THE COLUMBIA
SCIENCE &
TECHNOLOGY
LAW REVIEW

VOL. XXII STLR.ORG FALL 2020

ARTICLE

BIG DATA'S EXPLOITATION OF SOCIAL DETERMINANTS OF HEALTH: HUMAN RIGHTS IMPLICATIONS

Sarah Wood*

This Article acknowledges the necessity of including social determinants of health (SDH) data in healthcare planning and treatment but highlights the lack of regulation around the collection of SDH data and potential for violating consumers' basic rights to be treated equally, protected from discrimination, and to have their privacy respected. The Article analyzes different approaches from the U.S. and EU and proffers the global application of the GDPR plus data human rights provisions as the most sustainable option in a world where technology is ever-changing.

I. Introduction 64
II. The Data Supply Chain 66
III. The Increasing Relevance of SDH Data and the Acceleration of Its Collection 67
IV. Threats to Consumers' Human Rights in Using Big Data to Collect SDH Information 73
A. Equal Protection and Safeguards Against Discrimination 73
B. Data Security 75

* Sarah Wood is a psychologist-attorney and consulting legal researcher at the University of Basel School of Law (Switzerland). She holds a doctorate of philosophy in clinical psychology from Palo Alto University and a doctorate of jurisprudence and master's degree in international legal studies from Golden Gate University School of Law. *She may be reached at swood@paloaltou.edu.* This Article was funded by a grant to the University of Basel from the Swiss National Science Foundation (Big Data NFP75). The author would like to thank Professor Sabine Gless at the University of Basel for her help and encouragement in writing this article.

C. Datafication and the Misuse of Data 78
V. Laws to Safeguard the Collection and Use of SDH Data: EU and U.S. 79
A. United States 81
B. Approaches Taken by the European Union and California 84
VI. Looking Forward: Broad Application of a GDPR "Plus" 86
VII. Conclusion 89

I. INTRODUCTION

According to the World Health Organization (WHO), social determinants of health (SDH) are "the conditions in which people are born, grow, live, work and age."[1] This idea that psychosocial, economic, and environmental factors affect health outcomes is not a new one. In fact, following the adoption of the Ottawa Charter in 1986,[2] the WHO developed the Healthy Cities program in 1987 based upon the idea of creating settings in which people's health is maximized through a holistic approach.[3] In the 1990s, Dahlgren and Whitehead advocated for worldwide health policies addressing health inequities, citing the thousands of lives in Europe alone that could be spared if opportunities to live healthy lives were made more equal across socioeconomic groups.[4] Recent estimates indicate that SDH account for 80-90 percent of health outcomes.[5]

Fast-forward two or three decades and, while the same inequities continue to exist, we now have access to a never-ending flow of data confirming the importance of SDH and its potential to be used to address social and healthcare disparities. For purposes of this paper, SDH data means data that is collected, combined, or analyzed to predict health outcomes of individuals. This definition does not discriminate between data used for commercial or governmental

---

1. *Social Determinants of Health: About Social Determinants of Health*, WHO, https://www.who.int/social_determinannts/sdh_definition/en/ (last visited Oct. 3, 2020).

2. International Conference on Health Promotion, *OTTAWA Charter*, WHO Doc. WHO/HPR/HEP/95.1 (Nov. 21, 1986).

3. WHO REG'L OFF. FOR EUR., ADDRESSING THE SOCIAL DETERMINANTS OF HEALTH: THE URBAN DIMENSION AND THE ROLE OF LOCAL ENVIRONMENT 1 (2012); *see also Healthy Settings*, WORLD HEALTH ORGANIZATION, https://www.who.int/healthpromotion/healthy-settings/en (last visited Nov. 22, 2020).

4. Göran Dahlgren et al., *Policies and Strategies to Promote Social Equity in Health: Background Document to WHO-Strategy Paper for Europe*, WHO REG'L OFF. FOR EUR., Sept. 1991, at 5.

5. Sanne Magnan, *Social Determinants of Health 101 for Health Care: Five Plus Five*, NAT'L ACAD. OF MED., Oct. 9, 2017, at 1, https://nam.edu/social-determinants-of-health-101-for-health-care-five-plus-five.

purposes, data publicly available, protected health data, or data collected by private organizations. As such, SDH data is not a subset of health data, but rather a seemingly benign collection of data points about an individual's lifestyle and life circumstances that, through big data analytics, are amalgamated into predictive tools. This big data amalgamation allows SDH data to be more powerful as a predictive tool than traditional health data.

There have been a number of successful projects promoting global health using big data analytics, including those seeking to reduce the incidence of communicable diseases in Uganda and Haiti.[6] However, extending the use of SDH data to other domains like social welfare programs and educational and occupational opportunities comes with significant risk to human rights, including equal protection and the right to privacy. This is especially poignant given that the breadth and depth of big data analytics are rapidly increasing and are now "poised to affect every aspect of our lives and environments."[7] While the collection, analysis, and use of data are exceedingly unequal across the world,[8] the fact remains that data *is* collected and used worldwide, typically through a sort of "data supply chain."

But a solution for protecting SDH data is complicated. Unlike personally identifiable information such as a person's social security number, SDH data is not always distinguishable from other types of data. Indeed, corporations today gather SDH data attached to personal and identifiable user data. Therefore, this paper asserts that, at a minimum, the European Union's General Data Protection Regulation[9] (GDPR) should be applied to SDH data, but that additional protections against surveillance and data manipulation, as suggested by Martin Tisne, must also be in place so that fundamental rights to privacy and health, as well as the right to not be discriminated against, are protected.

---

6. Galit A. Sarfaty, *Can Big Data Revolutionize International Human Rights Law?*, 39 U. PA. J. INT'L L. 73, 84 (2017).

7. Luca Belli, *The Need for a RIoT (Responsible Internet of Things): A Human Rights Perspective on IoT Systems*, *in* NAVIGATING A NEW ERA OF BUSINESS AND HUMAN RIGHTS 181, 184 (Matthew Mullen et al. eds., 2019), https://www.business-humanrights.org/sites/default/files/documents/a_new_era%20%281%29_0.pdf.

8. U.N. Secretary-General, *Special Edition: Progress Towards the Sustainable Development Goals*, ¶¶ 90–92, U.N. Doc. E/2019/68 (May 8, 2019), https://undocs.org/E/2019/68; *see also* Press Release, Secretary General, Access to Timely, Relevant, Disaggregated Data Remains Major Hurdle, Deputy Secretary General Tells 'Data4Now' Event, U.N. Press Release DSG/SM/1339-ENV/DEV/2000 (Sept. 25, 2019), https://www.un.org/press/en/2019/dsgsm1339.doc.htm.

9. Commission Regulation 2016/679, 2016 O.J. (L 119) [hereinafter GDPR].

This Article examines the problematic collection and use of social determinants of health data, as well as the lack of existing law to protect consumers. In doing so, it acknowledges the necessity of including SDH in healthcare planning and treatment but highlights the lack of regulation around the collection of SDH data and the potential for violating consumers' basic rights to be treated equally, protected from discrimination, and to have their privacy respected.[10] First, the Article introduces SDH data and discusses its collection. It then explores how that collection and use can be problematic and analyzes where U.S. and international law might be relevant but not adequately utilized. Finally, the Article concludes that legal reforms could ameliorate some of the problems around collection of such data. Specifically, it proffers the global application of the GDPR plus data human rights provisions as the most sustainable option in an ever-changing world.

II. THE DATA SUPPLY CHAIN

The collection of SDH data is not that dissimilar from a typical supply chain where traditional goods and services are transferred. However, data is unique in that without analytics it is not particularly valuable.[11] So it is likely that big data analytics will be at the helm of the digitalization of supply chains,[12] particularly given the rise of artificial intelligence and machine learning technologies. Organizations, including healthcare companies and hospitals, are increasingly outsourcing technology services, such as software development. While efficient, this creates "a cybersecurity blindspot"[13] that can be exploited when companies fail to make cybersecurity an organization-wide priority and allow unrestricted third-party access to their data.[14] These blind spots apply to every organization in a supply chain and, as such, cyber-attacks are magnified due to the sheer number of businesses involved and cybercriminals' abilities to find and exploit the weakest link.[15]

---

10. G.A. Res. 217 (III) A, Universal Declaration of Human Rights (Dec. 10, 1948) [hereinafter UDHR].

11. Yasaman Kazemi, *AI, Big Data & Advanced Analytics in the Supply Chain*, FORBES (Jan. 29, 2019, 11:15 AM), https://www.forbes.com/sites/yasamankazemi/2019/01/29/ai-big-data-advanced-analytics-in-the-supply-chain/#ef17a20244ff.

