

DISCRIMINATION AND PRIVACY CONCERNS AT THE INTERSECTION OF HEALTHCARE AND BIG DATA

James Koenderman[†]

TABLE OF CONTENTS

INTRODUCTION	2118
I. BACKGROUND.....	2123
A. <i>The Where, What, and Why of Data Brokers</i>	2125
1. From Where Is Data Taken and What Kind of Data Is Collected?.....	2125
2. How Is the Data Used?.....	2129
B. <i>Specialized Services</i>	2132
C. <i>The Good of Big Data</i>	2133
1. General Advances	2133
2. Big Data Positives in Health Care	2135
II. ANALYSIS.....	2138
A. <i>HIPAA and HITECH</i>	2138
B. <i>Affordable Care Act</i>	2141
C. <i>GINA</i>	2143
1. The Rationale of GINA.....	2143
2. How Genetic Information Is Analogous to Unreliable Health-Related Data	2146

[†] Notes Editor, *Cardozo Law Review*, Volume 41. J.D. Candidate, Benjamin N. Cardozo School of Law, May 2020; B.A., St. John's University, 2015. I would like to thank Professor Felix Wu, my faculty advisor, for his generous feedback and guidance; my Notes Editor, Ava Schein, for her sincerity, advice, and encouragement; and all the editors of Volumes 40 and 41 of *Cardozo Law Review* for their diligence in preparing this Note for publication.

a. Incorrect Predictors.....	2147
b. Fear of Discrimination.....	2150
III. PROPOSAL: EXPAND DISCRIMINATION PROTECTION BEFORE PRIVACY PROTECTION	2151
A. <i>GDPR, CCPA, and the Privacy Approach</i>	2151
1. Potential Flaws in the GDPR and CCPA.....	2153
B. <i>The Anti-Discrimination Approach</i>	2156
1. The Issue of Preemptive Action	2157
2. The Issue of Risk Assessment	2158
CONCLUSION.....	2160

INTRODUCTION

In the age of the “Internet of Things,”¹ where American adults spend nearly six hours a day interacting with digital media,² questions and concerns arise regarding the growth of Big Data³ and its potential impact on privacy and data security.⁴ While there are significant concerns in

¹ “Simply put, this is the concept of basically connecting any device with an on and off switch to the Internet (and/or to each other).” Jacob Morgan, *A Simple Explanation of ‘the Internet of Things,’* FORBES (May 13, 2014, 12:05 AM), <https://www.forbes.com/sites/jacobmorgan/2014/05/13/simple-explanation-internet-things-that-anyone-can-understand/#663bfcc91d09> [<https://perma.cc/6PWG-BDMK>].

² This includes smartphones, desktops, laptops, and other connected devices. Rob Marvin, *Tech Addiction by the Numbers: How Much Time We Spend Online*, PCMAG (June 11, 2018), <https://www.pcmag.com/news/tech-addiction-by-the-numbers-how-much-time-we-spend-online> [<https://perma.cc/E9DG-FMD4>]. Nielsen’s first-quarter 2018 Total Audience Report indicates that the average American spends over eleven hours a day interacting with media in general, which includes radio and television. *Time Flies: U.S. Adults Now Spend Nearly Half a Day Interacting with Media*, NIELSEN (July 31, 2018), <https://www.nielsen.com/us/en/insights/article/2018/time-flies-us-adults-now-spend-nearly-half-a-day-interacting-with-media> [<https://perma.cc/M5UB-4UE2>].

³ “Big Data” is defined as “data sets, typically consisting of billions or trillions of records, that are so vast and complex that they require new and powerful computational resources to process.” *Big Data*, DICTIONARY.COM, <https://www.dictionary.com/browse/big-data> [<https://perma.cc/DD2S-VZ3G>].

⁴ Bernard Marr, *3 Massive Big Data Problems Everyone Should Know About*, FORBES (June 15, 2017, 12:29 AM), <https://www.forbes.com/sites/bernardmarr/2017/06/15/3-massive-big-data-problems-everyone-should-know-about/#f396fc861862> [<https://perma.cc/YRW5-52CU>].

many areas of privacy,⁵ health privacy concerns are among the most foreboding.⁶ Traditionally, the zone of Americans' health privacy has been highly legislated and regulated.⁷ Acts like the Genetic Information Nondisclosure Act (GINA), the Health Information Portability and Accountability Act (HIPAA), and the Health Information Technology for Economic and Clinical Health Act (HITECH) are all in part focused on the protection of Americans' health privacy.⁸ But in the burgeoning world of Big Data, these acts have failed to keep up with the times, becoming porous shields over consumer health privacy.⁹

Substantial legal gaps and loopholes have led to the commodification of “non-traditional,”¹⁰ but “health-related”¹¹ lifestyle data, including information garnered from phone applications, social

(identifying data privacy, data security, and data discrimination as the Big Data concerns “that should keep people up at night”).

⁵ See Olivia Solon, *'Data Is a Fingerprint': Why You Aren't as Anonymous as You Think Online*, GUARDIAN (July 13, 2018, 4:00 AM), <https://www.theguardian.com/world/2018/jul/13/anonymous-browsing-data-medical-records-identity-privacy> [<https://perma.cc/EF5L-4MFR>].

Nameless New York taxi logs were compared with paparazzi shots at locations around the city to reveal that Bradley Cooper and Jessica Alba were bad tippers. In 2017 German researchers were able to identify people based on their “anonymous” web browsing patterns. This week University College London researchers showed how they could identify an individual Twitter user based on the metadata associated with their tweets, while the fitness tracking app Polar revealed the homes and in some cases names of soldiers and spies.

Id.

⁶ See Marshall Allen, *Health Insurers Are Vacuuming Up Details About You—and It Could Raise Your Rates*, PROPUBLICA (July 17, 2018, 5:00 AM), <https://www.propublica.org/article/health-insurers-are-vacuuming-up-details-about-you-and-it-could-raise-your-rates> [<https://perma.cc/NSE3-EMJ3>]; Jamie Wells, *Call It What It Is: Health Insurers Use Your Data to Discriminate Against You*, AM. COUNCIL ON SCI. & HEALTH (July 25, 2018), <https://www.acsh.org/news/2018/07/25/call-it-what-it-health-insurers-use-your-data-discriminate-against-you-13227> [<https://perma.cc/R9FY-KNH6>].

⁷ See *infra* Part III.

⁸ See *infra* Part III.

⁹ See *infra* Part III.

¹⁰ Effy Vayena et al., *Policy Implications of Big Data in the Health Sector*, 96 BULL. WORLD HEALTH ORG. 66, 66–68 (2018), <http://www.who.int/bulletin/volumes/96/1/17-197426/en> [<https://perma.cc/JZ5B-GWDL>].

¹¹ See generally Stacy-Ann Elvy, *Commodifying Consumer Data in the Era of the Internet of Things*, 59 B.C. L. REV. 423 (2018).

media, credit card records, public records, and more.¹² Alone, each tidbit of information can provide some marginal insight into an individual's habits, vices, and general wellbeing; but when consolidated, these resources can paint a purportedly accurate picture of a person's lifestyle.¹³ In the hands of health insurers, this data can be inserted into predictive analytics software to determine an individual's present and future health risks.¹⁴ Ideally, this analysis could work to guide better decision-making in medicine, allowing insurers and doctors to identify high-risk patients and offer individualized treatments.¹⁵ This would increase efficiency in the healthcare sector and reduce costs for everyone involved.¹⁶ In reality, however, there is a concern among consumers that this data will not always be used in furtherance of such honest and good intentions,¹⁷ but as tools nefarious insurers may exploit for their own benefit.¹⁸

In other words, if a person's social media activity, health application data, credit card history, and much more can be employed to predict how much healthcare they might need in the future, then insurers could potentially discriminate against that individual.¹⁹ If a woman purchases

¹² See, e.g., Allen, *supra* note 6; Louise Matsakis, *The WIRED Guide to Your Personal Data (and Who Is Using It)*, WIRED (Feb. 15, 2019, 7:00 AM), <https://www.wired.com/story/wired-guide-personal-data-collection> [<https://perma.cc/3ETP-PADD>].

¹³ C. Mitchell Shaw, *Big Data Knows You Better than You Know Yourself*, NEW AM. (Aug. 10, 2016), <https://www.thenewamerican.com/tech/computers/item/23836-big-data-knows-you-better-than-you-know-yourself> [<https://perma.cc/95XS-KXAE>].

¹⁴ Janine S. Hiller, *Healthy Predictions? Questions for Data Analytics in Health Care*, 53 AM. BUS. L.J. 251, 252–53 (2016).

¹⁵ *Id.* at 251; see also Nathan Cortez, *Substantiating Big Data in Health Care*, 14 I/S: J.L. & POL'Y FOR INFO. SOC'Y 61, 61–62 (2017).

¹⁶ Lisa Hedges, *What Is Big Data in Healthcare and How Is It Already Being Used?*, SOFTWARE ADVICE (Oct. 25, 2019), <https://www.softwareadvice.com/resources/what-is-big-data-in-healthcare-and-whos-already-doing-it> [<https://perma.cc/ET8Z-XJ97>].

¹⁷ See, e.g., Allen, *supra* note 6; see also Harris Poll: *Only Nine Percent of U.S. Consumers Believe Pharma and Biotechnology Put Patients over Profits; Only 16 Percent Believe Health Insurers Do*, HARRIS POLL, <https://theharrispoll.com/only-nine-percent-of-u-s-consumers-believe-pharmaceutical-and-biotechnology-companies-put-patients-over-profits-while-only-16-percent-believe-health-insurance-companies-do-according-to-a-harris-pol> [<https://perma.cc/2ZTP-8LN2>] (finding that only sixteen percent of Americans believe health insurers put patients over profits).

¹⁸ See Allen, *supra* note 6.

¹⁹ *Id.*; see also JANE SARASOHN-KAHN, CAL. HEALTHCARE FOUND., *HERE'S LOOKING AT YOU: HOW PERSONAL HEALTH INFORMATION IS BEING TRACKED AND USED* 3–7 (2014).

plus-size clothing, there is a chance she may be at risk of depression.²⁰ A health application could reveal that a man is much less active of late, pointing to a possibility of cardiac issues in the future.²¹ A woman who changes her name could have gone through a recent divorce or may be newly married and expecting children soon, both of which might lead to higher medical bills.²² A man who eats fast food twice a week may be a costly heart-attack waiting to happen.²³ A woman who purchases allergy medication regularly might be at risk of asthmatic attacks.²⁴ A man's social media account may suggest poor sleeping habits, which could be a sign of mental illness.²⁵ Browsing histories and symptom searches can point to personal concerns about STDs and other diseases.²⁶ Public records could reveal whether a person drives a Ford or a Subaru, which may suggest how likely it is they will be injured in an accident.²⁷

²⁰ Allen, *supra* note 6.

²¹ See Alexandra Troiano, Note, *Wearables and Personal Health Data: Putting a Premium on Your Privacy*, 82 BROOK. L. REV. 1715, 1715–18 (2017); Andrew Boyd, *Could Your Fitbit Data Be Used to Deny You Health Insurance?*, CONVERSATION (Feb. 16, 2017), <http://theconversation.com/could-your-fitbit-data-be-used-to-deny-you-health-insurance-72565> [<https://perma.cc/2XD7-GQ3L>].

²² Allen, *supra* note 6.

²³ Maggie Fox, *Junk Food Causes a Third of Heart Attacks*, REUTERS (Oct. 20, 2008, 9:38 PM), <https://www.reuters.com/article/us-heart-food-idUSTRE49J7T420081021> [<https://perma.cc/L9TP-A6TC>].

²⁴ See *Have Asthma? You Likely Have an Allergy as Well*, AM. COLL. ALLERGY, ASTHMA & IMMUNOLOGY (Apr. 2, 2013), <https://acaai.org/news/have-asthma-you-likely-have-allergy-well> [<https://perma.cc/5JEB-XAHQ>] (reporting that nearly two-thirds of asthmatics also have an allergy); see also STAFF OF S. COMM. ON COMMERCE, SCI. & TRANSP., 113TH CONG., A REVIEW OF THE DATA BROKER INDUSTRY: COLLECTION, USE, AND SALE OF CONSUMER DATA FOR MARKETING PURPOSES 22 (Comm. Print 2013) [hereinafter A REVIEW OF THE DATA BROKER INDUSTRY].

²⁵ *Sleep and Mental Health*, HARV. HEALTH PUB., https://www.health.harvard.edu/newsletter_article/sleep-and-mental-health [<https://perma.cc/TCR2-VL3D>] (last updated June 29, 2018) (exploring the intimate connection between sleep and mental health).

²⁶ Brian Merchant, *Looking Up Symptoms Online? These Companies Are Tracking You*, VICE (Feb. 23, 2015, 10:25 AM), https://motherboard.vice.com/en_us/article/539qzk/looking-up-symptoms-online-these-companies-are-collecting-your-data [<https://perma.cc/EV3S-WX5B>].

²⁷ Brian Naylor, *Firms Are Buying, Sharing Your Online Info. What Can You Do About It?*, NPR (July 11, 2016, 4:51 PM), <https://www.npr.org/sections/alltechconsidered/2016/07/11/485571291/firms-are-buying-sharing-your-online-info-what-can-you-do-about-it> [<https://perma.cc/WP7-C3K7>]; see also Tanya Mohn, *Safest Cars for 2019 Announced—Subaru Tops List, Domestic Brands Nearly Absent*, FORBES (Dec. 26, 2018, 8:27 AM), <https://www.forbes.com/sites/tanyamohn/2018/12/26/safest-cars-for-2019-announced-subaru-tops-list-domestic-brands-nearly-absent/#56e35f8b19f8> [<https://perma.cc/H42G-ESG3>].

Health insurers with access to this ocean of information may choose to deny an individual coverage or increase their premiums, turning the path to affordable healthcare arduous and bleak.²⁸ At present, this threat is largely theoretical, as little empirical evidence exists linking data analytics to discriminatory practices.²⁹ What is certain, however, is that health-related data is being utilized openly in some respects.³⁰ For example, data brokers like Acxiom and LexisNexis are selling this information directly to doctors and other healthcare providers so that these entities can make treatment decisions unique to patients.³¹ While arming doctors with patient-specific health-related data is arguably a novel and progressive idea, how can consumers trust that this same information will not lead to data discrimination in the healthcare insurance realm?³²

This Note argues that the current legislative scheme surrounding the use of data analytics in the healthcare sector is inadequate and that new protections against discrimination must be considered. Simultaneously, it remains important to balance these protections with the potential benefits that the data industry can offer to healthcare. Part I of this Note provides an introduction to the complex framework of the data industry, its relation to health insurers, and its successes and failures. Part II of this

(reporting that Subaru maintains seven of the top safest cars on the market, while domestic brands such as Ford failed to qualify).

