8-2022

Data Privacy in the Time of Plague

Cason Schmit

Brian N. Larson

Hye-Chung Kum

Data Privacy in the Time of Plague

Cason Schmit, Brian N. Larson & Hye-Chung Kum[*]

Abstract:

Data privacy is a life-or-death matter for public health. Beginning in late fall 2019, two series of events unfolded, one everyone talked about and one hardly anyone noticed: The greatest world-health crisis in at least 100 years, the COVID-19 pandemic; and the development of the Personal Data Protection Act Committee by the Uniform Law Commissioners (ULC) in the United States. By July 2021, each of these stories had reached a turning point. In the developed, Western world, most people who wanted to receive the vaccine against COVID-19 could do so. Meanwhile, the ULC adopted the Uniform Personal Data Protection Act (UPDPA) at its annual meeting, paving the way for state legislatures to adopt it beginning in 2022. It has so far been introduced in three jurisdictions.

These stories intersect in public health. Public health researchers struggled with COVID-19 in the United States because they lacked information about individuals who were exposed, among other matters. Understanding other public health threats (e.g., obesity, opioid abuse, racism) also requires linking diverse data on contributing social, environmental, and economic factors. The UPDPA removes some barriers to public health practice and research resulting from the lack of comprehensive federal privacy laws. Its full potential, however, can be achieved only with involvement of public health researchers and professionals. This article analyzes the UPDPA and other comprehensive state privacy statutes, noting the ways that they could promote—and hinder—public health. It concludes with recommendations for public health researchers and professionals to get involved in upcoming legislative debates on data privacy. Lives will depend on the outcomes.

[*] Cason Schmit, JD, Assistant Professor of Public Health, Texas A&M University; Brian Larson, JD, PhD, Associate Professor of Law and Arts & Humanities Fellow, Texas A&M University; Hye-Chung Kum, MSW, PhD, Professor of Public Health and Director, Population Informatics Lab, Texas A&M University. For comments on earlier versions of this article, we are deeply grateful to Professors William Henning, Hannah Bloch-Wehba, and Brendan S. Maher, and Denise Chrysler from the Network for Public Health Law, and to anonymous peer reviewers and staff at *Yale Journal of Health Policy, Law & Ethics*. Professor James G. Hodge, Jr., provided valuable feedback on an early draft/concept of this work. We are grateful to David Morrison for his able research assistance. Funding from Texas A&M University's Triads for Transformation (T3) program supported this project.

INTRODUCTION

It is a commonplace and a cliché in legal scholarship and the broader culture that American data privacy laws are a "patchwork" of solutions to discrete privacy issues that leave significant gaps and open questions about which personal data are subject to protection and to what extent.[1] There is no blanket of privacy law that covers all subjects, types, and users of data. Patches cover some, overlapping in some cases with each other, but in other cases leaving large parts of the body of data uncovered.[2] One impetus for this Article grows from a series of events in 2021 that respond to this patchwork: Adoption by Virginia and Colorado of comprehensive data privacy legislation and approval by the Uniform Law Commissioners (ULC)[3] of the Uniform Personal Data Protection Act (UPDPA).[4] These developments occurred against the backdrop of significant changes to California's 2018 comprehensive privacy act resulting from a 2020 referendum. This Article is the first to our knowledge to critically assess the

[1] *E.g.,* Thorin Klosowski, *The State of Consumer Data Privacy Laws in the US (And Why It Matters)*, N.Y. TIMES (Sep. 6, 2021), https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us/ [https://perma.cc/RG26-2CPC] ("The United States doesn't have a singular law that covers the privacy of all types of data. Instead, it has a mix of laws that go by acronyms and initialisms like HIPAA, FCRA, FERPA, GLBA, ECPA, COPPA, and VPPA"); Brouse McDowell, Craig S. Horbus & Jarman J. Smith, *Corporate TIPS: U.S. Data Privacy Law Patchwork Grows as States Enact New Legislation*, LEXOLOGY (Aug. 18, 2021), https://www.lexology.com/library/detail.aspx?g=e24fedac-cea7-412a-a5eb-5d736276e8d6 [https://perma.cc/RLQ2-J4GE]; Natasha Singer, *An American Quilt of Privacy Laws, Incomplete,* N.Y. TIMES (Mar. 30, 2013), https://www.nytimes.com/2013/03/31/technology/in-privacy-laws-an-incomplete-american-quilt.html [https://perma.cc/G4UU-DLMZ]; Anthony Jones, *Autonomous Cars: Navigating the Patchwork of Data Privacy Laws That Could Impact the Industry*, 25 CATH. U.J.L. & TECH. 180 (2017); Kiran K. Jeevanjee, *Nice Thought, Poor Execution: Why the Dormant Commerce Clause Precludes California's CCPA from Setting National Privacy Law*, 70 AM. U. L. REV. F. 75 (2020); Stephanie Comstock Ondrof, "*Senator, We Run Ads": Advocating for a US Self-Regulatory Response to the EU General Data Protection Regulation*, 28 GEO. MASON L. REV. 815, 819 (2021). The reporters for the American Law Institute describe it instead as an "interrelated amalgam of different types of law," PRINCIPLES OF THE LAW OF DATA PRIVACY § 1 (AM. L. INST. 2019) [hereinafter PRINCIPLES OF DATA PRIVACY], as "a complex aggregation of overlapping and inconsistent laws that represent an increasingly significant compliance burden," and as "sectoral," contrasting with "omnibus" regulatory regimes. *Id.* § 1, cmt. e. (We use "comprehensive" below to refer to "omnibus" regimes.) The reporters for ALI's initiative were leading data-privacy scholars Professors Paul M. Schwartz and Daniel J. Solove. *Id.* at vii.