12. *Id.*

13. Richard Summerfield, *Dealing with Cyber Breaches in the Supply Chain*, FINANCIER WORLDWIDE (June 2017), https://www.financierworldwide.com/dealing-with-cyber-breaches-in-the-supply-chain#.XmpyA6hKi70.

14. *Id.*

15. *Id.*

According to Kirsten Martin, the data supply chain follows a path akin to supply chains with tangible goods.[16] The data supply chain includes the passage of information from consumers to companies, which then give data to tracking companies where it is passed along to data aggregators who then sell that information to any number of purchasers, including government and research organizations, but also advertising networks.[17] The complexity and opacity of the data supply chain results in biased and potentially unauthorized data collection and may harm consumers.[18]

Culnan and Milberg assert that information provided to a merchant as a byproduct during a business exchange can be separated out as a secondary exchange between the parties and it is this exchange that has a greater risk of violating the consumer's privacy because information is not conveyed to the customer and because there is a dearth of regulations regarding the disclosure of such information.[19] For example, employers across the U.S. encourage their employees to participate in wellness programs which involve the provision of health data in exchange for financial incentives.[20] Depending on who offers the program (an employer or health insurer, for instance) the data may or may not be regulated by the Health Insurance Portability and Accountability Act (HIPAA) and, even where HIPAA applies, copies of employee health data may be passed along to businesses that do not fall under the auspices of HIPAA.[21] It is these transfers to third parties that become especially worrisome given that consumers may not know the identities of all the secondary businesses involved nor the specific use and purpose of their personal data. So, the typical "supply chain" is increasingly outsourced to third party aggregators and sellers, leading to cybersecurity vulnerabilities and an inability for consumers to control their data.

III. THE INCREASING RELEVANCE OF SDH DATA AND THE ACCELERATION OF ITS COLLECTION

The risks and opacity inherent in the aforementioned data supply chain are especially problematic in view of the growing relevance of SDH data, the increasing prevalence of its collection, and the plethora

16. Kirsten E. Martin, *Ethical Issues in the Big Data Industry*, 14 MIS Q. EXEC. 67, 70–72 (2015).

17. *Id.* at 70–71.

18. *Id.* at 70–72.

19. MARY J. CULNAN ET AL., THE SECOND EXCHANGE: MANAGING CUSTOMER INFORMATION IN MARKETING RELATIONSHIPS 5–8 (1998).

20. Thorin Klosowski, *What to Consider Before Trading Your Health Data for Cash*, WIRECUTTER (Nov. 20, 2019), https://thewirecutter.com/blog/what-to-consider-before-trading-your-health-data-for-cash.

21. *Id.*

of uses employed by the medical field, public sector, and tech industry. Governments and health experts have realized the importance of SDH while at the same time, the tech sector has enabled the collection and use of SDH data at a previously unknown scale. As will be explored in the following Part, this combination of SDH data and big data analytics leads to a number of problems, including implications for the fundamental right to privacy,[22] the right to health,[23] and the right to equal protection under the law.[24]

The right to health, regardless of social status, was deemed a fundamental human right in the Constitution of the World Health Organization in 1946.[25] Even then, "the absence of disease" was not the benchmark for health, but rather, "complete physical, mental and social well-being."[26] Efforts to address the socioeconomic and environmental contributions to health continue[27] and are visible in several of the goals of the UN 2030 Agenda for Sustainable Development, including optimizing urban safety and inclusivity.[28] Goal 3 of the UN 2030 Agenda involves the promotion of well-being for people of all ages within the framework of leaving no one behind.[29] To this end, the WHO European Region and Health Evidence Network established specific policies to aid in reducing health inequities by addressing social determinants of health, including early childhood education, employment opportunities and improved working conditions, social protection through the use of social cash transfers, and improved living environments.[30]

In addition to these international organizations, the importance of SDH is also recognized at a national level in the United States. The National Academies of Sciences, Engineering, and Medicine has undertaken to educate medical providers on the importance of addressing SDH.[31] The same is true of the U.S. government's Healthy People Initiative, which seeks to improve public health by reducing disparities in literacy rates, high school

22. UDHR art. 12; U.S. CONST. amend. IV.

23. UDHR art. 25.

24. UDHR art. 7; U.S. CONST. amend. XIV.

25. CONST. OF THE WORLD HEALTH ORGANIZATION pmbl.

26. *Id.*

27. WHO COMM'N ON SOCIAL DETERMINANTS OF HEALTH, CLOSING THE GAP IN A GENERATION: HEALTH EQUITY THROUGH ACTION ON THE SOCIAL DETERMINANTS OF HEALTH 1 (2008).

28. G.A. Res. 70/1, at 14 (Oct. 21, 2015).

29. *Id.*

30. MATTHEW SAUNDERS ET AL., HEALTH EVIDENCE NETWORK SYNTHESIS REPORT 52: KEY POLICIES FOR ADDRESSING THE SOCIAL DETERMINANTS OF HEALTH AND HEALTH INEQUITIES, at vii–x (2017).

31. NAT'L ACAD. OF SCI. ENG'G & MED., A FRAMEWORK FOR EDUCATING HEALTH PROFESSIONALS TO ADDRESS THE SOCIAL DETERMINANTS OF HEALTH 1–7 (2016).

graduation rates, access to health services, and other metrics.[32] The Robert Wood Johnson Foundation also provides funding to reduce health inequity in a variety of areas, with an emphasis on SDH.[33] Despite these and a slew of other national and state initiatives,[34] significant health disparities in the U.S. remain.[35]

While there is clear consensus that SDH should be taken into account in healthcare systems globally, approaches obviously vary from country to country and the availability of resources to capitalize on SDH data differs between developing and developed nations. Further, the need for large quantities of data to make the link between various social, economic, and environmental factors and individuals' health is clear. Enter big data. Nowadays, people all over the world constantly produce "digital exhaust" in the form of consumer data (internet search histories, social networking data, shopping habits, wearable fitness tracker data, etc.) that is quickly swept up by large corporations, analyzed, and sold to the health care industry.[36] That said, as SDH data is becoming increasingly commercialized, risks to consumers' privacy and the potential for bias in data collection and analysis have become an urgent human rights issue.

Currently, collection of SDH data by public health organizations in the U.S. varies by state and not all share the same priorities with respect to the use of SDH data.[37] However, the problem of how to collect and share data goes beyond state lines, as does the technical difficulty associated with creating and maintaining SDH datasets.[38] The volume, velocity (the speed at which data is generated), and variety of big data—the original '3Vs' commonly referred to in the literature—are at the core of both its challenges and potential rewards.

---

32. *Social Determinants*, OFF. OF DISEASE PREVENTION AND HEALTH PROMOTION, https://www.healthypeople.gov/2020/leading-health-indicators/2020-lhi-topics/Social-Determinants (last visited Nov. 22, 2020).

33. *Social Determinants of Health,* ROBERT WOOD JOHNSON FOUND., https://www.rwjf.org/en/our-focus-areas/topics/social-determinants-of-health.html (last visited Nov. 22, 2020).

34. Samantha Artiga et al., *Beyond Health Care: The Role of Social Determinants in Promoting Health and Health Equity*, HENRY J. KAISER FAM. FOUND. (May 10, 2018), https://www.kff.org/disparities-policy/issue-brief/beyond-health-care-the-role-of-social-determinants-in-promoting-health-and-health-equity/.

35. Kristin Voigt, *Social Justice, Equality and Primary Care: (How) Can 'Big Data' Help?*, 32 PHIL. & TECH. 57, 59–60 (2019).

36. Kirsten Ostherr, *You Don't Want Facebook Involved in Your Healthcare: Big Tech Companies Want to Share Data About You with Your Doctors*, SLATE (Sept. 19, 2019, 7:30 AM), https://slate.com/technology/2019/09/social-determinants-health-facebook-google.html.

37. Anna Spencer et al., *Measuring Social Determinants of Health Among Medicaid Beneficiaries: Early State Lessons*, CENTER FOR HEALTH CARE STRATEGIES, INC. (Dec. 2016), https://www.chcs.org/media/CHCS-SDOH-Measures-Brief_120716_FINAL.pdf.

38. *Id.*

While SDH can predict health outcomes, SDH data has not traditionally been considered medical data, but rather commercial data.[39] This is no longer an appropriate classification when corporations have access to additional collateral data like location and online search and purchasing history, which allows SDH data to easily be linked to a single person. Using SDH to predict health outcomes triggers issues of data protection given that personal medical data is generally governed by different (and more stringent) standards. Further, the context in which databases are created across professions reveals different methodologies, in addition to standards and norms.[40] The Institute of Medicine's recommendation of including SDH data in electronic health records (EHR)[41] offers a potentially safer means of collecting and using SDH as health data, but does not address the issue of corporate collection of the same data without consumer consent, nor does it address the bias against marginalized populations within algorithms even when data is collected with the consumer's consent.