²⁸ Nicolas P. Terry, *Regulatory Disruption and Arbitrage in Health-Care Data Protection*, 17 *YALE J. HEALTH POL'Y L. & ETHICS* 143, 180 (2017) ("Health data acquired by data brokers can . . . be looped back into the health-care space for discriminatory purposes."); Wells, *supra* note 6 ("Under the guise of helping patients, the industry's use of data brokers for extensive mining of your personal lifestyle choices and socioeconomic details to determine pricing or premium rates creates a murky terrain that existing, presumed protected health information laws don't explicitly cover.").

²⁹ Maddalena Favaretto et al., *Big Data and Discrimination: Perils, Promises and Solutions*, 6 *J. BIG DATA* 12, 24 (2019) ("Due to the risk of discrimination in data mining and predictive analytics . . . more empirical research is needed to assess how discriminatory practices are deliberately and accidentally emerging from their increased use in . . . sectors such as healthcare . . ."). *But see* Allen, *supra* note 6 (quoting a research scientist on the possibility of data discrimination: "I can't say it hasn't happened").

³⁰ Mohana Ravindranath, *Does Your Doctor Need to Know What You Buy on Amazon?*, *POLITICO* (Oct. 30, 2018, 1:15 PM), <https://www.politico.com/story/2018/10/30/the-doctor-will-see-through-you-now-893437> [<https://perma.cc/SW5K-PEHV>].

³¹ *Id.*

³² Hiller, *supra* note 14, at 251–53.

Note examines the current legislative scheme behind healthcare privacy and discrimination, its inadequacies, and whether and where new protections might be a logical fit. Part III of this Note looks to recent privacy law enactments in the United States and abroad, and why, ultimately, they may be either over-expansive or futile. Part III also poses the idea that expansions in healthcare discrimination law, rather than privacy law, will realize a productive middle-ground between consumers and the healthcare and data industries—allowing the positives of data analytics to continue to flourish while allaying concerns about its potential negatives.³³

I. BACKGROUND

By 2010, the amount of data being collected every two days equaled the amount collected from the beginning of time until 2003.³⁴ Today, each individual is expected to generate enough data in their lifetime to fill three hundred million books.³⁵ Data brokers—shadowy companies often unfamiliar to consumers—specialize in the business of data collection, peddling consumer data to whomever is willing to buy it.³⁶ Some brokers specialize in niche areas like people searches, while others, such as

³³ This Note assumes that Congress will pass antidiscrimination legislation directly targeting insurers, thus circumventing the McCarran-Ferguson Act—a 1945 Act stating that insurance regulation will generally be left to the several states. For more background on the McCarran-Ferguson Act and the insurance industry's exemption from federal anti-trust enforcement, see generally Jonathan R. Macey & Geoffrey P. Miller, *The McCarran-Ferguson Act of 1945: Reconceiving the Federal Role in Insurance Regulation*, 68 N.Y.U. L. REV. 13 (1993).

³⁴ Rick Braddy, *Data Retention: Tough Choices Ahead*, FORBES (Aug. 27, 2018, 7:00 AM), <https://www.forbes.com/sites/forbestechcouncil/2018/08/27/data-retention-tough-choices-ahead/#68147b8f2c94> [https://perma.cc/KRM3-UMEH]. The volume of datasets is now measured in petabytes. SARASOHN-KAHN, *supra* note 19, at 4 (reporting that a petabyte is large enough fit every American's DNA three times, and noting that Walmart had 2.5 petabytes worth of consumer data in 2010 alone).

³⁵ *Watson Health: Get the Facts*, IBM, <https://www.ibm.com/watson-health/about/get-the-facts> [https://perma.cc/4XSX-67KQ].

³⁶ Lois Beckett, *Big Data Brokers: They Know Everything About You and Sell it to the Highest Bidder*, GIZMODO (Mar. 18, 2013, 10:11 AM) <https://gizmodo.com/5991070/big-data-brokers-they-know-everything-about-you-and-sell-it-to-the-highest-bidder> [https://perma.cc/S9N5-2JJ3]; Yael Grauer, *What Are 'Data Brokers,' and Why Are They Scooping Up Information About You?*, VICE (Mar. 27, 2018, 10:00 AM), https://motherboard.vice.com/en_us/article/bjpx3w/what-are-data-brokers-and-how-to-stop-my-private-data-collection [https://perma.cc/4924-39D2].

Acxiom, boast their services to all types of industries, including marketing, finance, and healthcare.³⁷

Private health insurers can be found on many data brokers' clientele lists.³⁸ Insurers can use these companies' data troves to gain insight into the lifestyles and online habits of present and potential policyholders.³⁹ Insurers may also pay middleman companies for more advanced, tailored services. For example, a company called Carrot Health Inc. aggregates data from various sources, including data brokers, and plugs this data into predictive modeling software capable of churning out patient risk scores.⁴⁰ Other times, health insurers may gain a policyholder's consent to data-sharing by offering free products or services in exchange for access to data collected by a policyholder's wearable device, such as an Apple Watch.⁴¹

These partnerships and programs all present potential risks and rewards in the healthcare field. However, before considering the implications of this data-sharing, two central questions must be answered: (1) where do data brokers get their data; and (2) how is the data used?

³⁷ See Grauer, *supra* note 36; *Complete List of Client Categories*, ACXIOM, <https://www.acxiom.com/wp-content/uploads/2019/02/Complete-List-of-Client-Categories.pdf> [https://perma.cc/3KQU-6U35].

³⁸ See Allen, *supra* note 6.

³⁹ *Id.*

⁴⁰ Brian Gormley, *Health Firms Are Looking at Personal Data*, WALL ST. J. (Apr. 29, 2019, 10:03 PM), <https://www.wsj.com/articles/health-firms-are-looking-at-personal-data-11556589780> [https://perma.cc/MH8Q-3ZDL].

⁴¹ For example, Apple and Aetna collaborated on a health application called Attain. The application sets personalized health goals, such as daily calorie thresholds, and communicates with a user's Apple Watch to monitor user progress. The application rewards users who achieve their goals with points, which can then be redeemed to earn gift cards from popular retailers. This data is shared with Aetna, though they claim they will not use the information to set prices or increase premiums. See Rebecca Pifer, *Aetna Launches App for Apple Watch*, HEALTHCARE DIVE (May 8, 2019), <https://www.healthcaredive.com/news/aetna-launches-app-for-apple-watch/553987> [https://perma.cc/9JN7-MCSR]; see also *Frequently Asked Questions*, ATTAIN BY AETNA, <https://www.attainbyaetna.com/faq.html> [https://perma.cc/B69F-FQ6W].

A. *The Where, What, and Why of Data Brokers*

In recent years, controversies over data breaches and data hemorrhaging have been ushered into the national spotlight.⁴² While many of these controversies concern hackers stealing personal data, such as names, addresses, and credit card information, and have led to criticisms of companies' security measures, there is a glut of other, non-personal data being collected and analyzed by data companies, who then sell this information and the inferences garnered from it to advertisers, creditors, insurers, and others.⁴³

1. From Where Is Data Taken and What Kind of Data Is Collected?

Data is collected from just about everywhere.⁴⁴ Social media websites, phone applications, web browsers, email services, and credit cards are all used as tools for collecting data.⁴⁵ The end result is a vast marketplace of data reservoirs where companies buy, sell, or exchange data pursuant to business contracts.⁴⁶ Data brokers occupy a central role in the data marketplace,⁴⁷ compiling information from various public and private sources to build consumer profiles wherein individuals are linked to specific groups or placed in data segments.⁴⁸ Data brokers then sell this derived data to their clients.⁴⁹ However, due to the lack of transparency in the data industry, it remains unclear how many companies purchase

⁴² See, e.g., *Facebook Data Breach Update Reveals Theft of 30M Users' Data*, ZACKS (Oct. 16, 2018), <https://www.zacks.com/stock/news/328243/facebook-data-breach-update-reveals-theft-of-30m-users-data?cid=CS-ZC-FT-328243> [<https://perma.cc/CZ37-NYM7>].

⁴³ See Allen, *supra* note 6.

⁴⁴ Grauer, *supra* note 36.

⁴⁵ Hiller, *supra* note 14, at 252–53; Beckett, *supra* note 36.

⁴⁶ Amy Schmitz, *Secret Consumer Scores and Segmentations: Separating “Haves” from “Have-Nots,”* 2014 MICH. ST. L. REV. 1411, 1421 (2014) (“Indeed, data collection and aggregation transpires through an unimaginable labyrinth of information sharing among merchants and data brokers.”).

⁴⁷ Stacy-Ann Elvy, *Paying for Privacy and the Personal Data Economy*, 117 COLUM. L. REV. 1369, 1372 (2017).

⁴⁸ A REVIEW OF THE DATA BROKER INDUSTRY, *supra* note 24, at 1.

⁴⁹ John A. Rothchild, *Against Notice and Choice: The Manifest Failure of the Proceduralist Paradigm to Protect Privacy Online (Or Anywhere Else)*, 66 CLEV. ST. L. REV. 559, 582 (2018).

and sell consumer data,⁵⁰ and data brokers are not particularly forthcoming about their specific data sources.⁵¹

Nevertheless, company privacy policies across the internet and mobile applications display remarkable uniformity.⁵² Privacy policies, which are rarely read or understood by the users of the service,⁵³ often grant companies the right to collect and share consumer data with “third parties” and “affiliates.”⁵⁴ However, these parties are rarely specified by name, and even if consumers were able to discover the identities of these parties, they would be hard-pressed to find out who those parties pass the information onto, and so forth.⁵⁵ Therefore, any data that the collecting entity imparts to others could potentially end up in the hands of data brokers.⁵⁶

Though companies often shed personally identifiable information,⁵⁷ which presents significant issues, companies also attempt to justify disclosures by granting themselves the right to share only de-identified, non-personal information.⁵⁸ For example, Fitbit’s privacy policy reads: “We may share non-personal information that is aggregated or de-identified so that it cannot reasonably be used to identify an individual. We may disclose such information publicly and to third parties”⁵⁹ Similarly, Snapchat’s privacy policy states: “We may also share with third parties . . . aggregated, non-personally identifiable, or de-identified information.”⁶⁰ Neither privacy policy provides any clarity as to what

⁵⁰ Elvy, *supra* note 47, at 1398.

⁵¹ A REVIEW OF THE DATA BROKER INDUSTRY, *supra* note 24, at 12–34. Even when prompted to reveal their sources by a government committee, many data brokers refused, citing confidentiality clauses in their business contracts among other reasons. *Id.*

⁵² See generally Rothchild, *supra* note 49.

⁵³ Kevin Litman-Navarro, *We Read 150 Privacy Policies. They Were an Incomprehensible Disaster.*, N.Y. TIMES (June 12, 2019), <https://www.nytimes.com/interactive/2019/06/12/opinion/facebook-google-privacy-policies.html> [<https://perma.cc/BYN9-TZGH>].

⁵⁴ Rothchild, *supra* note 49, at 617.

⁵⁵ *Id.*

⁵⁶ *Id.*

⁵⁷ *Id.* at 564–81.

⁵⁸ Elvy, *supra* note 47, at 1415.

⁵⁹ *Privacy Policy*, FITBIT, <https://www.fitbit.com/legal/privacy-policy> [<https://perma.cc/9DQ4-3Z4R>].

⁶⁰ *Privacy Policy*, SNAP INC., <https://www.snap.com/en-US/privacy/privacy-policy> [<https://perma.cc/5M2Y-JW7F>].

constitutes de-identified information, nor do they specify the third parties with whom they share this information.⁶¹

Many companies follow this same anonymization and sharing model.⁶² However, anonymization can sometimes be an illusory justification for data sharing.⁶³ Scholars have argued that due to the amount and quality of data collected, it is possible for data brokers and other entities with access to multiple datasets to link anonymized data to individuals.⁶⁴ One recent study found that 99.98% of Americans could be identified in a dataset that contemplated only fifteen demographic characteristics.⁶⁵ Other scholars contend that the threat of re-identification is overstated, and that not every grain of anonymous data

⁶¹ Other examples include Google, whose policy provides: “We may share non-personally identifiable information publicly and with our partners” *Privacy Policy*, GOOGLE, <https://policies.google.com/privacy?hl=en-US> [<https://perma.cc/9JMC-8BH5>]. Twitter’s policy, which is more specific, provides:

We share or disclose non-personal data, such as aggregated information like the total number of times people engaged with a Tweet, demographics, the number of people who clicked on a particular link or voted on a poll in a Tweet (even if only one did), the topics that people are Tweeting about in a particular location, some inferred interests, or reports to advertisers about how many people saw or clicked on their ads.

Privacy Policy, TWITTER, <https://twitter.com/en/privacy> [<https://perma.cc/SW4J-A4KG>].

⁶² Jennifer Schlesinger & Andrea Day, *Most People Just Click and Accept Privacy Policies Without Reading Them—You Might Be Surprised at What They Allow Companies to Do*, CNBC (Feb. 7, 2019), <https://www.cnbc.com/2019/02/07/privacy-policies-give-companies-lots-of-room-to-collect-share-data.html> [<https://perma.cc/7VY6-UR42>].

⁶³ See *id.* (quoting a privacy expert who asserts that just a few points of location data could be used to deanonymize an individual); see also Elvy, *supra* note 47, at 1415; see generally Jennifer Valentino-DeVries et al., *Your Apps Know Where You Were Last Night, and They’re Not Keeping It Secret*, N.Y. TIMES (Dec. 10, 2018), <https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html> [<https://perma.cc/T4VH-F3A7>] (reporting that several companies track locations of up to two hundred million devices through phone applications, which sometimes update more than fourteen thousand times per day and can pinpoint a person’s location “within a few yards”).

⁶⁴ See, e.g. Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701 (2010). “[T]he power of reidentification will create and amplify privacy harms. Reidentification combines datasets that were meant to be kept apart, and in doing so, gains power through accretion: Every successful reidentification, even one that reveals seemingly nonsensitive data like movie ratings, abets future reidentification.” *Id.* at 1705.

⁶⁵ Yves-Alexandre de Montjoye et al., *Estimating the Success of Re-identifications in Incomplete Datasets Using Generative Models*, NATURE COMM., July 2019, at 5.

can be meaningfully re-identified.⁶⁶ However, because companies are intentionally vague about who they share non-personal data with, or what non-personal data is shared, it is certainly possible that this de-identified information is falling into the hands of data brokers, who can combine it with their own datasets to identify and draw inferences about consumers.