[2] *See infra* text accompanying note 85.

[3] The group is also known the National Conference of Commissioners on Uniform State Laws. *About Us*, UNIF. L. COMM'N, https://www.uniformlaws.org/aboutulc/overview [https://perma.cc/HD2M-PCDW] (last visited Sept. 14, 2021).

[4] Consumer Data Protection Act, VA. CODE ANN. §§ 59.1-575 to 59.1-585 (2021) (effective Jan. 1, 2023) [hereinafter VCDPA]; Colorado Privacy Act, COLO. REV. STAT. §§ 6-1-104 to 110, 6-1-1301 to 1313 (2021) (effective July 1, 2023) [hereinafter CPA] UNIF. PERS. DATA PROT. ACT (Unif. L. Comm'n 2021) [hereinafter UPDPA]; *see also* California Consumer Privacy Act, CAL. CIV. CODE §§ 1798.100–1798.199.100 (2021) (effective Jan. 1, 2023) [hereinafter CCPA].

UPDPA and the adopted comprehensive acts in California, Virginia, and Colorado—which we refer to as the "CAVACO statutes"—side by side. This analysis is timely, as the UPDPA has already been introduced in three U.S. jurisdictions as of February 2022 and may prove an influential model for state privacy law.

Personal data also play a critical role in public health interventions and research, and a second impetus for this Article grows from public health crises that have rocked the United States in 2020–21 and the need for researchers to have access to so-called "big data" to address these crises. Talk of COVID-19 has been ubiquitous in the media, of course, but a second set of newsworthy events highlights other equally pernicious public health crises: racism and health risks associated with the poverty that disproportionately afflicts persons of color in the United States. Furthermore, media coverage of these crises has overshadowed other persistent and growing public health threats, like obesity, opioid abuse, homelessness, climate change, and mental health. These crises plague America, and data privacy legislation holds the potential to make ameliorating them less—or more—difficult.

As a preliminary matter, data protection laws raise particular concerns for promoting public health. Readers might wonder why these statutes are of concern to public health researchers and professionals. After all, many public health agencies are arms of local and state governments, and the UPDPA and the CAVACO statutes exclude government agencies from their coverage. The point is well taken, but it does nothing to allay concerns of public health researchers who may be affiliated with private institutions. Furthermore, the key challenge here relates to "secondary uses." Primary uses are those that permit us to live in the digital world, the very uses for which the data are collected. Secondary uses are those where data are collected for one purpose and reused for a different purpose, particularly where private entities gather data for business purposes and public health researchers and practitioners seek access to those data for public health purposes.

There are various ways that personal data—not just health data—can be used to improve public health.[5] Of course, there is research for scientific purposes. University and non-profit researchers want data to understand if two things are related; for example, whether a public-health initiative—perhaps a "nudge" for consumers to choose to donate their organs[6]—is effective at achieving its goals. They also want to learn about how the world works; how poverty and racism relate to disease, for example. Research based on secondary use of existing data is much cheaper than research that requires collecting new data from individuals,

[5] For more details, *see* Part I(A).

[6] RICHARD H. THALER & CASS R. SUNSTEIN, NUDGE 254 (Final ed. 2021).

and it permits using sizable longitudinal datasets accumulated over time. In addition, some research is not possible without re-using existing data. Researchers and public health professionals can also make secondary use of aggregated data to promote population health and well-being. For example, personal data about health diagnoses and outcomes can be linked to other data to understand the cause of injuries, diseases, or poor health and to help officials develop prevention strategies.