Within the medical field, SDH have come to the forefront because, while pharmacotherapy continues to advance and genetic testing for various diseases has expanded, it has become apparent that the most common, chronic, and debilitating medical conditions such as heart disease, stroke, and diabetes cannot be explained by genetic factors alone[42] and disproportionately affect minority populations.[43] Aspects of socioeconomic status like education, living environment (including access to basic needs such as food and water), and employment (such as opportunities for work, health coverage, and working conditions) are far more predictive of health outcomes than genetic makeup.[44] Across the world, this knowledge is slowly

39. Roland Gamache et al., *Public and Population Health Informatics: The Bridging of Big Data to Benefit Communities*, 2018 IMIA Y.B. OF MED. INFORMATICS 199, 204 (2018).

40. Chloé Dimeglio et al., *Expectations and Boundaries for Big Data Approaches in Social Medicine*, 57 J. FORENSIC & L. MED., at 51, 53 (2018).

41. INST. OF MED., CAPTURING SOCIAL AND BEHAVIORAL DOMAINS AND MEASURES IN ELECTRONIC HEALTH RECORDS PHASE 2, at 227–35 (2014), https://www.ncbi.nlm.nih.gov/books/NBK268995/pdf/Bookshelf_NBK268995.pdf.

42. *See, e.g.*, Chris Carlsten et al., *Genes the Environment and Personalized Medicine: We Need to Harness Both Environmental and Genetic Data to Maximize Personal and Population Health*, 15 EMBO REP. 736, 736 (2014); Paul Braveman et al., *The Social Determinants of Health: Coming of Age*, 32 ANN. REV. PUB. HEALTH 381, 384 (2011).

43. Pamela A. Meyer et al., *Introduction: CDC Health Disparities and Inequalities Report – United States, 2013*, 62 MORBIDITY AND MORTALITY WKLY. REP. 3, Nov. 22, 2013, at 1, 3.

44. *Id.*; *see also* H. Jack Geiger, *Community-Oriented Primary Care: A Path to Community Development*, 92 AM. J. PUB. HEALTH 1713, 1713 (2002) (noting Sidney and Emily Kark's pioneering health care program in South Africa based on

translating into valuing patient outcome over patient volume and the development of incentivized payment systems.[45] However, the parallel movement in precision/personalized medicine has resulted in a funding shift away from public health to individualized genomic research despite the potential to develop population-based interventions.[46] Evidence of this shift is further supported by the fact that big data is already being used in personalized medicine and, even with the potential for black-box issues going forward,[47] will likely continue to develop because of its vast potential.

Projects through the UN's Global Pulse program are using big data in the form of call records,[48] postal data,[49] and satellite images of household roof type[50] to understand socioeconomic factors in nations around the world. While such endeavors are laudable in that collection of data is crucial to understanding the overall well-being of any society, the use of big data to do so comes with the risk of sacrificing some human rights like privacy and consent.[51] The UN Special Rapporteur on extreme poverty and human rights, Philip Alston, recently spoke out against the alarming practice of identifying and surveilling those seeking social assistance with software and devices from big tech companies without any requirement that the companies adhere to human rights standards.[52]

The tech industry has followed the healthcare industry's movement toward inclusion of SDH and has offered up its nearly limitless ability to mine and analyze people's data.[53] In fact, the amount of information available to health care providers has become

---

the belief that SDHs were the primary indicators of health status); WHO, *supra* note 27.

45. Michael Counte et al., *Global Advances in Value-Based Payment and Their Implications for Global Health Management Education, Development, and Practice*, FRONTIERS PUB. HEALTH, Jan. 18, 2019, at 1, 2.

46. Muin J. Khoury et al., *Will Precision Medicine Improve Population Health?*, 316 JAMA 1357, 1357 (2016).

47. W. Nicholson Price II, *Black-Box Medicine*, 28 HARV. J. L. & TECH. 420, 420 (2015).

48. *Estimating Socioeconomic Indicators From Mobile Phone Data in Vanuatu*, UN GLOBAL PULSE, https://www.unglobalpulse.org/projects/estimating-socioeconomic-indicators-mobile-phone-data-vanuatu (last visited Nov. 22, 2020).

49. *Building Proxy Indicators of National Wellbeing with Postal Data,* UN GLOBAL PULSE, https://beta.unglobalpulse.org/wp-content/uploads/2016/06/building-proxy-indicators-of-national-wellbeing-with-postal-data.pdf (last visited Nov. 22, 2020).

50. *Measuring Poverty with Machine Roof Counting*, UN GLOBAL PULSE, https://www.unglobalpulse.org/projects/measuring-poverty-machine-roof-counting (last visited Nov. 22, 2020).

51. Sarfaty, *supra* note 6.

52. Phillip Alston (Special Rapporteur on Extreme Poverty and Human Rights), *Report of the Special Rapporteur on Extreme Poverty and Human Rights*, U.N. Doc. A/74/48037 (Oct. 11, 2019).

53. Ostherr, *supra* note 36.

so immense that some advocate for an entirely new profession – so-called "health information counselors" to help providers weed through all of it.[54] Facebook, Google, Microsoft, and Amazon are all looking to cash in on providing data generated by their customers to healthcare entities. Offers of "the potential to improve care, save lives and lower costs"[55] are of course appealing, provided there are adequate mechanisms in place to protect the public and address issues of systemic discrimination.

The vast majority of large companies in the U.S. use big data analytics[56] and, while data may be purported to be collected for one purpose, connected devices like smart watches, fitness trackers, and even smart furniture automatically collect more information than advertised and often sell that information for alternate, undisclosed purposes.[57] Fitness trackers and health apps have been increasingly used in criminal trials,[58] home security cameras and virtual assistants have been found to record video and voice data without users' knowledge,[59] and so-called "smart cities" are on the rise, using facial recognition, GPS tracking, and other technology in an attempt to reduce crime rates, traffic congestion, and other issues plaguing urban environments.[60] The current use of technology to surveil marginalized populations at a significantly higher rate than those with greater wealth indicates that safeguards have not been put in place to ensure equal protection under the law.[61]

This trend is not limited to private companies within industrialized nations; both developed and developing nations are

---

54. Amelia Fiske et al., Health Information Counselors: A New Profession for the Age of Big Data, 94 ACAD. MED. 37, 37 (2019).

55. Wullianallur Raghupathi et al., *Big Data Analytics in Healthcare: Promise and Potential*, HEALTH INFO. SCI. & SYS., 2014, at 1, 1.

56. THERESA M. PAYTON ET AL., PRIVACY IN THE AGE OF BIG DATA, at vii–ix (2014).

57. Belli, *supra* note 7, at 169.

58. *Apple Health Data Used in Murder Trial*, BBC NEWS (Jan. 12, 2018), https://www.bbc.com/news/technology-42663297; *see also* Erin Moriarty, *The FitBit Alibi: 21st Century Technology Used to Help Solve Wisconsin Mom's Murder*, CBS NEWS (Oct. 20, 2018), https://www.cbsnews.com/news/the-fitbit-alibi-21st-century-technology-used-to-help-solve-wisconsin-moms-murder.

59. Paul Roberts, *Updated: Green Light or No, Nest Cam Never Stops Running*, SECURITY LEDGER (Nov. 24, 2015, 3:59 PM), https://securityledger.com/2015/11/green-light-or-no-nest-cam-never-stops-watching; *see also* Tom Simonite, *How Alexa, Siri, and Google Assistant Will Make Money Off You*, MIT TECH. REV. (May 31, 2016), https://www.technologyreview.com/s/601583/how-alexa-siri-and-google-assistant-will-make-money-off-you.

60. Eva Blum-Dumontet, *Smart Cities: Utopian Vision, Dystopian Reality,* PRIVACY INT'L (Oct. 31, 2017), https://privacyinternational.org/report/638/smart-cities-utopian-vision-dystopian-reality.

61. Alston, *supra* note 52.

embracing this technology with minimal oversight.[62] Data, including SDH data, is increasingly used to develop risk scores across a variety of domains, ranging anywhere from the private sector's determination of creditworthiness to a government or state agency's determination of an offender's risk of recidivism,[63] although the algorithms behind these determinations are largely inaccessible.[64] China has taken to using a combination of these types of risk scores to compile an overall social credit score that affects an individual's access to schooling, housing, and work promotions.[65]

In short, SDH has grown increasingly relevant. While the amount of SDH data collected has increased, so too has the purposes for which tech companies and governments put it to use. While this comes with some efficiencies, it also comes with significant challenges to international law, data security, and privacy.