Data brokers also find useful information in countless other ways, many of which are outlined in a Federal Trade Commission report from 2014.⁶⁷ Brokers can obtain motor vehicle and driving records from state governments, and demographic statistics for particular city blocks from the U.S. Census Bureau.⁶⁸ Brokers may also use programs known as “web crawlers” to scour publicly available webpages, such as social media sites and blogs, wherein they can garner information on users who maintain public profiles.⁶⁹ Brokers can use software to help them decide which sites to comb and how often, as well as specify what kind of data the crawlers should be picking up.⁷⁰ Brokers may also rely on commercial sources, such as retailers and catalog companies, for information on consumers’ individual product purchases.⁷¹ For example, a person who signs up for a store loyalty card or rewards program may share their name and telephone number with the retailer, who may pass this information—as well as the individual’s purchase history—along to data brokers.⁷² Data brokers also utilize “third-party cookies” to gain insight into an individual’s activities on the web.⁷³ Cookies and other technologies, such

⁶⁶ See Felix T. Wu, *Defining Privacy and Utility in Data Sets*, 84 U. COLO. L. REV. 1117 (2013). “One cannot talk about the success or failure of anonymization in the abstract. Anonymization encompasses a set of technical tools that are effective for some purposes, but not others.” *Id.* at 1126.

⁶⁷ FED. TRADE COMM’N, *DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY* (2014) [hereinafter *A CALL FOR TRANSPARENCY AND ACCOUNTABILITY*].

⁶⁸ *Id.* at 11–12. Other publicly available government information includes applications for professional and recreational licenses, real property records, voter registration information, driving records, and court records—though the use of some of this data is sometimes subject to restrictions. *Id.*

⁶⁹ *Id.* at 17.

⁷⁰ *Id.* For more information about web crawlers, which also play a vital role in the operation of Google and other search engines, see Zachary Gold & Mark Latonero, *Robots Welcome? Ethical and Legal Considerations for Web Crawling and Scraping*, 13 WASH. J.L., TECH. & ARTS 275 (2018).

⁷¹ Julia N. Mehlman, Note, *If You Give a Mouse a Cookie, It’s Going to Ask for Your Personally Identifiable Information*, 81 BROOK. L. REV. 329, 335 (2015).

⁷² *Id.*

⁷³ *Id.* at 333–34.

as web beacons, are embedded in websites and allow those sites to track users as they click around the web.⁷⁴ First-party cookies are cookies that the host site—the site the consumer intends to visit—places in the consumer’s browser.⁷⁵ However, many websites also allow third-parties to place cookies.⁷⁶ For example, Yahoo’s privacy policy lists the hundreds of third parties, including data brokers such as Datalogix and Experian, that are allowed to place cookies and web beacons from Yahoo’s websites.⁷⁷

And finally, to complicate matters further, data brokers often buy and sell information with each other, making it “virtually impossible for a consumer to determine the originator of a particular data element.”⁷⁸

2. How Is the Data Used?

The end result of all of this data mining from various sources is a mountain of raw data that data brokers and data analysts try to organize for their purchasers and clients.⁷⁹ In its 2014 study of nine of the top data brokers in the country, the Federal Trade Commission (FTC) found that one of the brokers had information on 1.4 billion consumer transactions and over 700 billion total data elements, while another broker was adding three billion data elements to its records each month.⁸⁰

⁷⁴ See Rothchild, *supra* note 49, at 567–68, 568 n.39.

⁷⁵ Michal Wlosik & Michael Sweeney, *What’s the Difference Between First-Party and Third-Party Cookies?*, CLEARCODE: ADTECH PROCESSES, DATA & PRIVACY BLOG, <https://clearcode.cc/blog/difference-between-first-party-third-party-cookies> [<https://perma.cc/K8SX-C58V>] (“First-party cookies are stored by the domain (website) you are visiting directly. They allow website owners to collect analytics data, remember language settings, and perform other useful functions that help provide a good user experience.”). “Third party cookies are created by domains other than the one you are visiting directly, hence the name third-party. They are used for cross-site tracking, retargeting and ad-serving.” *Id.*

⁷⁶ Mehlman, *supra* note 71, at 333–34.

⁷⁷ *Third Parties*, YAHOO!, <https://policies.yahoo.com/xa/en/yahoo/privacy/topics/thirdparties/index.htm> [<https://perma.cc/DST8-RTMD>].

⁷⁸ A CALL FOR TRANSPARENCY AND ACCOUNTABILITY, *supra* note 67, at 14.

⁷⁹ *Id.* at 19.

⁸⁰ *Id.* at iv. The nine companies that were the subject of the report were: Acxiom, Corelogic, Datalogix, eBureau, ID Analytics, Intelius, PeekYou, Rappleaf, and Recorded Future. *Id.* at ii. The number of data elements is comprised of both “actual data elements”—like names, ages, and

In order to make this information useful, brokers use these data elements to create consumer categories, or “segments.”⁸¹ These range from relatively unassuming segments, such as “Dog Owner,” to moderately presumptuous segments, such as “Expectant Parent,” “Married Sophisticates,” “Cholesterol Focus,” and others that focus on ethnicity, income, and potential health statuses.⁸² Clients request this information from data brokers and use it for any of the three main purposes identified by the FTC: marketing,⁸³ risk mitigation,⁸⁴ or people searches.⁸⁵ Insurers are identified as one of the clients that purchase data for use in the subcategory of marketing analytics.⁸⁶ As the FTC notes, this is a troublesome practice in the insurance context, as a consumer belonging to a “Biker Enthusiasts” data segment may gladly accept coupons from a motorcycle dealership, unaware that an insurer may be

addresses—as well as “derived data elements,” which are inferences that brokers make about consumers by analyzing data. *Id.* at 19.

⁸¹ *Id.*

⁸² *Id.* at iv–v.

⁸³ Marketing purposes include all kinds of online and offline tailored advertising—including newsletters, email solicitations, and advertisement banners on social media websites. *Id.* at ii–iii. Consumers have become wise to this kind of marketing, as they are constantly followed by advertisements for items and events they have searched for on Google or browsed for on Amazon. These creepy advertisements are smartly dubbed “stalker ads.” Brian X. Chen, *Are Targeted Ads Stalking You? Here’s How to Make Them Stop*, N.Y. TIMES (Aug. 15, 2018), <https://www.nytimes.com/2018/08/15/technology/personaltech/stop-targeted-stalker-ads.html> [<https://perma.cc/CL76-AXWV>]. Poetically, as I researched this issue, I was pummeled with advertisements that somehow incorporated “Big Data.”

⁸⁴ Risk mitigation purposes entail the use of data to verify individuals’ identities and to detect fraud. See A CALL FOR TRANSPARENCY AND ACCOUNTABILITY, *supra* note 67, at iii. Companies may, for example, use data to find out if a consumer is lying to them about their identity or has submitted false information to them. See *id.*; see also *Fighting Fraud with Big Data Analytics*, CIO REV. (Apr. 13, 2018), <https://www.cioreview.com/news/fighting-fraud-with-big-data-analytics-nid-26041-cid-141.html> [<https://perma.cc/E7UL-LR6W>] (“The use of big data analysis helps businesses to detect potential fraud instances and employ effective threat mitigation programs.”).

⁸⁵ People search products are websites that allow individuals to find information about others, whether they are old friends, new neighbors, love interests, or even business competitors. A CALL FOR TRANSPARENCY AND ACCOUNTABILITY, *supra* note 67, at iii; see also, Patrick Allan & Emily Long, *How to Opt Out of the Most Popular People Search Sites*, LIFEHACKER (Dec. 12, 2019, 4:28 PM), <https://lifehacker.com/how-to-opt-out-of-the-most-popular-people-search-sites-1791536533> [<https://perma.cc/W7QV-TMZB>] (“[A]ddresses, family members, and known associates, available for all to see.”).

⁸⁶ A CALL FOR TRANSPARENCY AND ACCOUNTABILITY, *supra* note 67, at 39. Marketing analytics are employed as “a way to predict consumers’ likely behavior.” *Id.* at 31.

using that same segment to draw a negative inference about the person's risky behavior.⁸⁷

Importantly, data brokers do not always allow individuals the opportunity to correct false information about them, regardless of the purpose for which the information is being used.⁸⁸ Further, data brokers that allow individuals to delete raw data may not allow individuals to delete the inferences derived from that data.⁸⁹ This means that potentially inaccurate information could be used to make vital decisions in the insurance realm.⁹⁰

The business of data is certainly booming. In 2012, the nine companies in the FTC study generated \$426 million in revenue combined.⁹¹ More recently, just one of the companies from the 2012 study, Acxiom, was expected to generate over \$900 million in 2018 alone.⁹² Equifax, a company that was not a party to the FTC's study, generated over three billion dollars in revenue in 2016.⁹³ This exponential

⁸⁷ *Id.* at vi.

⁸⁸ *Id.* at iii. Curious journalists have searched for and requested information about themselves from data brokers, only to find the information to be largely incorrect. See Caitlyn Renee Miller, *I Bought a Report on Everything That's Known About Me Online*, ATLANTIC (June 6, 2017), <https://www.theatlantic.com/technology/archive/2017/06/online-data-brokers/529281> [<https://perma.cc/S467-KZQ3>]; Kalev Leetaru, *The Data Brokers So Powerful Even Facebook Bought Their Data—but They Got Me Wildly Wrong*, FORBES (Apr. 5, 2018, 4:08 PM), <https://www.forbes.com/sites/kalevleetaru/2018/04/05/the-data-brokers-so-powerful-even-facebook-bought-their-data-but-they-got-me-wildly-wrong/#3b0d2c6e3107> [<https://perma.cc/M93A-6WH3>].

⁸⁹ Grauer, *supra* note 36.

⁹⁰ See *infra* Section II.B; see also Timo Elliott, *How Trustworthy Is Big Data?*, BRINK (Feb. 2, 2018), <https://www.brinknews.com/how-trustworthy-is-big-data> [<https://perma.cc/3XT4-HW3C>] (describing big data as “at best a fuzzy source of truth”).

⁹¹ A CALL FOR TRANSPARENCY AND ACCOUNTABILITY, *supra* note 67, at 23.

⁹² Wesley Brown, *Acxiom Shares Off 7% After Fourth Quarter, Yearly Results Miss Wall Street Targets*, TALK BUS. & POL. (May 17, 2017, 1:53 PM), <https://talkbusiness.net/2017/05/acxiom-shares-off-7-after-fourth-quarter-yearly-results-miss-wall-street-targets> [<https://perma.cc/TGK4-22LM>]. Indeed, data has been called “the new oil,” which is an appropriate analogy on more than just an economic level. Like crude oil, harvested raw data must be transformed into something useful. See Charles Arthur, *Tech Giants May Be Huge, but Nothing Matches Big Data*, GUARDIAN (Aug. 23, 2013, 3:21 PM), <https://www.theguardian.com/technology/2013/aug/23/tech-giants-data> [<https://perma.cc/KW6S-K39P>].

⁹³ Bernard Marr, *Where Can You Buy Big Data? Here Are the Biggest Consumer Data Brokers*, FORBES (Sept. 7, 2017, 12:28 AM), <https://www.forbes.com/sites/bernardmarr/2017/09/07/where->

growth is staggering, and somewhat startling, considering the fact that consumers are largely unfamiliar with this enormous, but shadowy, industry.⁹⁴

B. *Specialized Services*

Mammoth data brokers like Acxiom are not the only companies providing services for health insurers. For example, a Minneapolis-based software company called Carrot Health Inc. gathers information from various sources, including data brokers and public records, to calculate patient risk scores for its health insurer clients.⁹⁵ Its software, Carrot MarketView, “combines social and behavioral data” from over seventy sources “with predictive modeling to deliver actionable insights into consumers and their underlying determinants of health.”⁹⁶ A brochure for their MarketView product brags that the software can provide “an unprecedented view” into the social determinants of health of members and prospects in the insurer’s service area “down to the identified individual level.”⁹⁷ Carrot Health further claims that it has information on 250 million identified U.S. adults.⁹⁸

Other data companies like Optum, LexisNexis, and IBM Watson Health allow insurers to merge the insurer’s data, namely, current policyholders’ electronic health records (which include clinical and payment records), with the personal, socioeconomic details that data

can-you-buy-big-data-here-are-the-biggest-consumer-data-brokers [https://perma.cc/MZB5-DUHE].

⁹⁴ Sacha Molitorisz, *It’s Time for Third-Party Data Brokers to Emerge from the Shadows*, CONVERSATION (Apr. 4, 2018, 2:46 AM), <https://theconversation.com/its-time-for-third-party-data-brokers-to-emerge-from-the-shadows-94298> [https://perma.cc/5N2B-PCWC] (“To date, there has been insufficient transparency, insufficient fairness and insufficient respect for user consent. This applies to Facebook, but also to app developers, and to Acxiom, Experian, Quantum and other data brokers.”).

⁹⁵ Gormley, *supra* note 40.

⁹⁶ *Carrot MarketView: Bringing Consumerism to Healthcare*, CARROT HEALTH, <https://carrothealth.com/wp-content/uploads/2018/04/MarketView-For-Payers.pdf> [https://perma.cc/G8ZS-UNZR].

⁹⁷ *Id.*

⁹⁸ *Id.*

brokers farm.⁹⁹ In other words, information on patient medical outcomes can be linked to personal details like education, family structure, or net worth.¹⁰⁰ For example, at a 2018 health data conference in San Diego, Optum’s marketing materials boasted that the company was able to combine claims and clinical information with social media interactions.¹⁰¹ Optum claims that this socioeconomic information is not used to price health plans, but when asked by a reporter what this information was used for, Optum’s salespeople contended that they were not allowed to say.¹⁰² Insurers, meanwhile, promise that they only use broker information to identify health risks in patients and target them for services.¹⁰³ However, when asked whether pricing decisions were made based off of this data, one research scientist ominously responded: “I can’t say it hasn’t happened.”¹⁰⁴

C. *The Good of Big Data*

1. General Advances

Despite the astonishing revelation that much of one’s life “is an open book,”¹⁰⁵ there are plenty of positives that have come out of this all-encompassing ocean of information and out of the world of data analytics generally. In May 2014, the White House released a report regarding the future of data and how they believed it would revolutionize financial, manufacturing, healthcare, and other industries.¹⁰⁶ Since last April,

⁹⁹ Allen, *supra* note 6.