Personal data can also be used for interventions that seem less benign. For example, an employer could use data about employees to change their health insurance premiums based on whether the employees have been vaccinated against a disease.[7] The government could use contact-tracing information regarding a pandemic illness to identify carriers and potential carriers and impose isolation or quarantine orders.[8] As we explain below, these examples highlight differences between using data for what is often called "human subjects research" and for public health interventions. The Common Rule, the regulations for research using human subjects, which is supported by twenty federal agencies, governs research on human subjects in many settings.[9] An Institutional Review Board (IRB) that "has been formally designated to review and monitor" research generally supervises such research projects.[10]

The UPDPA and the other state acts apply to most such secondary data practices, so understanding *how* they do so is critical. They may have an especially significant potential to affect the use of personal data for public health interventions and research. For that reason, an evaluation of the UPDPA and the

[7] Niraj Chokshi, Margot Sanger-Katz & Tara Siegel Bernard, *Delta's Extra $200 Insurance Fee Shows Vaccine Dilemma for Employers*, N.Y. TIMES (Aug. 26, 2021), https://www.nytimes.com/2021/08/26/business/delta-insurance-fee-unvaccinated.html [https://perma.cc/ZY8B-7ECT].

[8] *See Frequently Asked Questions: Contact Tracing*, CTRS. DISEASE CONTROL, https://www.cdc.gov/coronavirus/2019-ncov/faq.html#:~:text=Discussions%20with%20health%20department%20staff,or%20local%20health%20department [https://perma.cc/8JZN-MAZX] (last visited Sept. 18, 2021).

[9] Federal Policy for the Protection of Human Subjects, 82 Fed. Reg. 7149 (Jan. 19, 2017) (to be codified at 6 C.F.R. Part 46 and several other points) [hereinafter Common Rule]. The Federal Food and Drug Administration regulations that govern human subjects research are also highly similar to the Common Rule regulations. 21 C.F.R. § 50.1 (2020).

[10] U.S. FOOD & DRUG ADMIN., *Institutional Review Boards (IRBs) and Protection of Human Subjects in Clinical Trials* (Sept. 11, 2019), https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/institutional-review-boards-irbs-and-protection-human-subjects-clinical-trials [https://perma.cc/GFU5-AJPZ]. Universities and similar research institutions typically have their own IRBs and subject most of their research to their supervision. *See, e.g.*, Committee on the Use of Human Subjects, HARVARD UNIVERSITY, https://cuhs.harvard.edu/ [https://perma.cc/67AR-2M69] (last visited Feb. 12, 2022); Institutional Review Boards (IRBMED), U. OF MICH. MED. SCH., https://research.medicine.umich.edu/our-units/institutional-review-boards-irbmed [https://perma.cc/XQJ2-B8CH] (last visited Feb. 12, 2022).

other state comprehensive acts from the perspective of public health is particularly important now. To our knowledge, this Article is the first to closely examine the effect on public health of any state comprehensive privacy statute, including California's now-four-year-old law.

Part I introduces the present landscape in public health and U.S. data protection law, considering both the existing laws and some proposals for reorganizing and reimagining the privacy paradigm within American law. We explain why we focus the balance of this Article on the "notice and choice" paradigm that is evident in the existing legislation. Our position is that an ethics of data privacy should focus on the autonomy of data subjects, their ability to know of and consent to data practices to which their personal data are subjected. At the same time, given that certain "defaults" are at play in modern consent processes—click-through privacy policies and the like—regulators should establish consent defaults that favor some secondary uses of personal data in line with public interests and preferences, uses that minimize social harms and maximize community benefits, including uses for public health and research.

Part II provides a conceptual framework for data protection law in the "notice and choice" paradigm. It defines terms and identifies important characteristics of any data-protection regime, providing an extension of existing conceptual frameworks, such as the American Law Institute's *Principles of the Law of Data Privacy*.[11] Part II analyzes the UPDPA and CAVACO statutes using this conceptual framework. The detailed analysis is essential for privacy-law theorists, legislators, and groups interested in proposed privacy legislation that is being deliberated today.[12]

Part III assesses the UPDPA and CAVACO statutes against the normative frameworks previously discussed and recommends ways in which public health researchers and professionals may wish to intervene in coming months and years in the deliberations on data protection statutes. As Table 1 shows, the Colorado Privacy Act is the most supportive of public health practices and research, exempting a wide swath of them from its coverage and permitting most others without the necessity of disclosing them to data subjects. Some ethicists might go as far as to say it is *too* friendly to public health because of this lack of disclosure, and we'd agree. The California Consumer Privacy Act broadly supports research, but generally requires that those collecting data from consumers for public health activities, like public health surveillance and interventions, must disclose the practices and give data subjects the chance to opt out. This most closely fits the normative frameworks we outline below. The

[11] PRINCIPLES OF DATA PRIVACY, *supra* note 1.

[12] Part II cannot claim, however, to provide a comprehensive analysis of all aspects of these acts.

UPDPA raises a concern regarding the need for data subjects to opt in for uses of sensitive personal data—the kind of data often at issue in public health practices and research. Finally, the Virginia Consumer Data Protection Act requires consumers to opt in for almost all public health data practices, which could gravely impair public health activities subject to that act. We propose that public health researchers and professionals should seek to amend the Colorado and Virginia acts and should seek to revise the UPDPA as it is adopted in states to conform them to the normative frameworks we provide. We offer other suggestions as well.