IV. THREATS TO CONSUMERS' HUMAN RIGHTS IN USING BIG DATA TO COLLECT SDH INFORMATION

A number of international and domestic laws are applicable to the collection of consumer SDH data. This Part argues that while some have kept up with the shift toward increased use of data analytics, most fall short. Issues around consent and privacy remain at the forefront of any discussion regarding the potential for rights violations through the use of big data analytics.

A. Equal Protection and Safeguards Against Discrimination

The practices described in the previous Part contravene both the WHO Constitution and Article 7 of the Universal Declaration on Human Rights granting all people equal protection under the law[66] through the targeting of marginalized groups and perpetuating socioeconomic class divisions.[67] Biases in algorithmic development and data collection lead to inequality in application across

62. *Id.*

63. Lina Dencik et al., *Data Scores as Governance: Investigating Uses of Citizen Scoring in Public Services Project Report*, DATA JUSTICE LAB (Dec. 6, 2018), https://datajusticelab.org/2018/12/06/data-scores-as-governance-final-report-published.

64. VIRGINIA EUBANKS, AUTOMATING INEQUALITY: HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR 142–47 (2017).

65. Dencik et al., *supra* note 63; *see also* Ivon Watson et al., *China's Xinjiang Camps: Leaked Records Expose How Uyghurs are Judged and Detained,* CNN (February 2020), https://edition.cnn.com/interactive/2020/02/asia/xinjiang-china-karakax-document-intl-hnk (detailing an extreme example of the lengths to which a government can monitor and track socioeconomic information and use it to segregate and detain an entire ethnic group).

66. UDHR art. 7.

67. Sarah Valentine, *Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control*, 46 FORDHAM URB. L.J. 364, 365 (2019).

socioeconomic classes. Authorities tend to view algorithms as infallible[68] yet data collection is not equal across socioeconomic classes: "[p]eople of color, migrants, unpopular religious groups, sexual minorities, the poor, and other oppressed and exploited populations bear a much higher burden of monitoring and tracking than advantaged groups."[69] Cathy O'Neil terms this phenomenon a feedback loop created by "weapons of math destruction," where biased algorithms remain unchecked for accuracy yet are assumed to be correct in their output.[70]

People are put into categories (for example, parolees likely to reoffend) prior to ever acting and despite the known statistical limitations in predicting behavior.[71] Challenging such predictive models is difficult, even if the logic or other evidence suggests the model is producing erroneous data. As such, those for whom the model predicted poor behavior face an uphill battle overcoming such predictions.[72] That is, of course, only the case if one is allowed to scrutinize the model. Nowadays, and particularly in the case of big tech companies, the models themselves are deemed protected intellectual property and, therefore, do not have to be disclosed, let alone scrutinized by outside parties.[73] Acknowledgement of the need for "technological due process" under federal law is critical to provide notice to citizens and the opportunity to challenge biased algorithms that result in them being treated unfairly under the law.[74]

In addition to biases in data collection, biases can arise from the humans who create (or pay for the creation of) the algorithmic models, all of whom have their own values, ideology, and goals for the model.[75] As one might predict, goals for algorithmic models in western society tend to be increased profits or status (political or otherwise), neither of which have the general public's interests in mind. However, even those algorithms designed to address a public issue such as crime[76] or healthcare needs[77] have been shown to be biased against marginalized populations.

---

68. *Id.* at 366.

69. EUBANKS, *supra* note 64, at 6.

70. CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY (2016).

71. PAYTON ET AL., *supra* note 56, at 32.

72. O'NEIL, *supra* note 70, at 9.

73. *Id.* at 199; *see also* Obermeyer et al., *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, 366 SCIENCE 447, 447 (2019).

74. Danielle Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1249–52 (2008).

75. *Id.*; *see also* Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54, 67 (2019).

76. Katyal, *supra* note 75, at 68, 75–78.

77. Obermeyer et al., *supra* note 73, at 447.

In the case of SDH data, algorithms act as proxies of health despite the fact that the algorithmic models used to collect data on the various SDH are generally not related to health outcomes. These algorithms arguably violate citizens' right to equal protection under international and federal law given the biases described above. Similarly, when disparate databases (e.g., health records, social services, records, financial records) are combined in an effort to address SDH, significant issues arise due to each database's unique purpose and use, as well as contextual and methodological factors.[78] Like the algorithmic goals described above, a database created by a social services agency to collect information related to the determination of benefits has vastly different goals than a hospital's electronic medical record (EMR) system devised to store patient data. The unintended use of such different databases creates what Friedman and Nissenbaum coined an "emergent bias" in the 1990s.[79] Additionally, while all datasets share the common problems of missing and erroneous data, these major flaws are not corrected for when using data analytics created by for-profit big tech companies; instead, public and private data are combined haphazardly and sold to create faulty and dangerous predictive analytics.[80] Inaccurate data and biased algorithms largely go unchallenged due to the opacity inherent in big data analytics and a culture in which governments and other organizations place blind faith in technology and its developers, whom typically do not come from marginalized socioeconomic groups.[81] These factors limit society's ability to collect SDH data and use it for the public good.

B. Data Security

The existing legal framework also harms the security of consumers' SDH data. While a thorough analysis of the security threats posed by the internet is beyond the scope of this paper, it should be noted that the magnitude of the security breaches worldwide contrasts starkly with the Universal Declaration's protection against arbitrary interference with privacy.[82] In 2019, 61 percent of firms surveyed in several EU countries and the U.S. reported at least one cyber-attack that year.[83] A similar number of

78. Dimeglio et al., *supra* note 40, at 52.

79. Batya Friedman et al., *Bias in Computer Systems*, 14 ACM TRANSACTIONS ON INFO. SYS. 330, 335 (1996).

80. Valentine, *supra* note 67, at 387–89; *see generally* EUBANKS, *supra* note 64.

81. Valentine, *supra* note 67, at 395–96*; see also,* Katyal, *supra* note 75.

82. UDHR, art. 12.

83. HISCOX LTD., HISCOX CYBER READINESS REPORT 1–2, (2019), https://www.hiscox.com/documents/2019-Hiscox-Cyber-Readiness-Report.pdf.

organizations reported insider attacks last year.[84] Ransomware attacks have grown over 350 percent annually[85] and the healthcare industry is among the most frequently targeted.[86] Because SDH data is often passed to the healthcare industry from big tech, cyberattackers targeting the healthcare industry can potentially access SDH data.

The definition of SDH is broad and encompasses a large variety of data which consumers provide to big tech companies on a daily basis (where you live and work, who your friends and family are, what you search and post online, etc.). Online searches and keystrokes are monitored and analyzed by big tech, with Google being the largest tracker (recent data indicates they account for two-thirds of internet traffic[87]). Advertisements can redirect you to content that a specific person or company wants you to see[88] and a recent investigation by the Wall Street Journal found that apps on your phone can and do send personal data about things like your physical and mental health to companies such as Facebook and Google without your knowledge or consent.[89] Importantly, this finding by the Wall Street Journal came well after the U.S. Federal Trade Commission's 2012 charges against Facebook around issues of deception and privacy violations and after probes into the Cambridge Analytica scandal had begun, suggesting little change came from the inquiry. Although the inquiry eventually resulted in the U.S. Federal

---

84. *Bitglass 2019 Insider Threat Report: 41 Percent of Organizations Do Not Monitor User Behavior Across Their Cloud Footprints*, BITGLASS (April 3, 2019), https://www.bitglass.com/press-releases/threatbusters-2019-insider-threat-report.

85. *Security Report: Health Care - Hospitals, Providers and More*, CORVUS INSURANCE (2020), https://info.corvusinsurance.com/hubfs/Security%20Report%202.2%20-%20Health%20Care%20.pdf

86. Rob Sobers, *110 Must-Know Cybersecurity Statistics for 2020*, VARONIS (Sept. 24, 2020), https://www.varonis.com/blog/cybersecurity-statistics/.

87. Bennett Cyphers, *Don't Play in Google's Privacy Sandbox*, ELECTRONIC FRONTIER FOUND. (Aug. 30, 2019), https://www.eff.org/deeplinks/2019/08/dont-play-googles-privacy-sandbox-1.

88. Patrick Berlinquette, *I Used Google Ads for Social Engineering. It Worked.*, N.Y. TIMES (July 7, 2019), https://www.nytimes.com/2019/07/07/opinion/google-ads.html.