¹⁰⁰ *Id.*

¹⁰¹ *Id.*

¹⁰² *Id.*

¹⁰³ *Id.*

¹⁰⁴ *Id.*

¹⁰⁵ See *Here’s the Real Score: Big Data Knows Everything About You!*, NBC NEWS (Apr. 9, 2014, 12:25 PM), <https://www.nbcnews.com/business/consumer/heres-real-score-big-data-knows-everything-about-you-n75741> [<https://perma.cc/ZYZ8-E8EH>]; see also Naylor, *supra* note 27 (“There are some big companies out there that you’ve probably never heard of, that know more about you than you can imagine.”).

¹⁰⁶ See JOHN PODESTA ET AL., BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES 39 (2014), https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2

fifteen cold murder and sexual assault cases have been solved using GEDMatch, a genealogy website normally used to find biological relatives and flesh out family trees.¹⁰⁷ Educational institutions have used data to identify exceptional students who can participate in advanced classes, generating opportunities for these students that would have been unavailable otherwise.¹⁰⁸ Bloomberg has developed a data initiative known as the “Data for Good Exchange,” an annual conference wherein data scientists, government agencies, and nonprofits convene to discuss how data can be utilized to solve core societal issues, such as climate change.¹⁰⁹ Even at Disney World, high-tech “MagicBands” capable of tracking guests’ locations and other activities can, among many other things, deliver free vouchers to guests who have waited on long lines.¹¹⁰

014.pdf [https://perma.cc/9M94-PVUM]. Indeed, the report outlines several blooming advantages of the data industry. Technology companies are using big data to analyze millions of voice samples to deliver more reliable and accurate voice interfaces. Banks are using big data techniques to improve fraud detection. Health care providers are leveraging more detailed data to improve patient treatment. Big data is being used by manufacturers to improve warranty management and equipment monitoring, as well as to optimize the logistics of getting their products to market. Retailers are harnessing a wide range of customer interactions, both online and offline, in order to provide more tailored recommendations and optimal pricing. *Id.*

¹⁰⁷ Heather Murphy, *How an Unlikely Family History Website Transformed Cold Case Investigations*, N.Y. TIMES (Oct. 15, 2018), <https://www.nytimes.com/2018/10/15/science/gedmatch-genealogy-cold-cases.html> [https://perma.cc/2PNK-LUUE].

¹⁰⁸ FED. TRADE COMM’N, *BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?* 6 (2016).

¹⁰⁹ See Randy Bean, *Bloomberg’s Data Initiative: Big Data for Social Good in 2018*, FORBES (Jan. 2, 2018, 6:31 PM), <https://www.forbes.com/sites/ciocentral/2018/01/02/bloombergs-datainitiative-big-data-for-social-good-in-2018/#4f5e1e603a44> [https://perma.cc/D9XG-3MEJ]; see generally *Data for Good Exchange 2018*, BLOOMBERG, <https://www.bloomberg.com/lp/d4gx-2018/> [https://perma.cc/S88P-9MY6]. A recent venture sought to use data to understand the causes of street litter and thereby inform policy recommendations to reduce litter and limit floatable trash in New York City’s waterways. Ester R. Fuchs, *Smart Cities, Stupid Cities, and How Data Can be Used to Solve Urban Policy Problems*, TECH AT BLOOMBERG (Aug. 25, 2017), <https://www.techatbloomberg.com/blog/smart-cities-stupid-cities-and-how-data-can-be-used-to-solve-urban-policy-problems> [https://perma.cc/RW9F-FBU6].

¹¹⁰ Bernard Marr, *Disney Uses Big Data, IoT and Machine Learning to Boost Customer Experience*, FORBES (Aug. 24, 2017, 12:28 AM), <https://www.forbes.com/sites/bernardmarr/2017/08/24/disney-uses-big-data-iot-and-machine-learning-to-boost-customer-experience/#2db87ccb3387> [https://perma.cc/J24R-QG7D] (“The bands act as hotel keys, credit cards, tickets, FastPasses and more. With a simple swipe of the band across sensors located throughout the park, the giant system knows where you are, what you’re doing and what you need.”).

2. Big Data Positives in Health Care

In the healthcare field, data analytics has led to significant improvements both abroad and in the United States. In the European Union, data is streamlining the healthcare industry, making it more efficient and guiding improvements in clinical decision-making.¹¹¹ In Hong Kong, Big Data is also being used to increase efficiency, shortening patient-doctor visits and curtailing the duplication of unnecessary clinical tests and procedures.¹¹² The data explosion has led to more constructive treatments of various “simple” diseases, and a Swiss company is now looking to use data to help find more effective treatments for “complex” health issues like dementia and cardiovascular disease.¹¹³ The World Health Organization has also supported and encouraged countries’ uses of Big Data in the health sector, noting that data obtained from “non-traditional” data streams, like social media and online purchases, helped to contain the 2015 Zika virus outbreak.¹¹⁴

¹¹¹ The European Commissioner of Health, Vytenis Andriukaitis, outlined some of these benefits at a health conference in Belgium, including: “[d]iscovering patterns which could lead to new prevention, diagnostic and therapeutic techniques; [e]mpowering citizens to address their own health challenges; [a]ddressing inefficiencies and reducing waste in care processes; and [i]mproving analysis to facilitate better clinical decisions, and enhance quality of care, patient safety and patient outcomes.” Aran Noden, *Andriukaitis: Big Data Is Shaping Health Policy in Europe*, HEALTH EUROPA (Mar. 29, 2018), <https://www.healtheuropa.eu/andriukaitis-big-data-is-shaping-health-policy/85154> [<https://perma.cc/B4YT-9R6Z>]. Notably, the flow of data in the European Union is expected to be somewhat mitigated after the passing of the General Data Protection Regulation (GDPR), which came into effect in May of 2018. So far, though, there have been some issues with compliance, particularly among healthcare organizations. Leontina Postelnicu, *GDPR Four Months in—What Has Changed?*, HEALTHCARE IT NEWS (Oct. 8, 2018, 11:13 AM), <https://www.healthcareitnews.com/news/gdpr-four-months-%E2%80%93-what-has-changed> [<https://perma.cc/679S-ULFJ>].

¹¹² Leung Pak-Yin, *In Hong Kong, Big Data Is Helping People Take Control of Their Health Care*, S. CHINA MORNING POST, <https://www.scmp.com/comment/insight-opinion/article/2094499/hong-kong-big-data-helping-people-take-control-their-health> [<https://perma.cc/8YJA-4VS8>].

¹¹³ Gary Finnegan, *Complex Diseases Get the Big Data Treatment*, HORIZON (May 23, 2018), <https://horizon-magazine.eu/article/complex-diseases-get-big-data-treatment.html> [<https://perma.cc/UH34-QETF>].

¹¹⁴ Vayena et al., *supra* note 10.

The value of health research based on non-traditional data streams from Internet-based applications, platforms, e.g. social media and services, e.g. email and online purchasing, has already been demonstrated. For example, during the Zika virus outbreak in 2015, analyses of reports in the online media helped to supplement existing information, close

In the United States, data has been important to healthcare for many years,¹¹⁵ but the advent of new data sources and better storage capabilities has led to fascinating healthcare projects in the public sector.¹¹⁶ For example, health-related data is at the center of certain government research projects. The Center for Disease Control and Prevention (CDC) developed a project called BioMosaic, which combines airline records, disease reports, and demographic data to alert public health officials to the threat of disease outbreaks in cities, counties, and even hospitals—all through a convenient phone application.¹¹⁷ The Department of Health and Human Services (HHS) released a newsletter in July 2017, summarizing a report that centered on wearable devices tracking physical activity around the world, which, they note, found that there is a direct correlation between more “walkable” cities and improved public health.¹¹⁸ Data analytics has also been useful to law enforcement in the healthcare sector. In 2007, HHS created a Medicare Fraud Strike Force under the Office of the Inspector General, which employs data analytics and combines the resources of federal, state, and local law enforcement to combat healthcare fraud.¹¹⁹ The Department of Justice and other

knowledge gaps and allow researchers to estimate transmission dynamics and plan response measures that extended beyond vector suppression.

Id. at 66. Importantly, at the end of the bulletin, the WHO also noted the privacy implications of using such health-related data. *Id.* at 68.

¹¹⁵ Kristin Madison, *Health Care Regulators as Data Stewards*, 92 N.C. L. REV. 1605, 1609–10 (2014).

¹¹⁶ *See id.*

¹¹⁷ Amy O’Leary, *In New Tools to Combat Epidemics, the Key Is Context*, N.Y. TIMES (June 19, 2013, 10:00 PM), <https://bits.blogs.nytimes.com/2013/06/19/in-new-tools-to-combat-epidemics-the-key-is-context/?smid=tw-share> [<https://perma.cc/2YB5-VAKV>]. Notably, Google experimented with a similar project, Google Flu Trends, which was intended to track flu outbreaks. This project failed miserably, with Google overestimating the 2013 flu outbreak by 140%. *See* David Lazer & Ryan Kennedy, *What We Can Learn from the Epic Failure of Google Flu Trends*, WIRED (Oct. 1, 2015, 7:00 AM), <https://www.wired.com/2015/10/can-learn-epic-failure-google-flu-trends> [<https://perma.cc/4F64-BT84>] (“When Google quietly euthanized the program . . . it turned the poster child of big data into the poster child of the foibles of big data.”).

¹¹⁸ U.S. DEP’T HEALTH & HUMAN SERVS., RESEARCH MATTERS—TRACKING PHYSICAL ACTIVITY AROUND THE WORLD (2017).

¹¹⁹ *Medicare Fraud Strike Force*, HHS OFFICE OF INSPECTOR GEN., <https://oig.hhs.gov/fraud/strike-force> [<https://perma.cc/3SDL-6MZM>]. This strike force has paid incredible dividends—with HHS predicting over five billion taxpayer dollars were saved in 2016 alone. *Taxpayers Could See*

agencies operate similar programs.¹²⁰ These initiatives were made possible by an expanding use of data—benefits which undoubtedly should not be mired by new, potentially crippling privacy laws.¹²¹

Data utilization is also a central focus in the private sector. One of the driving forces behind the recent CVS-Aetna merger was a mutual desire to combine the companies' respective datasets.¹²² Among other benefits, this interconnectedness makes it easier for consumers to understand which pharmaceuticals their health plans cover and what their out-of-pocket costs will be before they arrive at the pharmacy, and enables Aetna to track and limit the occurrence of unfilled prescriptions.¹²³ There are rumors of other collaborations, mergers, and buyouts that are centered around healthcare data as well.¹²⁴ Additionally,

Billions Saved as the Result of HHS OIG Work, HHS OFFICE OF INSPECTOR GEN. (Dec. 13, 2016), <https://oig.hhs.gov/newsroom/news-releases/2016/oig-work.asp> [<https://perma.cc/J4MW-2VLR>].

¹²⁰ See *Health Care Fraud Unit*, U.S. DEP'T JUSTICE, <https://www.justice.gov/criminal-fraud/health-care-fraud-unit> [<https://perma.cc/Y3YJ-P8YF>].

¹²¹ See *infra* Section III.A.

¹²² Gary Bloom, *Disrupting Health Care: From Amazon to CVS, Data Is at the Heart of It*, FORBES (July 6, 2018, 7:30 AM), <https://www.forbes.com/sites/forbestechcouncil/2018/07/06/disrupting-health-care-from-amazon-to-cvs-data-is-at-the-heart-of-it/#4425467d1c06> [<https://perma.cc/VTR3-BTL8>] (quoting CVS CEO Larry Merlo: "With the analytics of Aetna and CVS Health's human touch, we will create a health care platform built around individuals").

¹²³ See Jennifer Bresnick, *How the CVS, Aetna Deal Will Overhaul Healthcare Big Data Analytics*, HEALTH IT ANALYTICS (Dec. 5, 2017), <https://healthitanalytics.com/news/how-the-cvs-aetna-deal-will-overhaul-healthcare-big-data-analytics> [<https://perma.cc/Y4KD-VQ8P>]. Patients' failure to take prescribed medication is known as "medication non-adherence," which can lead to unnecessary hospital readmissions and emergency room visits resulting in hundreds of billions in downstream healthcare costs. *Id.*

¹²⁴ Bernard Marr, *Apple & IBM Team Up for New Big Data Health Platform*, LINKEDIN (Apr. 15, 2015), <https://www.linkedin.com/pulse/apple-ibm-team-up-new-big-data-health-platform-bernard-marr?trk=mp-reader-card> [<https://perma.cc/X46Y-ZENE>] (addressing a partnership between Apple and IBM which allows Apple's iPhone and Apple Watch users to share data with IBM's Watson Health, a cloud-based analytics service); Martha C. White, *Cigna Express Scripts Merger Could Be Welcome Medicine—for Shareholders*, NBC NEWS (Mar. 8, 2018, 1:25 PM), <https://www.nbcnews.com/business/business-news/cigna-express-scripts-merger-could-be-welcome-medicine-shareholders-n854881> [<https://perma.cc/UPL2-WK6T>] (addressing the Cigna-Express Scripts merger and how the merger would "give Cigna more data and better insight into the total scope of healthcare people get, reducing the fragmentation that can lead to less effective or duplicative care"). In the Spring of 2018, there was also speculation that Walmart would expand its relationship with Humana, which would have undoubtedly been motivated in part by data. Dana Mattioli et al., *Walmart in Early-Stage Acquisition Talks with Humana*, WALL ST. J. (Mar. 29, 2018, 8:02 PM), <https://www.wsj.com/articles/walmart-in-early-stage-acquisition-talks-with-humana>

a company called the Pittsburgh Health Alliance is using machine learning to analyze data from various sources, including wearable devices and social media, to better understand patient needs and offer more precise, personalized treatments for cancer patients.¹²⁵ While these private initiatives are industrious, they also toe a dangerous line. Mergers centered on the convergence of data resources and the dawn of data-driven clinical decisions would appear to be innocuous progress in the healthcare field, but these innovations could easily trigger a future of data discrimination in private healthcare.

II. ANALYSIS

An analysis of United States' health privacy and discrimination laws reveals unacceptable gaps in protection.

A. HIPAA and HITECH

HIPAA, enacted in 1996, protects individually identifiable health information garnered by "covered entities" from being shared with or disclosed to other entities.¹²⁶ Health information is given a relatively broad interpretation under HIPAA, and includes information that relates to an individual's past, present, or future medical condition, as well as the

1522365618 [https://perma.cc/N3CQ-6MXC]. Not to be outdone, online retail giant Amazon also took its first steps into the healthcare industry, scooping up online drugstore PillPack for close to one billion dollars. Mark Brohan, *Amazon Drops Hints on Why It Bought PillPack*, DIGITAL COM. 360 (July 30, 2018), https://www.digitalcommerce360.com/2018/07/30/amazon-drops-hints-on-why-it-bought-pillpack [https://perma.cc/NT8X-VSP7].