In theory, a comprehensive privacy law is a smooth blanket, covering all circumstances while permitting appropriate socially desirable and beneficial uses, like those for research and public health. Our review of the UPDPA and CAVACO statutes shows that they do privilege some public health activities, particularly generalizable research, but that public health professionals must involve themselves actively in legislative and regulatory activity surrounding future adoption of such acts to improve them and to ensure that legislators and regulators do not forget public health in their rush to protect private data. A comprehensive data protection framework should provide a protective blanket unmarred by patchwork holes—not merely a sheet to cover the bodies of the dead.

I. THE CONTEMPORARY PUBLIC HEALTH AND LEGAL LANDSCAPES

Analyzing and evaluating the UPDPA and the CAVACO statutes requires some background in the public health and legal landscapes in the United States. This includes a basic understanding of public health practices, an overview of U.S. data privacy and protection law, and a discussion of normative concerns at the boundaries of these two disciplines.

A. Public Health Research and Practices

Public health, as both a science and a practice, is data driven. Data inform epidemiologists about the nature of disease and conditions that affect health. These data can help public health practitioners understand whether a disease spreads through air, touch, bodily fluids, animal contact, or consumption of tainted food.[13] Data can also help build an understanding of how social and environmental factors—such as walkable communities, food deserts,

[13] *See* J.A. Magnuson et al., *Informatics in Disease Prevention and Epidemiology, in* PUBLIC DISEASE HEALTH INFORMATICS AND INFORMATION SYSTEMS 239, 239–57 (J.A. Magnuson & Brian E. Dixon, eds., 3d ed. 2020), https://link.springer.com/chapter/10.1007/978-3-030-41215-9_14 [https://perma.cc/BKU5-BPX4] (describing public health informatics and disease investigation generally).

environmental contamination, economic inequities, and structural racism—affect health.[14] We can divide the activities that use these data into public health research, which seeks generalized knowledge; surveillance, which monitors health data to enable and assess interventions; community interventions or health programs designed to improve population health; and individual interventions, intended to serve at-risk individuals or protect the rest of the population from them.

The COVID-19 pandemic exposed the limitations of the traditional public health system, as it was unable to acquire, ingest, and share the unprecedented volumes of data needed to understand and control a rapidly spreading virus.[15]

1. Public Health Research

The field of public health is grounded in scientific evidence. This body of evidence includes, but is not limited to, microbiology, physiology, sociology, and policy research.[16] Public health research aims to generalize the results from a

[14] Sandro Galea et al., *Estimated Deaths Attributable to Social Factors in the United States*, 101 AM. J. PUB. HEALTH 1456, 1462–63 (2011) (estimating hundreds of thousands of deaths associated with non-biological factors, including education, racism, and economic inequity); *see also* Paula Braveman & Laura Gottlieb, *The Social Determinants of Health: It's Time to Consider the Causes of the Causes*, 129 PUB. HEALTH REPS. 19, 27 (2014) (describing the difficulty obtaining the cross-sectoral data needed to study social determinants of health).

[15] *See generally* Willem G van Panhuis et al., *A Systematic Review of Barriers to Data Sharing in Public Health*, 14 BMC PUB. HEALTH 1144 (2014); Drew Armstrong, *Data Failures Keep the CDC From Seeing the Whole Picture on COVID*, BLOOMBERG (Dec. 21, 2021), https://www.bloomberg.com/news/articles/2021-12-21/cdc-public-health-data-failures-mean-u-s-lacks-whole-picture-on-covid [https://perma.cc/5LQS-ASVH]; Xenia Shih Bion, *Crumbling Data Infrastructure Undermines Nation's Pandemic Reponse*, CAL. HEALTH CARE FOUND. BLOG, https://www.chcf.org/blog/crumbling-data-infrastructure-undermines-nations-pandemic-response/ [https://perma.cc/ 42G3-NPXT] (last visited Apr. 11, 2022). Many of these deficiencies are due to the three challenges in the U.S. public health system. First, public health in the United States is chronically underfunded, particularly after state and local budget cuts following the 2008 Great Recession. Second, the decentralized U.S. public health system—a product of the Tenth Amendment of the Constitution—imposes legal, political, and relationship barriers between local, state, and federal public health partners seeking to share public health information. *See generally* Panhuis et al., *supra*. Third, many available data that are relevant to public health are subject to restrictive data protection laws. *See generally* Rachel Hulkower, Matthew Penn & Cason Schmit, *Privacy and Confidentiality of Public Health Information*, in PUBLIC HEALTH INFORMATICS AND INFORMATION SYSTEMS 147 (J.A. Magnuson & Brian E. Dixon eds., 3d ed. 2020). However, a comprehensive overview of the challenges facing public health informatics and public health data systems is beyond the scope of this work.

