer-dna-familytreedna-fbi/ [https://perma.cc/FX7Y-ASXN] (describing FamilyTreeDNA's cooperation with the FBI in providing them customer genetic data). The potential for abuse or misuse of genetic data is underscored by the difficulty in interpreting whether a DNA sample is truly a "match" or "partial-match" with DNA available in genetic information databases. *See* Erin E. Murphy, *The Dark Side of DNA Databases*, ATLANTIC (Oct. 8, 2015), https://www.theatlantic.com/science/archive/2015/10/the-dark-side-of-dna-databases/408709/ [https://perma.cc/56FX-F8ED].

153 Sarah Zhang, *How a Genealogy Website Led to the Alleged Golden State Killer*, ATLANTIC (Apr. 28, 2018), https://www.theatlantic.com/science/archive/2018/04/golden-state-killer-east-area-rapist-dna-genealogy/559070/ [https://perma.cc/68Z8-N4T2]; *see also* Michael Balsamo & Jonathan J. Cooper, *Serial Killer Search Led to Wrong Man in 2017*, ASSOCIATED PRESS (Apr. 27, 2018), https://apnews.com/de2a1166d5664125858cb7b5eed209a5/Use-of-DNA-in-serial-killer-probe-sparks-privacy-concerns [https://perma.cc/EUX7-CB7C] (reporting that police investigating the Golden State Killer had used information from genetic testing companies that originally led to the wrong person as a suspect).

IV
THE NEED FOR TREATING PRIVATE GENETIC TESTING COMPANIES AS INFORMATION FIDUCIARIES

A. A Theory of Information Fiduciaries

Jack Balkin proposed the idea of "information fiduciaries" in 2014 as a potential way to address the peculiar sensitivity of personal data and the growing power of online service providers and cloud companies.[154] Generally, fiduciaries are professionals who owe duties of trustworthiness and loyalty to their clients.[155] Certain professions like doctors, lawyers, and accountants are subject to these duties because by nature these professions depend on a relationship of trust with clients.[156] Moreover, these professions involve a significant imbalance in knowledge and expertise between the professional and client. Clients put their trust or confidence in a fiduciary to look after their interests, and the fiduciary has a duty not to betray them.[157] For Balkin, the sharing of sensitive data was the common thread of professions marked legally as fiduciaries: "at their core, fiduciary relationships are relationships of trust and confidence that involve the use and exchange of information."[158]

Balkin argued the same was true for those whose business was based in collection of personal information—"certain kinds of information constitute matters of private concern not because of their *content*, but because of the *social relationships* that produce them."[159] Accordingly, professionals that deal in personal data have an implicit duty not to reveal, use, or sell that data if doing so is against the individual's interest or will pose a conflict of interest between the company and the indi-

---

154 *See* Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 U.C. DAVIS L. REV. 1183, 1186 (2016) [hereinafter *First Amendment*]; Jack Balkin, *Information Fiduciaries in the Digital Age*, BALKINIZATION (Mar. 5, 2014), https://balkin.blogspot.com/2014/03/information-fiduciaries-in-digital-age.html [https://perma.cc/278Q-X57J] [hereinafter *Digital Age*]; *see also* DANIEL J. SOLOVE, THE DIGITAL PERSON: TECHNOLOGY AND PRIVACY IN THE INFORMATION AGE 103–04 (2004) (similarly suggesting, a decade earlier, looking to the law of fiduciaries as a guide to policies regulating data brokers and other aggregators of sensitive personal information).

155 *First Amendment*, *supra* note 154, at 1207.

156 *See id.* at 1205–08.

157 *See First Amendment*, *supra* note 154, at 1186; *Digital Age*, *supra* note 154; *see also* SOLOVE, *supra* note 154, at 103–04 (similarly suggesting, a decade earlier, looking to the law of fiduciaries as a guide to policies regulating data brokers and other aggregators of sensitive personal information).

158 *First Amendment*, *supra* note 155, at 1186–1207.

159 *Id.* at 1205.

vidual.[160] Individuals like doctors and lawyers already fall into this category because they must use information learned from their relationship with clients in the client's best interests.[161] In Balkin's view, other professions dealing in personal information should be held to the same standard, most essentially in duties of care and of loyalty to their customers.