¹²⁵ Bernard Marr, *How Big Data Is Changing Healthcare*, FORBES (Apr. 21, 2015, 10:50 AM), https://www.forbes.com/sites/bernardmarr/2015/04/21/how-big-data-is-changing-healthcare/#1f6a10ee2873 [https://perma.cc/F8YT-KWYT]. See generally PITTSBURGH HEALTH DATA ALL., https://healthdataalliance.com [https://perma.cc/Y3N5-C5VU]. The Pittsburgh Health Data Alliance recently partnered with Amazon Web Services, whose sophisticated computing technology should accelerate research. See Fred Pennic, *Pittsburgh Health Data Alliance, Amazon Partner on Machine Learning Research Sponsorship*, HIT CONSULTANT (Aug. 8, 2019), https://hitconsultant.net/2019/08/08/pittsburgh-health-data-alliance-amazon-partner-on-machine-learning-research-sponsorship/#.Xa3nMOhKg2w [https://perma.cc/78F9-PP7A].

¹²⁶ 42 U.S.C. § 1320d-6(a) (2018).

care they have received.¹²⁷ If this information is connected to identifiers, like names, social security numbers, dates of birth or addresses, or if there is a “reasonable basis to believe” that the information can be linked to an individual, the information becomes protected under HIPAA.¹²⁸ Covered entities and other organizations subject to HIPAA may disclose information to business associates and other parties pursuant to certain de-identification requirements.¹²⁹ Under the 1996 provisions, covered entities include health plans (including insurers), healthcare clearinghouses,¹³⁰ and healthcare providers.¹³¹ The HITECH amendment expanded HIPAA to include certain activities of “business associates” of covered entities as well, the definition of which includes data aggregation institutions.¹³²

On the surface, HIPAA and HITECH (collectively, the Acts) appear to solve the problem of data discrimination by requiring data to be anonymized and by bringing data aggregation services under the umbrella of “business associates.” Additionally, the ability to share de-identified information is important to medical research, and the Acts

¹²⁷ “[H]ealth information is information, including demographic information, which relates to: the individual’s past, present, or future physical or mental health or condition, the provision of health care to the individual,” and their medical payment information. U.S. DEP’T HEALTH & HUM. SERVS., GUIDANCE REGARDING METHODS FOR DE-IDENTIFICATION OF PROTECTED HEALTH INFORMATION IN ACCORDANCE WITH THE HEALTH INSURANCE PORTABILITY AND ACCOUNTABILITY ACT (HIPAA) PRIVACY RULE 4 (2012); *see also* 45 C.F.R. § 160.103 (2019).

¹²⁸ 45 C.F.R. §§ 164.514(a), 164.502(a). 45 C.F.R. § 164.514(b)(2) lists the impermissible identifiers.

¹²⁹ 45 C.F.R. §§ 164.502(d), 164.502(e).

¹³⁰ Healthcare clearinghouses are entities that receive data from healthcare institutions and make it compliant with HIPAA. *See* Jonah Comstock, *Reintroduced Bill Could Give Data Brokers More Options Under HIPAA*, HEALTHCARE FIN. (Dec. 26, 2017), <https://www.healthcarefinancenews.com/news/reintroduced-bill-could-give-data-brokers-more-options-under-hipaa> [<https://perma.cc/BP2B-7V6V>].

¹³¹ 45 C.F.R. § 160.103; *see also* *Covered Entities and Business Associates*, U.S. DEP’T HEALTH & HUM. SERVS., <https://www.hhs.gov/hipaa/for-professionals/covered-entities/index.html> [<https://perma.cc/X48Q-BFFV>].

¹³² 45 C.F.R. § 164.302. In paraphrasing its definition from 45 C.F.R. 160.103, HHS describes a “business associate” as: “a person or entity that performs certain functions or activities that involve the use or disclosure of protected health information on behalf of, or provides services to, a covered entity.” U.S. DEP’T HEALTH & HUM. SERVS., OCR HIPAA PRIVACY: BUSINESS ASSOCIATES 1 (2003). Business associates include “legal; actuarial; accounting; consulting; data aggregation; management; administrative; accreditation; and financial” institutions providing services to the covered entity. *Id.* at 2; *see also* Hiller, *supra* note 14, at 283.

appear to preserve the balance between research and privacy.¹³³ However, the Acts ultimately fail to protect against data discrimination because data aggregators are only limited by the Acts when they perform work under contract with a covered entity, leaving the rest of their activity unregulated.

In other words, the Acts do not serve to protect the data aggregated by data brokers outside of their work as business associates. While there are deeper issues regarding HIPAA-permitted contractual provisions between covered entities and business associates that may allow for circumvention of laws regarding protected health information,¹³⁴ the dominant issue here is that consumer data generated outside of the healthcare context does not fall under HIPAA's or HITECH's purview.¹³⁵ As noted above, health information is only protected when it is attached to identifiers¹³⁶ and when it becomes part of a covered entity's records.¹³⁷ Business associates are only bound by the Acts when they are acting pursuant to their relationship with a covered entity.¹³⁸ Outside of this

¹³³ See Standards for Privacy of Individually Identifiable Health Information, 67 Fed. Reg 53,182, 53,226 (Aug. 14, 2002) (codified at C.F.R. pts. 160 & 164) ("Patient privacy must be balanced against other public goods, such as research and the risk of compromising such research projects if researchers could not continue to use such data."); see also *id.* at 53,232 ("The Privacy Rule . . . strives to balance the need to protect individuals' identities with the need to allow de-identified databases to be useful."). *But see* Sharona Hoffman & Andy Podgurski, *Balancing Privacy, Autonomy, and Scientific Needs in Electronic Health Records Research*, 65 SMU L. REV. 85, 123 (2012) (highlighting the fact that many researchers have found that HIPAA "severely hinder[s] research").

¹³⁴ HIPAA allows covered entities to disclose protected health information to business associates as long as the business associate is helping the covered entity carry out its healthcare functions. Business associates cannot employ this information for independent use, "except as needed for the proper management and administration of the business associate." OCR HIPAA PRIVACY, *supra* note 132, at 1. For more information on this topic, see Reece Hirsch & Heather Deixler, *HIPAA Business Associates and Health-Care Big Data: Big Promise, Little Guidance*, BLOOMBERG BNA (Feb. 17, 2014), https://www.morganlewis.com/-/media/files/publication/outside-publication/article/bna_hipaaandbigdata_17feb14.ashx [<https://perma.cc/UWP5-6Y5K>] (identifying the "management and administration" language as ambiguous and calling for further guidance).

¹³⁵ SARASOHN-KAHN, *supra* note 19, at 9.

¹³⁶ See *supra* note 127 and accompanying text.

¹³⁷ SARASOHN-KAHN, *supra* note 19, at 7.

¹³⁸ See Terry, *supra* note 28, at 194 ("If they are not determined to be health plans, health care clearinghouses, or health care providers conducting certain electronic transactions, and they are

context, data brokers remain free to collect, analyze, and sell “medically inflected” data points taken from social media, web trackers, online transactions, mobile phones, and more.¹³⁹ No law prevents health insurers, who are covered entities under HIPAA, from purchasing this data directly from data brokers.¹⁴⁰ For example, an insurer is not restricted from buying a list of people who, according to the data broker, suffer from depression or erectile dysfunction.¹⁴¹ HIPAA’s “original sin,” therefore, is that it introduced a privacy scheme structured around regulating certain custodians of health data, rather than protecting the health data itself.¹⁴²

B. *Affordable Care Act*

The Affordable Care Act (ACA) was passed in 2010 and prohibits health insurers from discriminating against people with pre-existing conditions.¹⁴³ The ACA also protects against discrimination on the basis of most of the familiar Title VII Civil Rights Act qualities, such as race, color, national origin, sex, age, or disability.¹⁴⁴ This closed many of the loop-holes insurance companies were used to working through. For example, insurance providers often considered just about any test or

not acting on behalf of, or providing a service to, a HIPAA covered entity, they are not subject to the HIPAA standards for covered entities and business associates.”).

¹³⁹ *Id.* at 179.

¹⁴⁰ See Allen, *supra* note 6.

¹⁴¹ Kelsey L. Zottnick, Note, *Secondary Data: A Primary Concern*, 18 VAND. J. ENT. & TECH. L. 193, 200 (2015).

¹⁴² Terry, *supra* note 28, at 164.

¹⁴³ 42 U.S.C. § 18001 (2018).

¹⁴⁴ *Id.* § 18116(a). Subsection (c) of this statute gives the Secretary of HHS the authority to promulgate regulations related to that section. *Id.* § 18116(c). In 2016, HHS issued a rule redefining discrimination “on the basis of sex” in order to include protections against discriminations on the basis of gender identity and the termination of pregnancy. Press Release, U.S. Dep’t of Health & Human Servs., HHS Proposes to Revise ACA Section 1557 Rule to Enforce Civil Rights in Healthcare, Conform to Law, and Eliminate Billions in Unnecessary Costs (May 24, 2019), <https://www.hhs.gov/about/news/2019/05/24/hhs-proposes-to-revise-aca-section-1557-rule.html> [<https://perma.cc/8AHJ-XLJB>]. In December 2016, the United States District Court for the Northern District of Texas enjoined prohibitions of this kind on a nationwide basis, finding that the new regulation contradicted existing law and exceeded the HHS’s statutory authority. See *Franciscan All., Inc. v. Burwell*, 227 F. Supp. 3d 660 (N.D. Tex. 2016).

diagnosis a pre-existing condition, with some of the most common being AIDS, alcoholism, Alzheimer's, arthritis, obesity, and even pregnancy.¹⁴⁵ If an otherwise healthy person suffered from even one of these conditions, they could be charged higher premiums or denied coverage.¹⁴⁶ While the ACA corrected this problem to a certain extent, it still remains possible for insurance companies to discriminate against those with pre-existing conditions.¹⁴⁷ Insurers do this through "risk classification by design," which is when companies try to force high-risk patients to seek different plans either by setting steep prices for the pharmaceuticals the patients need, or by manipulating plans to not include the best doctors for the patient's illness.¹⁴⁸ The ACA mitigates against this threat in some ways, such as by requiring plans to cover essential benefits,¹⁴⁹ but the Center for Health Law and Policy Innovation at Harvard Law claims that insurers still use these subtle tactics to discriminate.¹⁵⁰ If the ACA has failed to protect those with pre-existing conditions, it can be expected to be equally incapable of protecting those who have not yet developed a pre-existing condition, but whose unreliable online data profile might point to an unhealthy future.

¹⁴⁵ Kimberley Amadeo, *Obamacare Pre-Existing Conditions*, BALANCE, <https://www.thebalance.com/obamacare-pre-existing-conditions-3306072> [<https://perma.cc/D6CR-JNQS>] (last updated June 25, 2019).

¹⁴⁶ *Id.*

¹⁴⁷ John V. Jacobi et al., *Health Insurer Market Behavior After the Affordable Care Act: Assessing the Need for Monitoring, Targeted Enforcement, and Regulatory Reform*, 120 PENN. ST. L. REV. 109, 169–73 (2015).

¹⁴⁸ *Id.*

¹⁴⁹ *Id.* at 169–71.

¹⁵⁰ Allen, *supra* note 6; see also Vidushi Saxena, *Campaign Looks to Address Discrimination in Health Care*, BADGER HERALD (Sept. 8, 2016), <https://badgerherald.com/news/2016/09/08/campaign-looks-to-address-discrimination-in-health-care> [<https://perma.cc/N2TT-QUC3>]. In addition, the Trump Administration has sought to undermine the ACA repeatedly. See, e.g., Timothy S. Jost, *The Affordable Care Act Under the Trump Administration*, COMMONWEALTH FUND (Aug. 30, 2018), <https://www.commonwealthfund.org/blog/2018/affordable-care-act-under-trump-administration> [<https://perma.cc/JP6Q-RJPY>].

C. GINA

1. The Rationale of GINA

GINA, which was passed into law in 2008, has a singular purpose: “[t]o prohibit discrimination on the basis of genetic information with respect to health insurance and employment.”¹⁵¹ One of the key reasons for passing GINA was to allay concerns that advances in the use of genetic information in the field of medicine would lead to discrimination on that same basis in the health insurance sector.¹⁵² Generally speaking, Americans are aware of the prevalence of deoxyribonucleic acid—or DNA—in today’s society, and its status as the “blueprint” of every human being.¹⁵³ DNA and genetic information can indicate whether a person is susceptible to a genetic condition, how likely they are to develop that condition, how they will react to certain drugs, and more.¹⁵⁴

GINA serves to retain the benefits of individuals’ and researchers’ utilization of genetic testing, while keeping this information safe from potential discriminatory practice.¹⁵⁵ Title I of GINA, as an amendment to several existing acts, serves to outright prevent health insurers from discriminating on the basis of genetic information, from requesting or requiring genetic information, and from obtaining such information for certain purposes.¹⁵⁶ GINA carved out a research exception, which permits

¹⁵¹ Genetic Information Nondiscrimination Act of 2008, Pub. L. No. 110-233, 122 Stat. 881, 881. While GINA’s considerations regarding discrimination in employment are important, they are not relevant here. GINA was passed with unanimous support in the Senate (95-0) and nearly unanimous support in the House (414-1). See Perry W. Payne, Jr., *Genetic Information Nondiscrimination Act of 2008: The Federal Answer for Genetic Discrimination*, 5 J. HEALTH & BIOMEDICAL L. 33, 46 (2009).

¹⁵² Genetic Information Nondiscrimination Act § 2.

¹⁵³ Payne, Jr., *supra* note 151, at 33–34 (“The best and worst aspects of one’s heredity may be passed on to the next generation via DNA.”).

¹⁵⁴ *What Is Genetic Testing?*, U.S. NAT’L LIBR. MED.: GENETICS HOME REFERENCE (Jan. 7, 2020), <https://ghr.nlm.nih.gov/primer/testing/geneticstesting> [<https://perma.cc/SWF7-LRUR>]; Payne, Jr., *supra* note 151, at 35.

¹⁵⁵ Genetic Information Nondiscrimination Act § 2.