[16] *See generally* PUBLIC HEALTH INFORMATICS AND INFORMATION SYSTEMS (J.A. Magnuson & B.E. Dixon eds., 3d ed. 2020); Evan Anderson et al., *Measuring Statutory Law and Regulations for Empirical Research,* PUBLIC HEALTH LAW RESEARCH: THEORY AND METHODS 237 (A. C. Wagenaar & S. Burris eds., 1st ed. 2013); *see* PUBLIC HEALTH INFORMATICS, *supra*, at 71–73; Braveman & Gottlieb, *supra* note 14, at 27.

discrete study sample in a specified period in time to a broader population.[17] This is in contrast to the *practice* of public health, which involves ongoing efforts to monitor an entire community or population.[18] Public health research includes studies that require data collection (e.g., surveys, environmental sample collection) as well as studies that rely on pre-existing data (e.g., electronic health records).[19] Whenever public health research uses data from identifiable human data subjects, the Common Rule regulations protecting human subjects research will likely apply.[20]

2. *Surveillance*

There are several different types of public health surveillance that help public health professionals understand the threats to population health. Unlike health research, public health surveillance is "the *ongoing*, systematic collection, analysis, and interpretation of health-related data essential to *planning, implementation, and evaluation of public health practice.*"[21] Critically, the ongoing surveillance data-collection activities ensure that public health professionals have current data to inform public health activities. For example, healthcare providers are required by law to report if a patient has one or more conditions of public health concern.[22] These case reports assist public health professionals to understand where a disease is spreading within a community.

Importantly, these ongoing surveillance activities are not research under the Common Rule, so public health agencies can swiftly collect data and respond to public health threats within their statutory capacity without additional regulatory burdens.[23] Consequently, this surveillance information provides critical situational awareness required for deploying scarce public health resources

[17] James G. Hodge & Lawrence O. Gostin, *Public Health Practice vs. Research*, COUNCIL OF STATE & TERRITORIAL EPIDEMIOLOGISTS, May 24, 2004, at 14–20.

[18] *Id.* at 14–21.

[19] H. M. Xu et al., *Lead Concentrations in Fine Particulate Matter After the Phasing Out of Leaded Gasoline in Xi'an, China*, 47 ATMOSPHERIC ENV'T 217, 219–22 (2012) (describing an observed decrease in environmental lead concentrations associated with a decrease in the use of leaded gasoline); Tara I. Chang & Wolfgang C. Winkelmayer, *Comparative Effectiveness Research: What Is It and Why Do We Need It in Nephrology*, 27 NEPHROLOGY, DIALYSIS, TRANSPLANT 2156, 2156–60 (2012) (providing an overview of comparative effectiveness research, which often relies on electronic health records to evaluate the comparative health outcomes associated with different treatment options).

[20] *See supra* Introduction; Common Rule, *supra* note 9.

[21] *Introduction to Public Health Surveillance*, CTRS. DISEASE CONTROL, https://www.cdc.gov/training/publichealth101/surveillance.html [https://perma.cc/65PT-CXVL] (last visited Oct. 5, 2021) (emphasis added).

[22] Public health reporting is typically required by state law and requirements can vary substantively by jurisdiction. *See, e.g.*, N.M. CODE R. § 7.4.3 (LexisNexis 2021).

[23] Hodge & Gostin, *supra* note 17.

efficiently and effectively.[24]

In addition to acute public health threats, social, economic, and environmental factors may have a far greater impact on an individual's health than biological factors.[25] Public health professionals often have access to aggregate data on these factors (e.g., census data), but data records or person-level data—the type needed to link datasets and understand complex problems—are far more difficult to obtain.[26] Data on these social, economic, and environmental factors are nevertheless often abundant in commercial datasets, including data useful to market products and services or to determine things like loan eligibility.[27] Businesses sharing data about social, economic, and environmental factors with public health agencies is a promising but largely unexplored opportunity to better understand threats to public health, and by extension, develop viable interventions to address those threats.[28]

3. Public Health Programs and Population Interventions

Public health practice involves collective actions that assure the conditions for people to be healthy.[29] These actions, whether an ongoing program or new intervention, rely on data to ensure that scarce resources are used efficiently. Consequently, public health programs and interventions require data in the planning phase to determine the most effective deployment of limited resources; they require data throughout implementation to ensure activities are proceeding as intended; and they require data to evaluate whether, and to what extent, the

[24] For example, syndromic surveillance systems can detect symptom-based anomalies in local emergency rooms that can provide public health departments with rapid information of emerging infectious disease (e.g., influenza, anthrax). *See* Deborah W. Gould et al., *The Evolution of BioSense: Lessons Learned and Future Directions*, 132 PUB. HEALTH REPS. 7S, 7S–10S (2017); *see also* Matthias Linden et al., *Case Numbers Beyond Contact Tracing Capacity Are Endangering the Containment of COVID-19*, 117 DEUTSCHES ÄRZTEBLATT INT'L 790, 790–91 (2020) (describing the capacity limitations that hindered the public health response to COVID-19).