Balkin qualified that new fiduciaries of the digital age may not need to abide by identical standards as the professions established as fiduciaries in common law, nor need they stop monetizing data completely.[162] The reasons to identify a new class of individuals as fiduciaries are rooted in similar concerns as what inspired common law to mark doctors and lawyers as fiduciaries: an acute imbalance in expertise, relative customer dependence, and the necessity of personal information's disclosure.[163] Moreover, many companies offer digital services in exchange for personal data on the explicit premise that they are, in fact, trustworthy. "By presenting themselves as trustworthy collectors and keepers of our individual data, and by emphasizing that, for reasons of security and competitiveness, they cannot be fully transparent, digital organizations induce relations of trust from us, so that we will continue to use their services."[164] If this is their business model, what reason do these companies have to oppose being held to obligations they already promise to uphold? Balkin argues that the standard imposed on these companies should simply reflect their duty not to act as "con men"[165]—"[w]hat information fiduciaries may not do is use the data in unexpected ways to the disadvantage of people who use their services or in ways that violate some other important social norm."[166]

Professors Neil Richards and Woodrow Hartzog offer a vision of what fiduciary duties adapted to digital-age professions might look like.[167] Like Balkin, they argue that both by necessity and by promise, companies dealing in sensitive personal information have based their relationship to customers on trust, and it is only though trusted information relationships that a digital society can be sustained. To maintain trust, fidu-

160 *See id.* at 1206–07.
161 *Id.* at 1207–08.
162 *See id.* at 1221, 1226–27.
163 *See id.* at 1222–23.
164 *Id.* at 1223.
165 *Id.* at 1224.
166 *Id.* at 1227.
167 *See* Neil Richards & Woodrow Hartzog, *Taking Trust Seriously in Privacy Law*, 19 STAN. TECH. L. REV. 431, 433–34 (2016).

ciary duties or fiduciary-like duties should be imposed on these companies—sparingly where trust between customer and company is minimal but imposed stringently where customer trust (or potential for personal information exposure) is high.[168]

Much of their argument focuses on duties rooted in fiduciary duties but of particular concern to digital relationships, such as duties of discretion in data disclosure, transparency, stewardship, and loyalty. Given the widespread sharing of data and its utility, a company need not keep all personal data undisclosed but must exercise sound discretion in what data it releases, to whom, and under what circumstances.[169] Wrongful disclosure, not any disclosure, is the focus of the duty's prohibition.[170] Relatedly, companies dealing in private information must meet a minimum amount of transparency, related to the older concept of a fiduciary's duty of candor.[171] This duty is essential for an individual customer to be on notice of the company's practices, to properly tailor their disclosures to the company, and to have enough information to enforce and monitor company compliance with their other duties.[172] The common practice of burying relevant, material information in terms of service or privacy agreements is not sufficient—information fiduciaries have "an affirmative obligation of honesty to correct misinterpretations and to actively dispel notions of mistaken trust."[173]

Moreover, these companies must understand themselves to be data stewards, not merely guardians of a database—informational fiduciary duties follow the data itself.[174] This duty, along with the duty of discretion in disclosure, may require companies disclosing data to third parties to do so under agreements binding the third party to similar levels of care and loyalty to the individual.[175] Finally, private companies dealing in

168 *Id.* at 458.

169 *Id.* at 459 (comparing the characterization of confidentiality as non-disclosure to "characterizing safe sexual practices solely in terms of abstinence—it's effective, but risks overkill and is often too costly.").

170 *See id.* at 461; *see also* Jessica Litman, *Information Privacy/Information Property*, 52 STAN. L. REV. 1283, 1308–09 (2000) (discussing the expectation of confidentiality arising from many merchant-customer relationships that involve the exchange of sensitive personal information); Neil M. Richards & Daniel J. Solove, *Privacy's Other Path: Recovering the Law of Confidentiality*, 96 GEO. L.J. 123, 156–58 (2007) (tracing the duty of confidentiality to common law tort claims).

171 Richards & Hartzog, *supra* note 167, at 462.

172 *See id.* at 462–63.