¹⁵⁶ *Id.* §§ 101–106; Jessica L. Roberts, *Preempting Discrimination: Lessons from the Genetic Information Nondiscrimination Act*, 63 VAND. L. REV. 439, 451–52 (2010). One exception under GINA allows insurers to request the minimum amount of genetic information necessary to make a decision on whether they will pay for a test, treatment, or procedure. Once they have this

health insurers engaged in research to request, but not require, an individual to participate in genetic testing, subject to the Common Rule¹⁵⁷ and other federal and state laws and regulations.¹⁵⁸

Due to DNA's exceptional nature, and because it can allow analysts to draw conclusions about test subjects' health as well as their family members' health, GINA's definition of genetic information is broad.¹⁵⁹ Genetic information includes information about: (1) an individual's genetic tests, (2) the genetic tests of family members, and (3) the manifestation of a disease or disorder in family members of such individual.¹⁶⁰ GINA applies to "most" health insurers.¹⁶¹

Interestingly, when GINA was passed, there was little evidence of actual discrimination among insurers, and, although many states already had laws preventing discrimination on the basis of genetic information, there were no reported cases pursued under those laws.¹⁶² Despite this, according to Jessica Roberts, Congress alluded to two major justifications

information, GINA still prevents them from using it to discriminate. *GINA and You Information Sheet*, GINAHELP.ORG (June 2010), <http://www.ginahelp.org/GINAhelp.pdf> [<https://perma.cc/SD9C-9AWK>].

¹⁵⁷ The Common Rule was

established in 1991 to create uniform protection for human research subjects. The Rule reaches certain research "conducted, supported or otherwise subject to regulation" by fifteen federal agencies . . . as well as research by any institution claiming federal-wide assurance for the protection of human subjects by adopting the standards and rules articulated in the Common Rule.

William McGeeveran et al., *Deidentification and Reidentification in Returning Individual Findings from Biobank and Secondary Research*, 13 MINN. J.L. SCI. & TECH. 485, 490 (2012) (internal citations omitted). For studies that pose greater than a minimal risk to human subjects, an Institutional Review Board is required to review. *See id.* at 492. The Institutional Review Board must evaluate any ethical concerns before approving the research. *Id.* at 493. Two of the factors the Board will consider include whether the study's data monitoring is sufficient and whether it contemplates the protection of the confidentiality of the data. *Id.*; 45 C.F.R. § 46.111(a)(6)–(7) (2019).

¹⁵⁸ Caryn M. Silverman, *GINA's Impact on Clinical Research Trials*, LAW 360 (May 1, 2009, 12:00 AM), <https://www.law360.com/articles/99617/gina-s-impact-on-clinical-research-trials> [<https://perma.cc/7DS8-C3EG>].

¹⁵⁹ Payne, Jr., *supra* note 151, at 35.

¹⁶⁰ 42 U.S.C. § 2000ff(4)(A)(i)–(iii) (2018). *See generally* Sagit Ziskind, *The Genetic Information Nondiscrimination Act: A New Look at an Old Problem*, 35 RUTGERS COMPUTER & TECH. L.J. 163 (2009) (arguing that GINA helped to cure major deficiencies in the genetic legislative scheme).

¹⁶¹ *Gina and You Information Sheet*, *supra* note 156, at 2.

¹⁶² Roberts, *supra* note 156, at 466–73.

for passing GINA: (1) a research justification; and (2) an antidiscrimination justification.¹⁶³ In short, not only did Congress believe that fear of discrimination prevents people from pursuing genetic testing for various research or health purposes, but also that discrimination based on genetics—which are unchosen, unchangeable, and inherited—is fundamentally unfair.¹⁶⁴ At the heart of this second argument is the idea that Americans should not have to fear their own immutable genes will become a weapon of scrutiny and discrimination to be used against them.¹⁶⁵ This aligns with President Bush’s comments in June 2001, when he called genetic discrimination “little more than medical speculation,” and a “violat[ion of] our country’s belief in equal treatment and individual merit.”¹⁶⁶

There are arguments on both sides as to whether genetic information should be given such preferential treatment, especially when the lines drawn between genetic information, other medical tests, and an individual’s habits are so blurry.¹⁶⁷ Common medical tests, like those for cholesterol or blood pressure, are not illegal to request or require even though the results may be hopelessly tied to genetics.¹⁶⁸ Lifestyle choices, too, are intertwined with genetics.¹⁶⁹ The fundamental idea that we should be viewed as products of our choices—to smoke cigarettes, eat fast food, or drink five cups of coffee a day—rather than as products of our

¹⁶³ *Id.* at 471.

¹⁶⁴ *Id.* at 471–79.

¹⁶⁵ *Id.* at 478.

¹⁶⁶ President George W. Bush, Office of the Press Secretary, Radio Address by the President to the Nation (June 23, 2001), <https://www.c-span.org/video/?164891-1/presidents-radio-address> [<https://perma.cc/2F9R-G7TY>]; see also Michelle Meyer, *Re-Identification Is Not the Problem. The Delusion of De-Identification Is.*, HARV. L.: BILL OF HEALTH (May 22, 2013), <http://blog.petrieflom.law.harvard.edu/2013/05/22/re-identification-is-not-the-problem-the-delusion-of-de-identification-is-re-identification-symposium-2> [<https://perma.cc/H7FQ-F62Z>] (“My genome can tell you probabilistic information about meBut it won’t tell you how my phenotypes have developed over my lifetime or how they may have been altered My genomic data is of little research value without me”).

¹⁶⁷ See generally Deborah Hellman, *What Makes Genetic Discrimination Exceptional?*, 29 AM. J.L. & MED. 77 (2003).

¹⁶⁸ *Id.* at 80–81.

¹⁶⁹ *Id.* at 87–88.

genetic makeup is undermined when one considers that genetics greatly influence our cravings and desires.¹⁷⁰

2. How Genetic Information Is Analogous to Unreliable Health-Related Data

While genetic information is useful in making determinations regarding future health, it is only one of many factors. Other determinants include social and economic environments, lifestyle and habits (such as eating, drinking, and exercising), and relationships with friends and family.¹⁷¹ The relationship between all of these factors and their final impact on an individual's health is still being studied.¹⁷² Not surprisingly, some of these same factors are used by providers when determining premiums.¹⁷³ This risk assessment, at least in part, may involve conclusions made about a prospective insurance candidate's available health-related data pulled from the resources examined in Part II.¹⁷⁴ But in light of the legislative scheme surrounding healthcare privacy, is this a logical or fair use of that data? Does it seem appropriate for an insurance company to make policy decisions based on health inferences drawn by data brokers?¹⁷⁵ If Congress is willing to protect genetic information in part because genetic information is unchosen and genetic tests are "medical speculation" then they should be equally willing to protect consumers from data brokers' assumptions, which are frequently

¹⁷⁰ *Id.*; see also Alvin Powell, *Study Shows Coffee Drinking Habits Shaped by DNA Variations*, SCITECHDAILY (Oct. 7, 2014), <https://scitechdaily.com/study-shows-coffee-drinking-habits-shaped-dna-variations> [<https://perma.cc/CXZ4-3MU4>]. For an argument that GINA was the product of popular rather than logical support, and that genetic discrimination is more of a perceived than actual threat, see Jeffrey S. Morrow, *Insuring Fairness: The Popular Creation of Genetic Antidiscrimination*, 98 GEO. L.J. 215 (2009).

¹⁷¹ *Health Impact Assessment (HIA)*, WORLD HEALTH ORG., <https://www.who.int/hia/evidence/doh/en> [<https://perma.cc/S2CZ-VX8Y>].

¹⁷² *U.S. Government Seeking 1 Million People for Study of DNA, Health Habits*, CBS NEWS (May 1, 2018, 5:14 PM), <https://www.cbsnews.com/news/u-s-government-seeking-1-million-people-for-study-of-dna-health-habits> [<https://perma.cc/6QTA-C62Z>].

¹⁷³ Brian O'Connell, *What Are Insurance Premiums and How Much Do They Cost?*, STREET (Mar. 5, 2019, 10:06 AM), <https://www.thestreet.com/personal-finance/what-are-insurance-premiums-14875700> [<https://perma.cc/P9MX-VUGA>].

¹⁷⁴ See *supra* Section I.A.1; see also, Allen, *supra* note 6.

¹⁷⁵ See *supra* Section I.A.1; see also, Allen, *supra* note 6.

unchangeable and often downright incorrect.¹⁷⁶ Further, if HIPAA, the ACA, and GINA all point to a legislative scheme that values fairness and privacy while detesting speculation, then there should be additional protections against the use of data analytics for discriminatory purposes.

Genetic information shares some characteristics with the consumer data profiles compiled by data brokers and shared with insurers. These data profiles will be referred to as “digital DNA.” There are two central reasons, discussed below, why digital DNA is similar to genetic information and equally deserving of protection: (1) digital DNA is often unalterable and incorrect; and (2) if used by insurers against consumers, the stalking omnipresence of digital DNA can negatively influence the way people interact with the digital world.

a. Incorrect Predictors

While the American public generally perceives genetic tests as accurate, they are not always as correct as they may appear.¹⁷⁷ The FDA, through the last two administrations, has shown uncertainty in how to regulate genetic testing.¹⁷⁸ Many tests remain outside the purview of FDA regulations and are not independently reviewed for accuracy.¹⁷⁹ A Dutch study found that genetic tests often mislead people about their future health.¹⁸⁰ The study found that companies “predicted conflicting risks for some diseases,” and “in some cases arrived at medical predictions that were no better than flipping a coin.”¹⁸¹ In 2008, a genetic test misdiagnosed women with ovarian cancer, leading to the test being

¹⁷⁶ See Miller, *supra* note 88; Leetaru, *supra* note 88; see also Ravindranath, *supra* note 30 (“The data that brokers peddle to healthcare providers is often wrong.”).

¹⁷⁷ Beth Daley, *Genetic Tests Are Everywhere, but How Reliable Are They?*, BOS. GLOBE (Aug. 26, 2017, 7:00 AM), <https://www.bostonglobe.com/ideas/2017/08/26/genetics-for-everyone/9yLzbnyXoRvIAkul74DwDM/story.html> [<https://perma.cc/R58Q-FHS4>] (“Some tests that predict the likelihood a young pregnant woman will have a child with a genetic condition such as Down syndrome may only be correct only 60 percent of the time.”).

¹⁷⁸ *Id.*

¹⁷⁹ *Id.*

¹⁸⁰ Ian Sample, *Genetic Tests Flawed and Inaccurate, Say Dutch Scientists*, GUARDIAN (May 30, 2011, 6:00 PM), <https://www.theguardian.com/science/2011/may/30/genetics-tests-flawed-dutch-scientists> [<https://perma.cc/4SWV-UZ7R>].

¹⁸¹ *Id.*

pulled from the market, but not before some of the women had their ovaries removed.¹⁸²

Similarly, digital DNA can be just as, if not significantly more inaccurate than unregulated genetic testing. Curious journalists seeking to shed light on their own data-broker held consumer profiles have repeatedly discovered a wealth of incorrect information.¹⁸³ One journalist claims that seventy-eight percent of the categories a data broker called Oracle placed him in bore no resemblance to his life or interests.¹⁸⁴ Another journalist reported that fifty percent of the data she requested was wrong or out of date, including her zip code and occupation.¹⁸⁵ Another journalist documented the story of a deceased man who was still receiving health insurance solicitations fourteen years after his death.¹⁸⁶ Apparently, the insurers had purchased his name and address from a data broker.¹⁸⁷ In another more troubling illustration, a woman prescribed Prozac for an off label use was later denied coverage by an insurer that assumed the woman had a history of depression.¹⁸⁸

Because consumer information is bought, sold, and exchanged between data brokers and other entities, a single incorrect data point can have “a cascade effect through the cyberspace, creating erroneous files in potentially thousands of databases.”¹⁸⁹ One scholar has aptly analogized the data industry to the child’s game of “telephone,” where the original speaker’s whispered secret has changed significantly by the time it is said aloud at the end of the line.¹⁹⁰

It is not difficult to see how information garnered from a data broker’s sources can be misleading or incorrect. For example, a data broker or insurer that assesses an individual’s credit card purchases will

¹⁸² Daley, *supra* note 177.

¹⁸³ See Miller, *supra* note 88; Leetaru, *supra* note 88.

¹⁸⁴ Leetaru, *supra* note 88.

¹⁸⁵ Miller, *supra* note 88.

¹⁸⁶ David Lazarus, *Column: Why Health Insurance Companies Will Solicit You Even After You Die*, L.A. TIMES (Oct. 18, 2016, 3:00 AM), <https://www.latimes.com/business/lazarus/la-fi-lazarus-medicare-mailings-20161018-snap-story.html> [<https://perma.cc/635Q-P33A>].

¹⁸⁷ *Id.*

¹⁸⁸ Martha Tucker Ayres, Comment, *Confidentiality and Disclosure of Health Information in Arkansas*, 64 ARK. L. REV. 969, 972–73 (2011).

¹⁸⁹ Lazarus, *supra* note 186.

¹⁹⁰ Schmitz, *supra* note 46, at 1428–29.

not be able to consider that person's cash transactions as well.¹⁹¹ If someone were to buy all their fast-food with a credit card, and all their fruits and vegetables with cash, they may appear far less healthy to a data broker than they are in reality.¹⁹² Likewise, not all credit card purchases are made by the cardholder.¹⁹³ Additionally, social media interactions are far from accurate representations of a person's actual mental state.¹⁹⁴ Recently, BuzzFeed podcast host Tracy Clayton asked her Twitter followers to post an old picture of themselves where they appeared happy but were actually suffering through a difficult time in their lives.¹⁹⁵ The tweet quickly went viral and garnered hundreds of responses.¹⁹⁶ When asked what inspired the tweet, Clayton told a reporter that she was motivated by the idea that insurance companies evaluate claims by

¹⁹¹ Kelly Dilworth, *6 Ways to Outsmart Data Brokers*, CREDITCARDS.COM (Sept. 16, 2014), <https://www.creditcards.com/credit-card-news/six-ways-outsmart-data-broker-collection-privacy-1282.php> [<https://perma.cc/8PC4-9GX4>] (recommending that sensitive purchases be made with cash, and noting that combining cash with a loyalty card will make the purchase traceable).

¹⁹² See generally Eric Rosenbaum, *The Cashless Society Myth: PayPal, Square, and Bitcoin Have Not Stopped Cash from Being a Growth Business*, CNBC (Aug. 6, 2018, 9:32 AM), <https://www.cnbc.com/2018/08/06/spike-the-dollars-obit-cash-is-still-a-growth-business.html> [<https://perma.cc/3AT4-DJQD>] (reporting that thirty-one percent of transactions in the U.S. are cash—which is more than debit, credit, electronic, or check transactions).