[25] *See generally* Galea et al., *supra* note 14, at 1462–63.

[26] Braveman & Gottlieb, *supra* note 14, at 27.

[27] *See generally* CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION (2016); *Id.* at 68–83, 141–60 (describing the often discriminatory and destructive ways that data are used that nonetheless may be profitable to companies).

[28] Mattia Prosperi et al., *Big Data Hurdles in Precision Medicine and Precision Public Health*, BMC MED. INFORMATICS & DECISION MAKING 1, 5-10 (Dec. 29, 2018); Sonja A. Rasmussen et al., *Precision Public Health as a Key Tool in the COVID-19 Response*, 324 JAMA 933, 934 (2020); Cason Schmit et al., *Cross Sector Data Sharing: Necessity, Challenge, and Hope*, 47 J. L., MED. & ETHICS 83, 83 (2019); Braveman & Gottlieb, *supra* note 14, at 27.

[29] This reflects the Institute of Medicine's definition of public health: "Public health is what we, as a society, do collectively to assure the conditions in which people can be healthy." INST. MED., THE FUTURE OF PUBLIC HEALTH (1988). The definition implies a distinction between public health and healthcare. The former focuses on prevention and maintenance of health, the latter treats and mitigates existing ill health.

program or intervention is achieving population health benefits.[30]

For example, during the early deployment of the COVID-19 vaccinations, public health agencies relied on data to determine the most vulnerable sub-populations and used that data (in some cases) to deploy vaccines and set up vaccination sites.[31] Throughout vaccination deployment, public health agencies collected data to determine whether the clinics were indeed serving those vulnerable populations,[32] adjusting strategies as necessary.[33] Finally, public health agencies closely monitored case reports and hospital and mortality data to determine whether the vaccinations were affecting the spread of COVID-19 and its health outcomes.[34]

Increasingly, public health agencies are exploring and leveraging non-traditional public health data to inform population-based interventions. Traditional public health data include mandated case reports of infectious disease (e.g., drug-resistant tuberculosis, HIV), vital statistics, reports of foodborne illness, and other surveillance data.[35] The New York City Department of Health and Mental Hygiene, however, developed a program that scanned publicly available restaurant reviews—like those on Yelp!—for evidence of foodborne illness (e.g., "food made me sick").[36] Using big-data analytics, public health

[30] James Aspevig, *Project Management and Public Health Informatics,* PUBLIC HEALTH INFORMATICS AND INFORMATION SYSTEMS 211, 221–35 (J.A. Magnuson & Paul C. Fu, Jr. eds., 2014); *see also* CASON SCHMIT, PUBLIC HEALTH LAW AND POLICY INNOVATIONS: SOCIAL IMPACT BONDS 2–3, and generally (2014).

[31] *Ensuring Equity in COVID-19 Vaccine Distribution*, HEALTH RES. & SERVS. ADMIN., https://www.hrsa.gov/coronavirus/health-center-program [https://perma.cc/P6ZL-L77K] (last visited Oct. 10, 2021).

[32] As opposed to merely reaching "vaccine tourists." *See* Claire Gillespie, *What is Vaccine Tourism, and Is It Legal? Here's What You Need to Know*, HEALTH (Jan. 28, 2021), https://www.health.com/condition/infectious-diseases/coronavirus/what-is-vaccine-tourism [https://perma.cc/VWW3-KQBM] ("Vaccine tourism means visiting another country or state to get a vaccine not available to you at home.").

[33] *Strategies to Engage Communities Most Vulnerable to Covid-19*, NAT'L ACADS. SCIS. ENG'G MED., https://www.nap.edu/resource/26068/interactive/vulnerable-communities.html [https://perma.cc/8UHH-DLZT] (last visited Oct. 7, 2021); *see also* Megan Cerullo, *State Vaccine Incentives Do Little to Boost Vaccination Rates, Research Shows*, CBS NEWS (Sep. 8, 2021), https://www.cbsnews.com/news/statewide-vaccine-incentives-lotteries-do-not-boost-vaccination-rates/ [https://perma.cc/UVD9-WU5X] (describing evaluations of vaccine incentives).

[34] Dvir Aran, *Estimating real-world COVID-19 Vaccine Effectiveness in Israel Using Aggregated Counts* 1–6 (medRxiv, Working Paper, 2021), https://www.medrxiv.org/content/10.1101/2021.02.05.21251139v3.full.pdf [https://perma.cc/D8C7-FGB4].