173 *Id.* at 462.

174 *See id.* at 466.

175 *See id.* at 466–68.

personal information must have a duty of loyalty to their customers. This duty in the digital context may not bar all self-dealing or corporate benefit from the data, but it does mean that companies cannot use the data to undermine trust, such as by using the data in unexpected or directly adverse ways.[176] In other words, the kind of trust between company and customer will define the scope of the duty.

B. Applying Information Fiduciary Duties to Genetic Testing Companies

While academics have argued for the information fiduciary model to apply to those who handle personal data generally, I argue that the model is particularly well-suited to private genetic testing companies, especially to close (or at least narrow) the loopholes for de-identified and aggregated data. Much like the loopholes that created the genetic data market, fiduciary duties are applicable to these companies primarily because of the kind of relationship they have to their customers and because of the kind of data they collect. Holding these companies to their promises of customer privacy and autonomy is not just appropriate and ethically incumbent, it is also practically overdue. The market for individual genetic data is already booming, customer DNA profiles are already being widely shared without consent, and private genetic testing companies are already raking in proceeds.[177] Moreover, the putative privacy protection provided by de-identification and aggregation offer little-to-no guarantee against data being traced back to the individual. Soon, individuals may find their DNA or those of their relatives being commonplace aspects of government surveillance, identity theft, or private sector exploitation by insurers, banks, advertisers, or employers. Meanwhile, private genetic testing companies are quickly insulating themselves from any consequences, as evidenced from user agreements that grant corporate ownership of genetic data and rob customers of legal recourse.[178] These are all the hallmarks of a predatory industry; one whose business model is to make a private profit but pass on the lion share of risks to customers. As genetic databases grow and re-identification techniques become more sophisticated, the risks of the genetic testing industry will become ubiquitous.

---

176 *See id.* at 470.

177 *See supra* notes 130–131 and accompanying text.

178 *See supra* subpart III.A.

These circumstances are fertile ground for new legal regulation to spring up. The relationship between Ancestry or 23andMe and its customers bears all the signs of a fiduciary relationship. These companies transact with their customers from a position of advanced expertise in two arenas: genetics and data collection. Customers agree to user agreements and privacy policies with few ways to assess the quality of service being offered, the consequences of signing up for the service, and nearly no ability to monitor the company to ensure it is upholding what few duties it agreed to abide by.[179] The use of de-identification and aggregation makes this problem even more acute, because users will have little awareness that their information has even been shared to third-parties or know with whom their genetic data has ended up. Given company practices of selling de-identified and aggregated data with or without consent, many customers may not find out their DNA has been released until someone with access to their data decides to re-identify them.[180] Moreover, the information at issue is intrinsically sensitive, representing an individual's biological blueprint. Like a defendant approaching her lawyer or a cancer patient approaching his doctor, customers come to Ancestry or 23andMe in a position of relative ignorance and vulnerability. So far, that imbalance has been these companies' business model.

That imbalance can be addressed, however, by imposing on these companies the duties of a digital-age information fiduciary; namely duties of sound discretion, transparency, stewardship, and loyalty. Imposing these duties in truth just gives legal force to the private genetic companies' advertising promises of customer control and privacy. For an individual to truly be in control of her genetic data, she must know how it is being used and trust that the entity holding her data won't go behind her back to sell it. For 23andMe or Ancestry, a duty of discretion will not prevent them from selling their customers' data, either to research institutions or elsewhere,[181] but it does mean that they cannot rely on de-identification, aggregation, and arcane provisions in their terms of service to show "respect" for customer privacy and choice. Since genetic data is inherently identifiable, compliance with a duty of discretion will generally require them to form agreements with research partners prohibiting genetic data they receive from being shared,

179 *See supra* notes 170–171 and accompanying text.
180 *See supra* notes 8–9 and accompanying text.
181 *See supra* notes 173–174 and accompanying text.

re-identified, or used in a manner to the detriment of customers.