¹⁹³ Lorie Konish, *Your Child Could Be Running Up Your Credit Card Bill. You Can Fix That*, CNBC (Oct. 25, 2018, 11:00 AM), <https://www.cnbc.com/2018/10/25/your-child-could-be-running-up-your-credit-card-bill-you-can-fix-that.html> [<https://perma.cc/V3NZ-ARC7>] (reporting that twenty-nine percent of parents with children under eighteen have had their children use their card without permission).

¹⁹⁴ See, e.g., Grace Dent, *Social Media Is Full of Sad, Lonely People Pretending They're OK and Perfectly Fine Attention-Seekers Pretending to Be Sad*, INDEPENDENT (Mar. 6, 2017, 3:45 PM), <https://www.independent.co.uk/voices/facebook-instagram-twitter-social-media-makes-sad-lonely-attention-seekers-a7614396.html> [<https://perma.cc/NT23-Y56F>]. After doing some digging through Instagram's settings, I discovered a link to a list of my "Ad Interests," as perceived by Instagram. Of the many shocking products (Energy Drinks) and hobbies (Deer Hunting) that I did not know I was interested in, one stood out: "Boxing Training." I have never pursued or researched boxing training on my own, but a friend (who is in fact interested in boxing) sends boxing videos to me on Instagram more often than I would like. Instagram was clearly conflating our interests.

¹⁹⁵ Leah Fessler, *The Ultimate Proof that Looking Happy on Social Media Often Masks Real Pain*, QUARTZ (July 13, 2018), <https://qz.com/1326806/the-ultimate-proof-that-looking-happy-on-social-media-often-masks-real-pain> [<https://perma.cc/Q8FL-5LEW>].

¹⁹⁶ *Id.*

looking at social media, and how she thought that was “irresponsible and risky.”¹⁹⁷

To compound the issue, consumers often are not allowed to access, correct, or remove this information or the inferences data brokers have drawn from it.¹⁹⁸ Further, brokers may not always allow consumers to opt out of data sharing,¹⁹⁹ and if the broker does allow an opt-out, the process may require the consumer to make phone calls, send faxes, or use the postal service.²⁰⁰ Even the minority of people who spend time and effort opting out of data sharing may not realize these opt-outs are limited.²⁰¹ Many brokers use automated processes, so if the broker reobtains information from another source, the person opting out could be re-entered into the broker’s system.²⁰² In this sense, digital DNA is much like genetic information—unchosen by consumers, potentially incorrect, and effectively immutable.

b. Fear of Discrimination

One of the justifications behind the passing of GINA was Congress’s concern that Americans would not participate in any kind of genetic testing—for research or otherwise—due to a fear of discrimination.²⁰³ Representative Louise Slaughter described the dilemma of undergoing genetic testing as “a Hobson’s choice between learning vital health information and risking [one’s] health insurance.”²⁰⁴ It is easy to imagine a similar chilling effect on the American public when the revelation sets in that every decision or move one makes is recorded and compiled by some faceless, digital villain.²⁰⁵ Individuals might face the difficult choice

¹⁹⁷ *Id.*

¹⁹⁸ Grauer, *supra* note 36.

¹⁹⁹ Schmitz, *supra* note 46, at 1422–23.

²⁰⁰ Grauer, *supra* note 36.

²⁰¹ Schmitz, *supra* note 46, at 1423.

²⁰² Grauer, *supra* note 36.

²⁰³ Roberts, *supra* note 156, at 464–65.

²⁰⁴ *Id.* at 472 (quoting 144 CONG. REC. H3,262 (daily ed. May 14, 1998)).

²⁰⁵ Jay Stanley, *The Potential Chilling Effects of Big Data*, ACLU (Apr. 30, 2012, 11:46 AM), <https://www.aclu.org/blog/privacy-technology/consumer-privacy/potential-chilling-effects-big-data> [<https://perma.cc/P4ML-RRQA>] (“Over time, as the ramifications of big data analytics sink in, people will likely become much more conscious of the ways they’re being tracked, and the chilling effects on all sorts of behaviors could become considerable.”).

of researching their potential illness or remaining uninformed to avoid any chance of discrimination.²⁰⁶ Americans would soon become paranoid, “quivering, neurotic beings . . . in a psychologically oppressive world”²⁰⁷ This phenomenon, dubbed “social cooling,” could also suppress the public’s desire to take risks or exercise fundamental American rights such as free speech.²⁰⁸ Consumer privacy and security concerns have already been found to deter many Americans from engaging in online transactions or posting on social networks.²⁰⁹ As the age of the “Internet of Things”²¹⁰ continues to progress, more Americans may opt to remove themselves from the digital world, reducing economic activity and hampering the exchange of ideas.²¹¹

III. PROPOSAL: EXPAND DISCRIMINATION PROTECTION BEFORE PRIVACY PROTECTION

A. *GDPR, CCPA, and the Privacy Approach*

In 2016, the European Union passed the General Data Protection Regulation (GDPR), seeking to “harmonise the protection of fundamental rights and freedoms of natural persons in respect of processing activities and to ensure the free flow of personal data between Member States.”²¹² Considered one of the most contested laws in

²⁰⁶ Frank Pasquale & Tara Adams Ragone, *Protecting Health Privacy in an Era of Big Data Processing and Cloud Computing*, 17 STAN. TECH. L. REV. 595, 652 (2014).

²⁰⁷ Stanley, *supra* note 205.

²⁰⁸ SOCIAL COOLING, <https://www.socialcooling.com> [<https://perma.cc/JAG3-8JTK>] [hereinafter SOCIAL COOLING].

²⁰⁹ Rafi Goldberg, *Lack of Trust in Internet Privacy and Security May Deter Economic and Other Online Activities*, U.S. DEP’T COM.: NAT’L TELECOMM. & INFO. ADMIN. (May 13, 2016), <https://www.ntia.doc.gov/blog/2016/lack-trust-internet-privacy-and-security-may-deter-economic-and-other-online-activities> [<https://perma.cc/6J4B-JT4W>].

²¹⁰ *See supra* note 1.

²¹¹ *See* SOCIAL COOLING, *supra* note 208

²¹² OFFICIAL JOURNAL OF THE EUR. UNION, REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL OF 27 APRIL 2016: GENERAL DATA PROTECTION REGULATION, at cl. (3). GDPR became effective May 25, 2018. Mauricio F. Paez & Undine von Diemar, *GDPR: Understanding the New Landscape*, CORP. COUNSEL BUS. J. (April 2, 2018), <https://ccbjournal.com/articles/gdpr-understanding-new-landscape> [<https://perma.cc/BBB5-8TH7>].

European Union history, GDPR is the product of years of heated negotiations and thousands of amendments.²¹³ Some of the highlights of GDPR include stronger consumer consent requirements, a consumer right to access data and discover what it is being used for, a consumer right to erase data or cease its dissemination, and stiff penalties for non-compliant businesses.²¹⁴ Alone, GDPR has had a significant impact on American companies doing business in the European Union, and even those that are not.²¹⁵ However, in 2018, on the heels of GDPR's effective date, the California Consumer Privacy Act (CCPA) was passed, granting Californians a set of new controls over their personal data.²¹⁶ While the two bodies of law are not identical, drafters of the CCPA looked to GDPR for some of its data protection concepts.²¹⁷ Though no new federal privacy laws have been passed yet, discourse on the topic is growing, with the Commerce Department going so far as to issue a request for comments on new data privacy regulations.²¹⁸ Other state governments, in the absence of federal guidance, are also becoming increasingly active

²¹³ Julia Powles, *The G.D.P.R., Europe's New Privacy Law, and the Future of the Global Data Economy*, NEW YORKER (May 25, 2018), <https://www.newyorker.com/tech/annals-of-technology/the-gdpr-europes-new-privacy-law-and-the-future-of-the-global-data-economy> [https://perma.cc/BAJ3-9YNC].

²¹⁴ *What Is GDPR, The EU's New Data Protection Law?*, GDPR.EU, <https://gdpr.eu/what-is-gdpr> [https://perma.cc/5VT4-WBK5].

²¹⁵ Yaki Faitelson, *Yes, the GDPR Will Affect Your U.S.—Based Business*, FORBES (Dec. 4, 2017, 8:30 AM), <https://www.forbes.com/sites/forbestechcouncil/2017/12/04/yes-the-gdpr-will-affect-your-u-s-based-business/#6b3290e6ff26> [https://perma.cc/R7TC-8KD5].

²¹⁶ John Tomaszewski & Edward Murphree, *The CA Consumer Privacy Act: The NEW Elephant in the Room*, JD SUPRA (Jan. 4, 2019), <https://www.jdsupra.com/legalnews/the-ca-consumer-privacy-act-the-new-32726> [https://perma.cc/NC8G-MSHD]; see also Paez & von Diemar, *supra* note 212. The CCPA became effective January of 2020. Tomaszewski & Murphree, *supra*.

²¹⁷ Tomaszewski & Murphree, *supra* note 216. Unlike GDPR, California's bill was introduced just a week before it was passed with unanimous consent. See Issie Lapowsky, *California Unanimously Passes Historic Privacy Bill*, WIRED (June 28, 2018, 5:57 PM), <https://www.wired.com/story/california-unanimously-passes-historic-privacy-bill> [https://perma.cc/F9LJ-J42K].

²¹⁸ Yaki Faitelson, *Data Privacy Disruption in the U.S.*, FORBES (Dec. 12, 2018, 9:30 AM), <https://www.forbes.com/sites/forbestechcouncil/2018/12/12/data-privacy-disruption-in-the-u-s/#10f8e6a815cc> [https://perma.cc/74A8-M8A5].

in working to enact new legislation.²¹⁹ But in the healthcare sector, the question remains: is this the right approach?

1. Potential Flaws in the GDPR and CCPA

The CCPA is stricter than most existing privacy laws in the United States.²²⁰ It provides for a dramatic increase in consumers' access to and control over how their personal information is sold, used, or disclosed.²²¹ Notable improvements include: a right to access personal information collected by businesses, a right to erase personal information, a right to opt-out of future sales of their information, and a requirement that businesses make their privacy policies more transparent to users.²²² The CCPA does not apply to HIPAA-regulated personal health information, but it does apply to other data held by for-profit health organizations.²²³ This means that the information used by health insurers to draw inferences about individuals will come under the CCPA's purview.²²⁴ Due to the fact that it is extremely burdensome for companies to maintain different privacy standards within each state they operate, the CCPA could end up affecting more than just Californians.²²⁵

²¹⁹ Andrew Burt, *States Are Leading the Way on Data Privacy*, HILL (Aug. 21, 2018, 10:30 AM), <https://thehill.com/opinion/technology/402775-states-are-leading-the-way-on-data-privacy> [<https://perma.cc/HTH6-YDVH>].

²²⁰ *California Consumer Privacy Act of 2018*, GIBSON DUNN (July 12, 2018), <https://www.gibsondunn.com/wp-content/uploads/2018/07/california-consumer-privacy-act-of-2018.pdf> [<https://perma.cc/FEU3-VJA2>].

²²¹ Michael R. Overly, *Crossroads of Cybersecurity and the Law*, CSO ONLINE (Aug. 1, 2018 4:15 AM), <https://www.csoonline.com/article/3294297/regulation/is-californias-consumer-privacy-act-of-2018-going-to-be-gdpr-version-2.html> [<https://perma.cc/F755-ZE5K>].

²²² *Id.*

²²³ Rachel Z. Arndt, *California Law Introduces New Data Concerns for Healthcare Organizations*, MOD. HEALTHCARE (July 2, 2018, 1:00 AM), <https://www.modernhealthcare.com/article/20180702/NEWS/180709994/california-law-introduces-new-data-concerns-for-healthcare-organizations> [<https://perma.cc/84L8-SB52>].

²²⁴ Brandon P. Reilly, *What California's New Consumer Privacy Bill Means for Healthcare*, MANATT (July 19, 2018), <https://www.manatt.com/insights/newsletters/health-update/what-californias-new-consumer-privacy-bill-means> [<https://perma.cc/JR97-DPTZ>].

²²⁵ Carrie Nixon, *The New California Privacy Law*, NIXON L. GRP. (Aug. 31, 2018), <https://www.nixonlawgroup.com/nlg-blog/2018/8/31/the-new-california-privacy-law-why-healthcare-tech-companies-in-every-state-need-to-pay-attention> [<https://perma.cc/ZHB9-BDZ8>].

GDPR and the CCPA both seek to place data control back into the hands of consumers.²²⁶ At first glance, this appears to solve the problem of data discrimination in healthcare by increasing access and transparency. As of the date of this Note, GDPR is just over two years old, and the CCPA has only been in effect for six months, so it remains to be seen whether they will be vaunted as successful pillars of privacy law.²²⁷ However, when viewing the spectrum of potential consequences of these laws, two potentially devastating possibilities loom: (1) consumers take full advantage of their newfound privacy rights, causing a data crunch and hampering business and industry innovation;²²⁸ or (2) consumers fail to exercise control, rendering the laws little more than shells of red tape and leaving enforcement up to government actors.²²⁹

The first possibility would be unsurprising, considering consumers' growing desire to protect their data.²³⁰ A recent, post-GDPR survey found that fifty-five percent of United Kingdom citizens planned to exercise their GDPR rights within a year.²³¹ A separate survey of over seven thousand EU consumers revealed that ninety percent sought direct control over how companies use their data, although only twenty-one

²²⁶ See, e.g., Kassandra Coulombe, *GDPR: Giving Control Back to Consumers*, BOS. DIGITAL (May 2, 2018), <https://www.bostondigital.com/blog/gdpr-giving-control-back-consumers> [<https://perma.cc/7WLT-EJXR>]; Stuart D. Levi, *California Privacy Law: What Companies Should Do to Prepare in 2019*, SKADDEN (Jan. 17, 2019), <https://www.skadden.com/insights/publications/2019/01/2019-insights/california-privacy-law> [<https://perma.cc/2RFF-V2YD>].

²²⁷ Dipayan Ghosh, *How GDPR Will Transform Digital Marketing*, HARV. BUS. REV. (May 21, 2018), <https://hbr.org/2018/05/how-gdpr-will-transform-digital-marketing> [<https://perma.cc/YCC4-TDV6>].

²²⁸ Larry Dobrow, *How Pharma Is Tightening Up Data Privacy and Security*, MM&M (Feb. 5, 2019), <https://www.mmm-online.com/home/channel/features/how-pharma-is-tightening-up-data-privacy-and-security> [<https://perma.cc/L2JS-P8YJ>].