89. Sam Schechner & Mark Secada, *You Give Apps Sensitive Personal Information. Then They Tell Facebook,* WALL ST. J. (Feb. 22, 2019, 11:07 AM), https://www.wsj.com/articles/you-give-apps-sensitive-personal-information-then-they-tell-facebook-11550851636; *see also* Kit Huckvale et al., *Assessment of the Data Sharing and Privacy Practices of Smartphone Apps for Depression and Smoking Cessation*, JAMA NETWORK OPEN, Apr. 19, 2019, at 1, https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2730782 (using depression and smoking cessation apps as examples)*; see also* Sam Schechner, *Popular Apps Cease Sharing Data with Facebook*, WALL ST. J. (Feb. 24, 2019, 5:46 PM), https://www.wsj.com/articles/popular-apps-cease-sharing-data-with-facebook-11551044791 (listing several apps that changed their practices following the publishing of the WSJ article and Facebook's request that they not send private information).

Trade Commission lodging a $5 billion fine against Facebook in 2019 for its part in the Cambridge Analytica case,[90] Facebook's stock actually rose after the announcement.[91] The lack of adequate privacy laws in the U.S and any discernable change by corporations despite hefty fines. remains a concern for many citizens: a recent survey indicated that the majority of Americans feel their data is not private as it is collected by the government and corporate America.[92]

Although consent and privacy are heavily intertwined, in the U.S. HIPAA requires a patient's consent to allow the transfer or disclosure of medical data. As mentioned above, however, consumer data and SDH data typically do not fall within this protection and big tech is therefore able to use it without consumers' knowledge. Big tech is also getting its hands on citizens' medical data because HIPAA doesn't regulate tech companies,[93] nor are smart technology devices considered medical devices.[94] Even where hospitals or medical centers are involved in the collection and/or distribution of data, they invoke the exception that allows for the use of de-identified data in research.[95] Alarmingly, re-identification of specific individuals from anonymized data points is not as difficult as one would hope: the ease of re-identification increases with the number of data points in a particular entry.[96] Some researchers have found just four anonymous mobility data points were needed to reidentify nearly all of the members of a particular dataset.[97] Citing the fundamental right to

90. Lesley Fair, *FTC's $5 Billion Facebook Settlement: Record-Breaking and History-Making*, FED. TRADE COMM'N. (July 24, 2019), https://www.ftc.gov/news-events/blogs/business-blog/2019/07/ftcs-5-billion-facebook-settlement-record-breaking-history.

91. Rob Price, *Why Facebook's Stock Jumped Despite Facing A Record-Breaking $5 Billion FTC Penalty: 'A Slap on the Wrist'*, BUS. INSIDER (July 12, 2019, 7:51 PM), https://www.businessinsider.com/facebook-stock-rose-news-5-billion-ftc-settlement-why-critics-2019-7?r=US&IR=T.

92. Brooke Auxier et al., *Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information*, PEW RSCH. CTR. (Nov. 15, 2019), https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information.

93. Mona Sobhani et al., *All Our Data is Health Data. And the Tech Companies Have It All,* MEDIUM (Aug. 14, 2019), https://medium.com/@usccbc/all-our-data-is-health-data-57d3cf0f336d; *see also* PAYTON ET AL., *supra* note 56.

94. Charlie Warzel, *All Your Data is Health Data. And the Big Tech Has it All,* N.Y. TIMES (Aug. 13, 3019), https://www.nytimes.com/2019/08/13/opinion/health-data.html.

95. Sobhani et al., *supra* note 93.

96. Price II, *supra* note 47, at 457.

97. Yves-Alexandre de Montjoye et al., *Unique in the Crowd: The Privacy Bounds of Human Mobility*, SCI. REP., Mar. 25, 2013, at 1, 1; *see also* Yves-Alexandre de Montjoye et al., *Unique in the Shopping Mall: On the Reidentifiability of Credit Card Metadata,* SCI., Jan. 29, 2015, at 536, 538. *But see* David Sánchez et al., *Comment on 'Unique in the Shopping Mall: On the*

privacy codified in the European Convention, the European Court of Justice recently required Google to limit its processing of certain types of personal data.[98] While some disagree as to whether the decision has created a fundamental right to be forgotten, the decision highlights the relevance and importance of human rights as they pertain to the collection and use of consumer data.

In sum, while the cybersecurity threats to SDH data have grown it has remained remarkably easy to share such data without consumer consent. Though de-identification and the 'right to be forgotten' may seem promising, it is unclear that these methods will be sufficient to ensure the security of rapidly spreading SDH data.

C. Datafication and the Misuse of Data

The collection and use of SDH data poses further harms through datafication. Many in modern society deem big data analytics the future of research—despite significant issues around privacy, security, and bias. Decades ago, David Shenk discussed the "data smog" and highlighted the significant psychological effects (primarily anxiety) of being inundated with too much information in an increasingly datafied world.[99] Today, while societies continue to grapple with the impossibility of keeping up with all the available data out there, increased datafication[100] has created a different kind of problem in big data analytics, namely the beliefs that big data knows no limits of competence and that exceedingly large datasets, simply because of their size, adequately and accurately represent reality.[101]

Unfortunately, today, "personal data are treated solely as an economic asset, with proliferation of data viewed positively."[102] The collectors of data, like Google, Facebook, and Twitter, and those seeking to use such data tout it as objective truth akin to raw data collected by scientists in traditional experiments. This is a fallacy—data is not collected in a vacuum, or a controlled laboratory. The data is not collected randomly, but via a process akin to convenience

---

Reidentifiability of Credit Card Metadata,' SCI., Mar. 18, 2016, at 1274 (arguing existing anonymization techniques are adequate).

98. Case C-131/12, Google Spain SL v. Agencia Española de Protección de Datos, ECLI:EU:C:2014:317, ¶¶ 113, 118–38 (May 13, 2014).

99. DAVID SHENK, DATA SMOG: SURVIVING THE INFORMATION GLUT 27–31 (1997).

100. Datafication, according to Mayer-Schönberger and Cukier, is the collection of and transformation of information (data) into something quantifiable. *See* VIKTOR MAYER-SCHÖNBERGER ET AL., BIG DATA: A REVOLUTION THAT WILL TRANSFORM HOW WE LIVE, WORK, AND THINK 15 (2013).

101. MIREILLE HILDEBRANDT, SMART TECHNOLOGIES AND THE END(S) OF LAW 36–40 (2015).

102. Orla Lynskey, *A Legal Response to Data-Driven Mergers, in* BEING PROFILED: COGITAS ERGO SUM: 10 YEARS OF PROFILING THE EUROPEAN CITIZEN 78, 79 (Emre Bayamlioğu et al. eds., 2018).

sampling; only those who use the internet, and in many cases, social media more specifically, are taken into account. José van Dijck highlighted a 2012 Pew Research Center study that found only 15% of Americans used Twitter.[103] A poll published in 2019 by the same organization found that the number has grown slightly (22%) but also found that Twitter's users are largely comprised of younger and well-educated Americans, and 80% of tweets come from just 10% of those users.[104] So, regardless of how much data is collected from Twitter or other tech platforms, that data will not be representative of Americans as a whole. It would be inappropriate to make broad generalizations about American society based on this data and to do so may result in misguided governmental policies and corporate strategies.

With increased datafication comes unintended and unanticipated use of data previously collected for a specific (and different) purpose.[105] Alarmingly, service providers may elect to repurpose data they have collected and sell it to a third party or may collect additional data not necessarily related to the service but available to the provider because of the access it has to its users.[106] This concept, known as dataveillance,[107] is particularly relevant to the collection and use of SDH data. Dataveillance involves the gathering of metadata by corporations (and governments) without a predefined purpose. Because dataveillance has no predefined purpose, actors engaging in dataveillance can use the data collected in a variety of ways without informing the consumer.[108] Further, those collecting the data have little interest in transparency and their actions are never subjected to public scrutiny.

V. LAWS TO SAFEGUARD THE COLLECTION AND USE OF SDH DATA: EU AND U.S.

The fact that law has been unable to keep up with technology is particularly important in a day and age where so much data is being

---

103. Josè van Dijck, *Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology,* 12 SURVEILLANCE & SOC'Y 197, 199 n.2 (2014).

104. Adam Hughes et al., *10 Facts About Americans and Twitter,* PEW RES. CTR. (Aug. 2, 2019), https://www.pewresearch.org/fact-tank/2019/08/02/10-facts-about-americans-and-twitter.