These companies' alternative to more comprehensive regulation of their business partners is covered by a duty to be transparent. If companies want to sell sensitive customer data and have no ability to guarantee the safety of that data, their bare minimum duty to a customer is tell them the risks inherent in sharing the data. Informed consent is a bedrock principle of medical research because it is one of the few ways to procedurally protect individual autonomy.[182] A duty of transparency can make informed consent uniformly applicable to genetic data, no matter the techniques used to alter the data before it is sold. Such a duty can also assure that customers are truly informed, by placing an affirmative obligation on companies to make sure that their customers understand the risks of a given decision. Indeed, when the sharing of personal data has inherently unpredictable consequences, complete transparency is the only way to respect customer autonomy.

Moreover, this kind of concern for where the genetic data ends up is mandated by a duty of stewardship over the data itself. By agreeing to store consumer genetic data, private genetic companies are implicitly agreeing that they are capable of safeguarding it. And given the advertising of Ancestry and 23andMe, this agreement is hardly implicit.[183] Both companies sell their services with the explicit promise of privacy—"privacy by design," and the like. Customers should not be expected to guess that the realities of genetic data and contractual fine print mean that their data is not necessarily private and that the company many alter its promise at will. A legally enforceable duty of stewardship forces these companies to abide by a promise of privacy not only at the point of collection, but also wherever the data is sent.

All of these duties are interconnected in their purpose and animated by an overarching duty of loyalty. Ancestry and 23andMe sell themselves as committed to their customers' interests, even if their data operations may belie the claim. Yet a duty of loyalty does not preclude these companies' involvement in a genetic data market, it merely limits it according to the kind of trust between company and customer. 23andMe's 80% agreement rate to using genetic data for research purposes may demonstrate that there really is desire and expectation

182 *See supra* notes 79–81 and accompanying text.

183 *See supra* subpart III.A.

among customers that DNA data will be used for research.[184] But many of those agreeing to research may not expect it to include pharmaceutical research in the service of private profit. Fewer still may realize that their data used in research may end up with their insurers, employers, or law enforcement. The duty of loyalty, in combination with duties of discretion, transparency, and stewardship, require that a company's business practices do not advance the company's interest to the detriment of their customers'.

The result of imposing these duties will doubtless involve more time, effort, and expense on the part of private genetic companies. And of course, imposing fiduciary duties exposes them to legal liability if they breach their duties. These costs are however inherent to the imposition of any duty, and really are what define them as duties to begin with. To argue that these duties are not financially feasible or may excessively impede research risks arguing too much. All of these duties rest on promises that private genetic companies already make to customers in marketing their services. If companies cannot fulfill promises of consumer privacy and control without becoming insolvent, customers need to be notified of that more effectively than through an online contract's fine print. Indeed, to the extent that client choice or privacy in genetic data inhibits research, a much broader discussion is needed about our valuation of individual autonomy compared to scientific progress. That is a conversation that needs to be held by members of the public, not the heads of private genetics companies or research institutions.

Imposing fiduciary duties on these companies is also appropriate because of the problems it does not solve. Most fundamentally, it does not answer the question of whether property or privacy is more implicated by genetic data. Fiduciary duties simply require that whatever the division of rights over the data, a company given genetic data may not use it to the disadvantage of the individual it is derived from. Indeed, these duties offer relatively wide discretion in the actual practices individual companies or jurisdictions adopt. They do not even necessarily force genetic testing companies to substantially alter their business model. These companies just have to abide by the promises they've already been making and to be frank about the potential risks of purchasing their services.

184 *See supra* note 121 and accompanying text.

Imposing fiduciary duties on 23andMe, Ancestry, and their ilk is also essential for narrowing the loophole exploited by de-identification and aggregation. These techniques for altering data have become indispensable because of the economic and legal framework that's grown around them. Recent research has made it increasingly clear that even as de-identification and aggregation give companies legal protection, they offer diminishing privacy protection for individuals. Regardless of HIPAA and other statute's exemption for de-identified and aggregated data, fiduciary duties would require genetic testing companies to disclose only the information they can assure protection of, to be clear about the risks involved, and to seek consent before disclosing genetic data to third parties. For better or worse, the genetic data market is here to stay. Companies are already making millions by selling their customers' genetic profiles. The least we have a right to demand is that they not betray our confidence in the process.