[35] John R. Lumpkin & J.A. Magnuson, *History of Public Health Information Systems and Informatics, in* PUBLIC HEALTH INFORMATICS AND INFORMATION SYSTEMS 17–29 (J.A. Magnuson & Paul C. Fu, Jr. eds., 2014).

[36] *See generally* Cassandra Harrison et al., *Using Online Reviews by Restaurant Patrons to Identify Unreported Cases of Foodborne Illness—New York City, 2012-2013*, 441 MORBIDITY & MORTALITY WKLY. REP. 63 (2014); Elaine O. Nsoesie, *Online Reports of Foodborne Illness*

professionals were able to identify previously unreported outbreaks.[37] With this information, they were able to focus their limited enforcement budget only on highly probable events.

4. *Individual-based Interventions*

Prevention is a central focus for public health practitioners. Preventing adverse health outcomes—as opposed to treating those that develop—is often less expensive and leads to better population health.[38] While prevention efforts can target entire communities, such as building sidewalks to promote active living, many preventative interventions require identifying at-risk individuals who stand to benefit the most.[39]

For example, maternal and child health is a critical ongoing public health issue. Prenatal contact with expectant mothers can have a tremendously beneficial effect on birth outcomes and maternal health.[40] Moreover, the benefits can extend far into a family's future.[41] In commercial settings, advanced data analytics can predict whether a customer is pregnant based on changes to purchasing behavior.[42] These predictions are immensely valuable to companies seeking to gain loyal customers at a point when purchasing behavior will change substantially.[43] For public health, this predictive ability can help direct scarce

Capture Foods Implicated in Official Foodborne Outbreak Reports, 67 PREVENTATIVE MED. 264–69 (Aug. 11, 2014).

[37] Harrison et al., *supra* note 36.

[38] *See generally* Thomas R. Frieden, *A Framework for Public Health Action: The Health Impact Pyramid*, 100(4) AM. J. PUB. HEALTH 590 (2010).

[39] *See* Karen A. Monsen et al., *Public Health Nurses Tailor Interventions for Families at Risk*, 28 PUB. HEALTH NURSING 119, 119–21 (Mar.–Apr. 2011); *see generally* R. J. Donovan et al., *TARPARE: A Method for Selecting Target Audiences for Public Health Interventions*, 23-3 AUSTL. & N.Z. J. PUB. HEALTH 280 (June 23, 1999).

[40] In some cases, the benefits of prevention can be leveraged to support profitable investments. *See* ASS'N STATE AND TERRITORIAL HEALTH OFFICIALS, FINANCING PUBLIC HEALTH INTERVENTIONS THROUGH PAY FOR SUCCESS (2017) https://opioidspreparedness.org/Health-Systems-Transformation/Pay-for-Success-South-Carolina-Issue-Brief/ [https://perma.cc/UKZ4-629Y].

[41] *Id.*; Monsen et al., *supra* note 39.

[42] Kashmir Hill, *How Target Figured Out a Teen Girl Was Pregnant Before Her Father Did*, FORBES (Feb. 16, 2012, 11:02 AM), https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/?sh=a82d6c66686d [https://perma.cc/7GP4-NS3T].

[43] Charles Duhigg, *How Companies Learn Your Secrets*, N.Y. TIMES (Feb. 16, 2012) https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html [https://perma.cc/Z5HU-VFC8]. Importantly, the scarcity of resources for public health interventions is a substantial limitation to the public health use of these data. While the ability to identify at-risk individuals can help public health agencies be more efficient with existing resources, these data are less useful when public health practitioners lack the capacity to act. For example, public health surveillance techniques can monitor trends of suicide ideation in near-real time, but if public health agencies

resources to at-risk individuals for programs and benefits.[44] Recent advances in machine learning and artificial intelligence have the capacity to further amplify these benefits but also raise concerns about unacceptable uses.[45] For example, commercial data brokers have increasingly detailed information about individuals that they sell to businesses, individuals, and governments, using artificial intelligence and machine learning tools to identify groups of people with certain health conditions, such as diabetes, HIV, depression, and pregnancy, based on their aggregated consumer data,[46] and enabling businesses to target these individuals with goods or services they might want or need. Certainly, these practices are problematic when they enable exploitation of the vulnerable, but these data can also facilitate interventions that promote social, economic, and health equity.

In public health contexts, it is important to identify and address population health threats, which can span varied domains, including hazardous products, environmental contamination, occupational hazards, infectious disease, law, and policies. The value of non-traditional public health data in advancing these aims is becoming increasingly clear. It might be important to identify individuals with an infectious disease who might pose a risk to others. In the case of sexually transmitted infections, contact-tracing efforts can be essential to identify and notify individuals of this risk.[47] This contact-tracing can enable timely treatment and inform people of the need for precautions.[48] In the COVID-19 pandemic, contact-tracing apps were developed to notify individuals if they were near someone who tested positive for the virus.[49] This information can prompt individuals to get a test to confirm infection and notify them of the need for

lack the financial, human, or political capital to enact preventative interventions, the surveillance data is not useful beyond informing the community of the public health issue. *See generally* Marissa L. Zwald et al., *Monitoring Suicide-Related Events Using National Syndromic Surveillance Program Data*, 11 ONLINE J. PUB. HEALTH INFORMATICS (2019); Deb Stone et al., *Preventing Suicide: A Technical Package of Policy, Programs, and Practices*, CTRS. DISEASE CONTROL (2017) https://www.cdc.gov/violenceprevention/pdf/suicidetechnicalpackage.pdf [https://perma.cc/HMX8-LFMK].