²²⁹ Mindaugas Kiskis, *GDPR Is Eroding Our Privacy, Not Protecting It*, TNW (Aug. 5, 2018), <https://thenextweb.com/contributors/2018/08/05/gdpr-privacy-eroding-bad> [<https://perma.cc/LR4L-MKE3>].

²³⁰ *Watch Out for GDPR-Related Claims—It's Not All About Fines . . .*, WALKER MORRIS (July 25, 2018), <https://www.walkermorris.co.uk/publications/watch-out-for-gdpr-related-claims-its-not-all-about-fines> [<https://perma.cc/QUY4-MYQU>] (“[C]onsumers are becoming increasingly aware of their status as data subjects and emboldened in relation to the exercise of their [personal] data protection rights.”).

²³¹ Anthony Spadafora, *British Consumers Are Already Taking Advantage of GDPR*, IT PRO PORTAL (Aug. 6, 2018), <https://www.itproportal.com/news/british-consumers-are-already-taking-advantage-of-gdpr> [<https://perma.cc/5XEF-VABV>].

percent understood their rights under GDPR.²³² Nevertheless, if these consumers actively inform themselves of the legislation's parameters, companies could face an enormous influx of onerous data requests.²³³ In the United States, under the CCPA or similar consumer-controlled privacy legislation, a high-profile data breach in the healthcare industry could provoke thousands of consumers to opt-out of information sharing.²³⁴ If massive amounts of data evaporate, then the data supply chain becomes unstable,²³⁵ crippling many modern businesses and data initiatives.²³⁶

The second possibility would also be catastrophic, resulting in a privacy scheme untenable for businesses and useless for consumers. One of the pointed criticisms against GDPR is that it could make personal data more vulnerable to bad actors.²³⁷ GDPR's right to access allows a consumer to request a copy of all the data that an organization has collected from them.²³⁸ However, the GDPR does not provide guidance to companies as to how to verify the authenticity of a data request.²³⁹ This means that identity thieves could hack an account and make a request on behalf of an unassuming consumer.²⁴⁰ If this practice were to become prevalent, then consumers would be greatly discouraged from exercising their rights under GDPR. Further, there is no guarantee that consumers will take any deliberate action under GDPR or the CCPA. Recall from the paragraph above that despite European consumers' desire to control their data, only twenty-one percent were actually aware of what their GDPR rights were.²⁴¹ In America, only five percent of people successfully change

²³² Kim Davis, *GDPR: What Matters to Consumers*, DMN (May 28, 2018), <https://www.dmnews.com/retail/article/13034544/gdpr-what-matters-to-consumers> [<https://perma.cc/3ATH-QBPR>].

²³³ *Id.*

²³⁴ See Dobrow, *supra* note 228.

²³⁵ *Id.*

²³⁶ See *supra* Section I.C.

²³⁷ Frank Ready, *Could the GDPR 'Right to Access' Make Personal Data More Vulnerable?*, LAW.COM: LEGALTECH NEWS (Jan. 23, 2019, 10:30 AM), <https://www.law.com/legaltechnews/2019/01/23/could-the-gdpr-right-to-access-make-personal-data-more-vulnerable/?slreturn=20200222232331> [<https://perma.cc/A3MA-47ER>].

²³⁸ *Id.*

²³⁹ *Id.*

²⁴⁰ *Id.*

²⁴¹ Davis, *supra* note 232.

their privacy defaults on popular websites despite over sixty percent of people confessing that data privacy is a concern.²⁴² Indeed, despite the revolutionary nature of GDPR, a survey showed that consumers are actually more confused as to what happens to their data since GDPR became effective.²⁴³ It has been argued elsewhere that GDPR has done little more than “embolden[] [a] few data protection activists.”²⁴⁴ If consumers do not take advantage of privacy laws, government actors will be burdened with the brunt of enforcement, forced to choose between imposing steep penalties that could impair business,²⁴⁵ or letting violations go without consequence—leaving consumers as helpless as they started.

B. *The Anti-Discrimination Approach*

While far-reaching privacy laws may end up throttling the growth of data analytics, which may in turn inhibit efficiency and the advent of new healthcare technologies, an expansion of anti-discrimination laws could result in a healthcare system mutually beneficial to consumers and insurers alike. Legislation that prohibits an insurer from discriminating against its policyholders, or would-be policyholders, would calm consumer concerns that their rates will go up or they will be denied coverage. Simultaneously, anti-discrimination legislation would not prevent health organizations from using a patient or policyholder’s data to anticipate and more effectively address their health needs. This would

²⁴² See Geoffrey A. Fowler, *Hands Off My Data! 15 Default Privacy Settings You Should Change Right Now*, WASH. POST (June 1, 2018, 12:59 PM) <https://www.washingtonpost.com/news/the-switch/wp/2018/06/01/hands-off-my-data-15-default-privacy-settings-you-should-change-right-now> [https://perma.cc/53UE-PYWJ].

²⁴³ Sarah Vizard, *Despite GDPR, Consumers Still Don’t Understand How Brands Use Their Data*, MARKETING WEEK (May 25, 2018), <https://www.marketingweek.com/2018/05/25/despite-gdpr-consumers-dont-understand-how-brands-use-data> [https://perma.cc/35PM-G7ZS] (reporting that forty-eight percent of consumers do not know how their data is used, up from thirty-one percent in 2016).

²⁴⁴ See Kiskis, *supra* note 229 (also noting that when GDPR became effective, most companies simply updated their privacy policies—which annoyed consumers simply clicked through).

²⁴⁵ David O. Klein, *Comparing the California Consumer Privacy Act (CCPA) and the EU’s General Data Protection Regulation (GDPR)*, LEXOLOGY (Oct. 15, 2018), <https://www.lexology.com/library/detail.aspx?g=f3845305-195c-419c-ae17-63cb5a6f7fef> [https://perma.cc/49V3-CUDH].

ensure that the positives of data analytics in healthcare remain and progress.²⁴⁶ Additionally, anti-discrimination may allow for open and honest discourse surrounding consumer data, encouraging doctors and insurers to discuss potential inaccuracies²⁴⁷ in data profiles with their patients and policyholders. There are, however, two potential arguments against this approach, which are discussed and responded to below.

1. The Issue of Preemptive Action

Currently, there is little evidence that data discrimination exists in the healthcare industry.²⁴⁸ An argument can be made that it would be a waste of time to legislate where a problem may or may not exist.²⁴⁹ However, this argument fails for two reasons. First, the lack of evidence is likely due to the fact that the data analytics industry remains a shadowy figure, whose power is not appreciated or fully understood by most Americans.²⁵⁰ Without control over or knowledge about who is buying and selling data, there is little chance that consumers will be able to unearth any evidence of discrimination. Second, this would not be the first time that Congress legislated to preemptively address a discrimination issue.²⁵¹ As we saw with GINA, there can be other justifications for expanding the law before injustice occurs.²⁵² As long as data and predictive analytics are used by health insurance companies while remaining less than perfectly accurate,²⁵³ there is a chance that Americans could face unfair discrimination.

²⁴⁶ See *supra* Section I.C.2.

²⁴⁷ See *supra* Section II.C.2.a; see also Ravindranath, *supra* note 30.

²⁴⁸ See Allen, *supra* note 6.

²⁴⁹ This was an argument made by opponents of GINA. See Roberts, *supra* note 156, at 462–63.

²⁵⁰ Sacha Molitorisz, *It's Time for Third-Party Data Brokers to Emerge from the Shadows*, CONVERSATION (Apr. 4, 2018, 2:46 AM), <https://theconversation.com/its-time-for-third-party-data-brokers-to-emerge-from-the-shadows-94298> [<https://perma.cc/J4JE-QDPX>].

²⁵¹ Roberts, *supra* note 156, at 466–73.

²⁵² See *supra* Section II.C.1.

²⁵³ Timo Elliot, *How Trustworthy Is Big Data?*, BRINK (Feb. 2, 2018), <https://www.brinknews.com/how-trustworthy-is-big-data> [<https://perma.cc/8SPQ-FHLY>] (“Big data often comes from systems not directly controlled by an organization and may contain inherent biases or outright false values . . .”).

2. The Issue of Risk Assessment

As with genetic information,²⁵⁴ it could be said that analysis of consumer data is simply an element of risk assessment.²⁵⁵ Without access to genetic information (protected under GINA) or personal health information (protected under HIPAA), and an inability to discriminate on the basis of pre-existing conditions (ACA) or data profiles (new legislation), insurers will be left toothless when setting anything other than a standard premium. However, this new legislation would not affect an insurer's ability to have a policyholder or potential policyholder answer questions about their current state of health and habits, as is current practice.²⁵⁶ Additionally, many insurance companies already have programs that offer discounts or similar benefits in exchange for access to data obtained from wearable devices.²⁵⁷ In particular, John Hancock—a life insurance company, among other things—currently employs a program in which policyholders can get discounts for wearing Fitbits and reporting the information it collects.²⁵⁸ These kinds of programs can benefit insurance companies while also motivating policyholders to lead active, healthier lives.²⁵⁹ Healthier choices by consumers will in turn lead to fewer expensive medical payouts for insurers.²⁶⁰ Additionally, because there is no middleman or data broker

²⁵⁴ See Morrow, *supra* note 170, at 227–28.

²⁵⁵ See Ray Lehmann, *Why 'Big Data' Will Force Insurance Companies to Think Hard About Race*, INS. J. (Mar. 27, 2018), <https://www.insurancejournal.com/blogs/right-street/2018/03/27/484530.htm> [<https://perma.cc/KZ35-932J>] (“Unsurprisingly, insurers’ biggest use case for big data lies at the very heart of the business of insurance—the ability to assess risk.”).

²⁵⁶ Ann Carrns, *A Health Insurer Calls, with Questions*, N.Y. TIMES (June 10, 2014), <https://www.nytimes.com/2014/06/11/your-money/a-health-insurer-calls-with-questions.html> [<https://perma.cc/F2PY-5XYD>].

²⁵⁷ Tara Siegel Bernard, *Giving Out Private Data for Discount in Insurance*, N.Y. TIMES (Apr. 8, 2015), <http://www.nytimes.com/2015/04/08/your-money/giving-out-private-data-for-discount-in-insurance.html?> [<https://perma.cc/7Z9T-V86D>].

²⁵⁸ Andrew Dawson, *John Hancock Life Insurance Will Now Reward You for Being Active*, RUNNER'S WORLD (Sept. 20, 2018), <https://www.runnersworld.com/news/a23334787/john-hancock-life-insurance-rewards-activity-tracker-data> [<https://perma.cc/AT2H-E6D2>]; see also Troiano, *supra* note 21, at 1723–25.

²⁵⁹ Bernard, *supra* note 257.

²⁶⁰ Bruce Japsen, *To Keep You Healthy, Health Insurers May Soon Pay Your Rent*, FORBES (Aug. 14, 2018, 8:00 AM), <https://www.forbes.com/sites/brucejapsen/2018/08/14/to-keep-you-healthy->

in between the policyholder and the insurance company, the data would likely be far more accurate. This would arguably resolve Congress's concerns regarding fundamental unfairness,²⁶¹ and align closely with President Bush's comments about "individual merit."²⁶² New anti-discrimination legislation would not interfere with these programs, provided that they outline clear health goals and activity thresholds²⁶³ and the insurer obtains the policyholder's informed consent regarding the parameters and the potential downsides of participation, such as an increase in premiums resulting from a failure to reach particular goals. However, to prevent insurers from drawing a negative inference based off of a person's refusal to consent to data sharing, these programs should only be offered after a policy is already secured. Insurers would also not be prevented from using data to make big picture decisions regarding allocation of resources and targeting high-need communities, where many insurers already institute educational programs.²⁶⁴ Finally, considering the significant expenses associated with adapting to new privacy laws,²⁶⁵ any marginal loss in profits associated with rectifying premiums may be a welcome alternative to insurers.

health-insurers-may-soon-pay-your-rent/#44261adf67ce [https://perma.cc/YU2J-AKXA] (reporting that UnitedHealth and other insurers have invested heavily in affordable housing in recent years as part of an effort to keep indigent people healthy); *see also* Cheryl Wischhover, *A Life Insurance Company Wants to Track Your Fitness Data*, VOX (Sept. 20, 2018, 5:10 PM), <https://www.vox.com/the-goods/2018/9/20/17883720/fitbit-john-hancock-interactive-life-insurance> [https://perma.cc/M3TR-AC2K] (quoting the president of John Hancock: "The longer people live, the more money we make. If we can collectively help our customers live just a bit longer, it's quite advantageous for us as a company").

²⁶¹ Roberts, *supra* note 156, at 471–79.

²⁶² *See supra* note 166 and accompanying text.

²⁶³ For example, Attain, Aetna's health application, outlines specific daily and weekly activity goals based on a participant's age, sex, and weight. *See* Christina Farr, *Apple and Aetna Are Teaming Up on a New App to Help Track and Reward Healthy Behavior*, CNBC (Jan. 29, 2019, 5:00 AM), <https://www.cnbc.com/2019/01/28/apple-aetna-team-up-on-attain-health-tracking-app.html> [https://perma.cc/W9YS-GCQ2].

²⁶⁴ Charlotte Hu, *Health Insurer Cigna Is Zeroing in on Targeted Communities in an Effort to Combat the Opioid Crisis*, BUS. INSIDER (June 21, 2018, 10:37 AM), <https://www.businessinsider.com/cignas-plan-to-reduce-opioid-use-in-communities-with-higher-overdose-rates-2018-6> [https://perma.cc/QE6T-JVPP].

²⁶⁵ Oliver Smith, *The GDPR Racket: Who's Making Money from This \$9bn Business Shakedown*, FORBES (May 2, 2018, 2:30 AM), <https://www.forbes.com/sites/oliversmith/2018/05/02/the-gdpr-racket-whos-making-money-from-this-9bn-business-shakedown/#4d81c1ec34a2> [https://perma.cc/H5VA-VTZN] (reporting that it will cost Fortune 500 and FTSE 350 companies nine billion dollars to become compliant with GDPR).

CONCLUSION

The world of Big Data has bloomed rapidly and, in many respects, in secret. The rise of data analytics in the healthcare industry has provoked consumers' justifiable fears of discriminatory treatment. In the United States, the legislative scheme surrounding health privacy is entirely inadequate in the face of this new threat. A natural reaction to this fear is to advocate for privacy expansions rooted in consumer consent and transparency laws. This solution, however, could have unintended, debilitating consequences in the healthcare and many other industries. New anti-discrimination laws would be a snug fit in the current legislative scheme, preserving a balance between consumer protections and the benefits of a data-driven business landscape.