105. Katherine J. Strandburg, *Monitoring, Datafication, and Consent: Legal Approaches to Privacy in the Big Data Context, in* PRIVACY, BIG DATA, AND THE PUBLIC GOOD FRAMEWORKS FOR ENGAGEMENT 5, 10 (Julia Lane et al. eds., 2018); *see also* Kate Conger et al., *Facebook's Suspension of 'Tens of Thousands' of Apps Reveals Wider Privacy Issues*, N.Y. TIMES (Sept. 20, 2019), https://www.nytimes.com/2019/09/20/technology/facebook-data-privacy-suspension.html.

106. Strandburg, *supra* note 105, at 11–12.

107. van Dijck, *supra* note 103, at 198–99.

108. *Id.* at 205.

collected from consumers around the world without their knowledge. As mentioned above, the classification of SDH information as commercial data triggers significantly weaker protection than it would receive if it were deemed health data, and this is particularly true in the U.S. where a GDPR equivalent does not exist. Even the GDPR, however, has obvious limitations in its scope, including the fact that the protection of personal data still allows machine learning to use anonymized aggregated data in order to arrive at biased conclusions that negatively affect individuals and groups alike.[109]

In the U.S., a number of bills have been introduced by a variety of senators on the topic of data privacy,[110] including one seeking to categorize wearable devices and consumer genetic testing services as personal health data, but none have passed yet and, in an effort to achieve bipartisan support, most do not offer the sweeping protections around the use of personal data afforded by the GDPR. In the EU, while there are specific regulations for health data, the GDPR includes provisions addressing public health research[111] and its application extends to non-EU entities such as Facebook and Google who process the data of EU persons. Following the 2018 Cambridge Analytica scandal, California passed the Consumer Privacy Protection Act (CPPA), which went into effect January 2020.[112] It applies to all California customers and is based on the state's constitutional right to privacy, granting consumers the right to know what information is collected about them, the right to know what is done with that information, the right to opt out of allowing a business to sell their information (without retaliation), and, in some cases, the right to delete personal information by request.[113] In order to assess the current state of laws protecting SDH and identify potential gaps that could put individual rights at risk, a comparison of the California CPPA and GDPR follows.

---

109. Anton Vedder, *Why Data Protection and Transparency are not Enough when Facing Social Problems of Machine Learning in a Big Data Context, in* BEING PROFILED: COGITAS ERGO SUM: 10 YEARS OF PROFILING THE EUROPEAN CITIZEN 42 (Emre Bayamlioğu et al. eds., 2018).

110. *See, e.g.*, Consumer Online Privacy Rights Act, S. 2968, 116th Cong. (2019) [hereinafter COPRA]; Data Care Act, S. 2961, 116th Cong. (2019); Mind Your Own Business Act, S. 2637, 116th Cong. (2019); Data Breach Prevention and Compensation Act, H.R. 2545, 116th Cong. (2019); Corporate Executive Accountability Act, S. 1010, 116th Cong. (2019); American Data Dissemination Act, S. 142, 116th Cong. (2019); Social Media Privacy and Consumer Rights Act, S. 189, 116th Cong. (2019); Protecting Personal Health Data Act, S. 1842, 116th Cong. (2019).

111. Anita Burgun et al., *Health Data for Public Health: Towards New Ways of Combining Data Sources to Support Research Efforts in Europe,* 26 IMIA Y.B. OF MED. INFORMATICS 235, 239 (2017).

112. Cal. Civ. Code § 1798.100 (West 2020).

113. *Id.*

A. United States

In the U.S., privacy laws address the collection, disclosure, and use of information, but offer significantly less regulation regarding data use.[114] This area of law is complicated by a sectoral approach whereby various U.S. industries (e.g., health, finance, education) have their own separate laws and anonymized data is largely free from regulation. [115] As a result, case law addressing privacy is also compartmentalized and issues around the collection and use of big data remain a relatively recent phenomenon for the courts. This is a problem because modern big data analytics can take siloed data from different industries, and, through various algorithms, draw accurate conclusions about people. SDH data provides a good example of this phenomenon. Although the indicators of health come from a number of separate data points, data about one's finances, living situation, and social network could be combined to develop an overall risk score similar to the "social credit score" utilized in China—even though the laws governing the collection and use of each piece of data are different.

As Katherine Strandburg described, privacy law related to data collection in the U.S. has morphed into an inadequate notice and consent system.[116] Likewise, the Federal Trade Commission requires companies repurposing data to provide notice and obtain consumer consent and, as a result, corporations have responded with lengthy yet vague privacy policies that leave the consumer with many questions about precisely how their data is being used.[117] Historically, this may have been adequate when the information was used solely for advertising or similar purposes but when SDH data is collected and used to classify people by risk, simple consent becomes wholly insufficient. Not only are privacy policies unyieldingly lengthy, but potential uses are couched in vague terms and advanced vocabulary, and notifiers do not have to identify specific third parties that might subsequently gain access to their data.[118] Such unintended uses of personal data have resulted in an uptick in cases filed by consumers over alleged privacy violations and improper use of data in the last several years. However, it can be difficult for plaintiffs to establish

---

114. Paul Ohm, *Changing the Rules: General Principles for Data Use and Analysis, in* PRIVACY, BIG DATA, AND THE PUBLIC GOOD: FRAMEWORK FOR ENGAGEMENT 96, 97-99 (Julia Lane et al. eds, 2014).

115. *Id.* at 98, 103.

116. Strandburg, *supra* note 105, at 8.

117. *Id.*

118. 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; *see also* Thomas Calver et al., *What Tech Giants Really Do with Your Data*, BBC NEWS (July 5, 2018), https://www.bbc.com/news/business-44702483.

Article III standing in federal courts for privacy violations due to the injury-in-fact requirement[119] and the well-established limitation preventing Congress from abrogating this requirement.[120]

In *Spokeo*, the U.S. Supreme Court emphasized that the injury-in-fact requirement needed to demonstrate standing requires a particularized *and* concrete injury, but admitted an intangible injury may still be concrete so long as there is a "risk of real harm."[121] On remand, the Ninth Circuit concluded that potential lost employment opportunities and anxiety due to inaccurate information provided in a credit report in violation of the Fair Credit Reporting Act were sufficiently concrete harms that satisfied Article III standing.[122] This idea that "*some* statutory violations alone do establish concrete harm"[123] has been adopted by a number of other circuits,[124] and privacy torts, in particular, "do not always require additional consequences to be actionable."[125] In the case of SDH data, consumers may be able to assert identifiable harms resulting from privacy violations in the collection of such data, but it will depend on the nature of the allegedly violated statute and the harm asserted.

Recent federal cases have certified class actions in a number of district courts that challenge companies' collection and dissemination of data. A federal judge in the Southern District of New York certified a class action lawsuit against Corelogic for selling allegedly inaccurate credit reports in violation of the Fair Credit Report Act (as well as other statutes).[126] In Virginia, the U.S. District Court for the Eastern District took similar action with a class action suit against Equifax for its unauthorized data collection policies in credit reporting.[127]

---

119. *See* Spokeo, Inc. v. Robins, 136 S. Ct. 1540, 1547 (2016) (citing Lujan v. Defenders of Wildlife, 504 U.S. 555, 560-561 (1992) (holding that the constitutional minimum to establish standing requires that the plaintiff "(1) suffered an injury in fact, (2) fairly traceable to the challenged conduct of the defendant, and (3) likely to be redressed by a favorable judicial decision.")).

120. Spokeo, 136 S.Ct. at 1548 (citing Gladstone, Realtors v. Village of Bellwood, 441 U.S. 91, 100).

121. Spokeo, 136 S.Ct. at 1549.

122. Robins v. Spokeo, Inc., 867 F.3d 1108, 1114–17 (9th Cir. 2017).

123. *Id.* at 1113 (emphasis in original).

124. *See* Strubel v. Comenity Bank, 842 F.3d 181 (2nd Cir. 2016) (a plaintiff's assertion that certain notice disclosures violated the Truth in Lending Act alleged a sufficiently particular and concrete injury-in-fact to her informed use of credit); Dreher v. Experian Info. Sols., Inc., 856 F.3d 337 (4th Cir. 2017) (intangible informational injuries can be sufficiently concrete where they result in an actual harm but inaccurately listing a source of information does not amount to such an injury); Lyshe v. Levy, 854 F.3d 855 (6th Cir. 2017) (bare procedural violations can amount to concrete injuries-in-fact, but an alleged procedural violation not required under the Fair Debt Collection Practices Act does not).

125. Eichenberger v. ESPN, Inc., 876 F.3d 979, 983 (9th Cir. 2017).

126. Feliciano v. Corelogic Rental Prop. Sols., LLC, 332 F.R.D. 98 (S.D.N.Y 2019).

127. Soutter v. Equifax Info. Servs., LLC, 307 F.R.D 183 (E.D. Va. 2015).

Facebook's collection of users' call and text logs unbeknownst to users is the subject of a class action lawsuit in the United States District Court of Northern California.[128] In the same court, Facebook also faces litigation related to the Cambridge Analytica scandal wherein plaintiffs are asserting a variety of privacy claims under California tort law, negligence, breach of contract, and other causes of action.[129] Massachusetts has launched a similar investigation into Facebook's privacy policies in the wake of Cambridge Analytica.[130] The United States District Court of Northern California will also hear a class action against Disney and other app makers for the unauthorized collection of behavioral data for the purposes of profit allegedly in violation of California and Massachusetts laws (the tort of 'intrusion upon seclusion') as well as New York, California, and Massachusetts consumer protection laws.[131] Some of the information allegedly obtained in these cases (e.g., location data, device data, fingerprint data, responses to advertisements, name, and gender) can be used as SDH data and, in combination, are the types of information some assert can be used to re-identify a single individual even if anonymized.[132] It is notable that in addition to class actions, state and federal officials have been increasingly willing to pursue a variety of legal actions against big tech companies.[133]

With respect to claims around the use of data, state and federal antidiscrimination laws may be invoked where applicable. Recently,

---

128. Olin v. Facebook, Inc., No. 18-cv-01881-RS (TSH), 2019 U.S. Dist. LEXIS 195155 (N.D. Cal. Nov. 7, 2019).

129. In re Facebook, Inc., 402 F. Supp. 3d 767 (N.D. Cal. 2019).

130. Nate Raymond, *Facebook Fights Disclosing App Records in Massachusetts Privacy Probe*, REUTERS (Nov. 7, 2019), https://www.reuters.com/article/us-facebook-privacy-massachusetts/facebook-fights-disclosing-app-records-in-massachusetts-privacy-probe-idUSKBN1XH2WC.

131. McDonald v. Aps, 385 F. Supp. 3d 1022 (N.D. Cal. 2019).

132. *See* Gina Kolata, *Your Data Were 'Anonymized'? These Scientists Can Still Identify You,* N.Y. TIMES (July 23, 2019), https://www.nytimes.com/2019/07/23/health/data-privacy-protection.html; *see also* de Montejoy et al., *supra* note 97.

133. *See* Press Release, Xavier Becerra, Attorney General, State of California, *Attorney General Becerra Petitions Court to Compel Facebook to Comply with Outstanding Investigative Subpoena Issued by California* (Nov. 6, 2019), https://oag.ca.gov/news/press-releases/attorney-general-becerra-petitions-court-compel-facebook-comply-outstanding; *see also* Jim Brunner, *Facebook, Google to Pay Washington $450,000 to Settle Lawsuits Over Political-Ad Transparency*, THE SEATTLE TIMES (December 18, 2018), https://www.seattletimes.com/seattle-news/politics/facebook-google-to-pay-washington-450000-to-settle-lawsuits-over-political-ad-transparency; THE ASSOCIATED PRESS, *Justice Department Launches Antitrust Probe of Big Tech*, FORTUNE (July 23, 2019), https://fortune.com/2019/07/23/justice-department-antitrust-probe-big-tech; Press Release, Letitia James, Attorney General, State of New York, *Attorney General James Gives Update on Facebook Antitrust Investigation* (Oct. 22 2019), https://ag.ny.gov/press-release/2019/attorney-general-james-gives-update-facebook-antitrust-investigation.

some national advocacy groups and state attorneys general have had success following this strategy. In 2019, Facebook agreed to a settlement requiring them to overhaul targeted marketing ads for housing following a suit by the National Fair Housing Alliance, ACLU, and Communications Workers of America for alleged violations of the Fair Housing Act.[134] The complaint asserted that housing advertisers could use SDH data, including age, zip code, family size, gender, and even ethnicity, to filter their ads and target select groups they wanted to buy or rent their properties.[135] Facebook agreed to remove zip code, gender, and age targeting options, as well as "direct descriptors of, or semantically or conceptually related to, a person or group of people based on Protected Classes."[136] This agreement is limited to housing advertisements (with some exceptions), while a broader agreement to discontinue targeting ad options based on ethnicity was made by Facebook with respect to advertisers of housing, employment, credit, and insurance or public accommodations in the state of Washington.[137]

B. Approaches Taken by the European Union and California

The GDPR, grounded in the right to privacy granted by the Charter of Fundamental Rights of the European Union[138] and the Treaty on the Functioning of the European Union,[139] went into force in 2018 and provides a number of protections around the use of the personal data of natural persons in the EU.[140] Personal data is defined as "any information relating to an identified or identifiable natural person," including things such as "a name, an identification number, location data, [and] an online identifier," but also "one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person."[141] This definition has many similarities to California's new law, although it might be argued that the language of the California law is broader in

134. Press Release, National Fair Housing Alliance, *Fair Housing Groups Settle Historic Lawsuit with Facebook: Transforms Facebook's Ad Platform Impacting Millions of Users* (March 19, 2019), https://nationalfairhousing.org/2019/03/18/national-fair-housing-alliance-settles-lawsuit-with-facebook-transforms-facebooks-ad-platform-impacting-millions-of-users.

135. Complaint, Nat'l Fair Hous. All. v. Facebook, Inc., No. 18 Civ. 2689 (S.D.N.Y Mar. 27, 2018).

136. *Id.* at 40–41.

137. Assurance of Discontinuance at 3–4, In re Facebook, Inc., No. 18-2-18287-5 (Wash. Super. Ct. July 24, 2018).

138. Charter of Fundamental Rights of the European Union, 2000 O.J. (C 364) 10.

139. Consolidated Version of the Treaty on the Functioning of the European Union, Oct. 26, 2012 O.J. (C 326) 47.

140. GDPR, *supra* note 9.

141. *Id.* art. 4.

that it includes "information that identifies, relates to, describes, is capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household."[142] The California law also provides additional examples and specific covered categories, including biometric information, educational information, network activity information, purchase history, as well as "audio, electronic, visual, thermal, olfactory, or similar information."[143] However, the GDPR also expressly restricts the processing of sensitive information, including biometric and health data that divulge a person's identity, as well as personal data that will reveal information like race, ethnic origin, and political beliefs.[144] Additionally, recent European Court of Justice decisions have confirmed that Article 9's open-ended definition of personal information is quite broad and includes medical injuries (health data) and an employee's work time, but not their dynamic IP addresses.[145]

Those making decisions about the processing of personal data (so-called data controllers), as well as those doing the actual processing (data processors) must abide by the GDPR principles of lawfulness, fairness, and transparency. This is not the case under the California law, which only applies to businesses that collect consumers' personal information,[146] and to businesses that sell consumers' personal information to third parties.[147] Although the GDPR applies to data controllers and data processors of EU residents' information around the world, fewer businesses trigger the California law. That law only applies to those that process California residents' information and either: (1) have annual gross revenues above $25,000,000; (2) buy, sell, or share for commercial purposes the personal information of 50,000 or more consumers, households, or devices; *or* (3) obtain more than 50% of their annual revenue from selling consumers' personal information.[148] Because of the California law's narrow scope, concerns that it does not apply to enough data controllers and data processors are warranted.

142. Cal. Civ. Code § 1798.140(o)(1) (West 2020).

143. *Id.*

144. GDPR, *supra* note 9, at art. 9.

145. W. Gregory Voss et al., *Personal Data and the GDPR: Providing a Competitive Advantage for U.S. Companies*, 56 AM. BUS. L.J. 287, 316–21 (2019).

146. Cal. Civ. Code § 1798.140(c)(1) (West 2020).

147. Cal. Civ. Code § 1798.120 (West 2020).

148. Cal. Civ. Code § 1798.140(c)(1) (West 2020); *see also* Geert Somers et al., *The California Consumer Privacy Act and the GDPR: Two of a Kind?*, FINANCIER (November 2018), https://www.financierworldwide.com/the-california-consumer-privacy-act-and-the-gdpr-two-of-a-kind#.XfEBoOhKi70 (describing the major differences between the two regulations).

VI. LOOKING FORWARD: BROAD APPLICATION OF A GDPR "PLUS"

Given the significant financial incentives to collect SDH data, citing ethical codes and hoping that corporations will self-monitor is naïve and inadequate.[149] Exploitation of such data can potentially create societies where citizens are excessively monitored and discriminated against based on their so-called risk. Importantly, SDH data fluctuates[150] and it is particularly problematic to base formative decisions like one's access to the labor or housing market on transient markers. Application of the GDPR in the U.S. would allow continuity between two of the world's largest economies and elevate American consumers' privacy protection under the law. Even so, it would not entirely address some of the issues surrounding the collection and use of SDH data described above, including the increased surveillance of citizens through the use of discriminatory data analytics and inappropriate data amalgamation.[151]

It is already the case that Facebook, Google, Microsoft, and Amazon all operate in Europe and have been required to comply with the GDPR since 2018, at least as it applies to EU residents. Each of these companies took different approaches to become GDPR compliant and both Facebook and Google were immediately sued when the regulation went into effect.[152] A final resolution of these cases is still pending, but in the meantime, Google was fined €50 million by France's data protection regulator in 2019 for having inadequate consent procedures and the company faces several other investigations for its use of location tracking data.[153] All that aside, the GDPR has initially benefitted the larger tech companies that are able to afford to make the required changes or pay the fines associated

149. Ben Wagner, *Ethics as an Escape from Regulation: From 'Ethics Washing' to Ethics Shopping?* in BEING PROFILED: COGITAS ERGO SUM: 10 YEARS OF PROFILING THE EUROPEAN CITIZEN 84 (Emre Bayamlioğu et al. eds., 2018).

150. Andrea C. Maciejewski, *Medical Records and Privacy Rights: The Unintended Consequences of Aggregated Data in Electronic Health Records*, 90 U. COLO. L. REV. 1111, 1129 (2019).

151. Howard Yu, *GDPR Isn't Enough to Protect Us in an Age of Smart Algorithms*, THE CONVERSATION UK (May 29, 2018), http://theconversation.com/gdpr-isnt-enough-to-protect-us-in-an-age-of-smart-algorithms-97389.

152. Sean Keane, *GDPR: Google and Facebook face up to $9.3B in Fines on First Day of New Privacy Law*, CNET (May 25, 2018), https://www.cnet.com/news/gdpr-google-and-facebook-face-up-to-9-3-billion-in-fines-on-first-day-of-new-privacy-law.

153. Jon Porter, *Google Fined €50 Million for GDPR Violation in France*, THE VERGE (January 21, 2019), https://www.theverge.com/2019/1/21/18191591/google-gdpr-fine-50-million-euros-data-consent-cnil.

with noncompliance, but the long-term outcome remains to be seen.[154] While this early pattern may have been an unintended consequence, it is not consistent with international human rights standards to allow smaller companies to violate consumers' right to privacy and equal treatment simply because they cannot afford to comply with the applicable law. In light of the fact that the GDPR is the governing data protection regulation in the world's largest economy, has already been broadly adopted, requires U.S.-type standing to sue, and permits monetary fines as remedies for violations, it is likely the most feasible option to achieve a minimum standard for privacy rights with respect to personal data, including SDH data.

However, the GDPR does not address more recent problems like the increased use of big data to monitor citizens, the implementation of data analysis techniques that do not treat socioeconomic groups equally, and the inappropriate merging of data sets. These issues must be included in any new federal legislation to prevent unwarranted data collection and discrimination. These issues are of particular concern given the EU's new data strategy,[155] under which data, including anonymized SDH data, could be made public and entered into a so-called single data market. A similar single market for health data was proposed in 2018 with the idea of promoting increased patient access to data and continuity of care, but privacy issues appeared to be an afterthought.[156] The 2020 data strategy mentions the vast amounts of data generated by IoT (Internet of things) devices, the accompanying significant security concerns, and a call to improve consumer tools to manage their own data, but fails to note that the GDPR does not adequately address surveillance of citizens, discriminatory data analytics, and re-identification of individuals through insufficiently anonymized data.[157]

One of the most recent iterations of a data privacy protection bill in the U.S. is the Consumer Online Privacy Rights Act (COPRA), introduced in late 2019, which, among other things, would require covered entities using algorithms (or helping other companies to use them) to conduct annual impact assessments where algorithmic decisions are used for educational, housing, credit, and employment advertising or eligibility decisions.[158] The law would also require impact assessments where algorithms are used to restrict access to

154. Nick Kostov et al., *GDPR has Been a Boon for Google and Facebook*, WALL ST. J. (June 17, 2019), https://www.wsj.com/articles/gdpr-has-been-a-boon-for-google-and-facebook-11560789219?mod=rsswn.

155. *Commission Communication on A European Strategy for Data,* COM (2020) 66 final (Feb. 19, 2020).

156. *Commission Communication on Enabling the Digital Transformation of Health and Care in the Digital Single Market; Empowering Citizens and Building a Healthier Society*, COM (2018) 126 final (Apr. 25, 2018).

157. *Id.*

158. COPRA, *supra* note 110.

places of public accommodation[159] and these assessments would be used to assess discriminatory impact. This concept is particularly important for reasons described earlier in this Article, but in all likelihood, the tech industry will eventually find a workaround. Instead, providing citizens with data rights under federal law and, thereby the standing to sue for their violation, could address not just discriminatory algorithms today, but also technology developed in the future that might result in similar negative consequences.

In addition to adopting the GDPR, the most sustainable option for the U.S. to ensure citizens' fundamental human rights to privacy and freedom from discrimination would be to a) recognize that they are fundamental and b) apply them to technology. Martin Tisne suggests a Bill of Data Rights that guarantees citizens the right to be free from being unreasonably surveilled, from having their data surreptitiously monitored, and from being discriminated against as a result of data.[160] Similarly, implementation of the technological due process framework suggested by Danielle Citron would provide notice to citizens and a means of reviewing biased algorithmic data frequently utilized by state and federal governments.[161] Codification of such rights would be consistent with the rights-based approach of the GDPR and would provide more specific means by which citizens could seek relief through the courts. This combined human-focused methodology is in line with Hartzog and Richards' suggestion of addressing the areas of data protection typically not mentioned: "power, relationships, abusive practices, and data externalities."[162]

This would allow for broader consumer protection and reduce the need to continuously redraft bills in a futile attempt to keep up with technology. Such an approach would also diminish the siloed nature of privacy protection and harmonize legal safeguards for information like SDH data which may have different classifications (e.g., health, commercial) across settings. Regulation should take place at the federal level because it should not be the case that a citizen of one state that regulates the collection of SDH data can lose this protection upon crossing into another state that does not. Given the power of data stored on cell phones and American mobility, leaving SDH data regulation in the hands of states alone is unrealistic and inadequate. This is particularly so given that SDH data do not fall under any specific protection under U.S. law like medical,

159. *Id.*

160. Martin Tisne, *It's Time for a Bill of Data Rights*, MIT TECH. REV. (December 14, 2018), https://www.technologyreview.com/s/612588/its-time-for-a-bill-of-data-rights.

161. Citron, *supra* note 74, at 57.

162. Woodrow Hartzog et al., *Privacy's Constitutional Moment and the Limits of Data Protection*, 61 B.C. L. REV. 1687, 1694 (2020).

educational, or financial data do—even though medical, educational, and financial decisions are being made using SDH data.

VII. CONCLUSION

With the ever-increasing pervasiveness of data in our everyday lives, it comes as no surprise that the reach and impact of SDH data has increased in kind. Already we have seen the rapid expansion of the use of SDH data into the insurance, healthcare, marketing, housing, and financial sectors. Although the negative impact of the use of SDH data may not be readily apparent to all consumers today, it is impossible to predict all the additional uses that businesses will find in the coming years, or what the consequences to society might be. As outlined in this paper, there are numerous weaknesses in the current governance of SDH data that need to be urgently addressed to define a safe, fair, and transparent data ecosystem for consumers and businesses. Unless properly controlled, opening the Pandora's box that is the unfettered collection of SDH data will have negative consequences for society through cybersecurity vulnerabilities, discriminatory practices, human rights violations, and supply chain blind spots in the procurement, transfer, and use of data. As is often the case with technological revolutions, the governance of SDH data has not kept pace with the speed of the industry's exploitation. To this point, the laws governing SDH data come from disparate regulatory frameworks that do not provide a clear unified strategy for communication to businesses, do not provide an enforcement mechanism, are not transparent to consumers, and do not evolve to keep pace with technology.

Federal adoption of the GDPR in the U.S. may be more feasible now that California's Consumer Privacy Protection Act has gone into effect, triggering many American businesses to make big changes. The large number of Senate bills introduced on this topic in recent years also highlight both the need and the desire of American citizens for new federal legislation. However, a federal rule needs to be in place as it is not practicable to have 50 separate state laws for businesses that regularly cross state lines. It is also inadequate to only have federal adoption of the GDPR as biased algorithms, unmonitored merging of data sets, and unwarranted citizen surveillance are real threats to the fundamental human rights of privacy and freedom from discrimination. Therefore, to address the health and privacy concerns around the collection and use of SDH data, the implementation of a data rights bill in addition to the adoption of the GDPR should be considered.