[44] Monsen et al., *supra* note 39, at 119–21.

[45] Alicia Solow-Niederman, *Information Privacy and the Inference Economy*, 117 NW. UNIV. L. REV., (forthcoming 2022) (manuscript at 45), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3921003 [https://perma.cc/YK5P-W3F7].

[46] DATA BROKERS: LAST WEEK TONIGHT WITH JOHN OLIVER, at 5:50 – 8:30, https://www.youtube.com/watch?v=wqn3gR1WTcA [https://perma.cc/7MZT-LBFL].

[47] Megan S. C. Lim et al., *SMS STI: A Review of the Uses of Mobile Phone Text Messaging in Sexual Health*, 19 INT'L J. STD AIDS 287, 288 (May 2008).

[48] *Id.*

[49] Nadeem Ahmed et al., *A Survey of COVID-19 Contact Tracing Apps*, 8 IEEE ACCESS 134577, 134578 (July 31, 2020); *see generally* Vittoria Colizza et al., *Time to Evaluate COVID-19 Contact-Tracing Apps*, 27 NATURE MED. 361 (Feb. 15, 2021).

precautions around others.[50] These contact-tracing apps had the potential to fill a critical gap in the early pandemic as professional public health contact tracers—chronically underfunded—were quickly overwhelmed by the highly contagious disease.[51] However, low adoption severely limited their utility.[52] Specifically, the apps often required users to opt in (e.g., downloading or turning the feature on). Since the contract tracing apps required a critical mass of users to be effective, the opt-in default settings—compounded by trust issues in the tech companies developing the apps—were substantial barriers to the effective use of these contract tracing apps in the U.S. response to COVID-19.[53]

Public health activities can have both positive and negative effects on individual interests. For example, identifying an expectant mother to enroll in a nurse-family partnership program will provide that person with services that will directly improve their health and welfare. However, identifying an individual with a dangerous infectious disease could lead to required isolation from vulnerable individuals, interfering with the individual's liberty interests. Regardless, public health interventions should always be intended to promote community health. Consequently, even public health actions that infringe on some individual interests should confer at least some indirect personal or community benefits.

Generally, public health agencies have been slow to adopt big data approaches and tools. Limited funding and capacity, heavily siloed data sources, complex data protection laws, and a decentralized public health system are substantial barriers to U.S. public health agencies modernizing public health informatics infrastructure.[54] Consequently, public health agencies rely heavily on traditional data sources, like disease reporting, surveys, public health registries,

[50] *Contact Tracer's Interview Tool: Notifying People About an Exposure to COVID-19,* CTRS. DISEASE CONTROL (Updated Sept. 22, 2021), https://www.cdc.gov/coronavirus/2019-ncov/php/notification-of-exposure.html [https://perma.cc/52RA-H6PK] (Updated Sept. 22, 2021).

[51] Linden et al., *supra* note 24, at 790.

[52] Ahmed, *supra* note 49, at 134598; Eugene Y. Chan & Najam U. Saqib, *Privacy Concerns Can Explain Unwillingness to Download and Use Contact Tracing Apps when COVID-19 Concerns are High*, COMPUT. HUM. BEHAV. (Jan. 28, 2021).

[53] De la Garza, A., *Why Aren't COVID-19 Contact Tracing Apps Working?* TIME (Nov. 10, 2020), https://time.com/5905772/covid-19-contact-tracing-apps/ [https://perma.cc/B2TB-2KVN]; J. Rich, *How Our Outdated Privacy Laws Doomed Contact-Tracing Apps*, BROOKINGS (2021), https://www.brookings.edu/blog/techtank/2021/01/28/how-our-outdated-privacy-laws-doomed-contact-tracing-apps/ [https://perma.cc/B2TB-2KVN].

[54] *See generally* Panhuis et al., *supra* note 15; *CDC Public Health Data Failures Mean U.S. Lacks Whole Picture on COVID*, BLOOMBERG, https://www.bloomberg.com/news/articles/2021-12-21/cdc-public-health-data-failures-mean-u-s-lacks-whole-picture-on-covid [https://perma.cc/5T4H-QXYE] (last visited Apr. 11, 2022); *Crumbling Data Infrastructure Undermines Nation's Pandemic Response*, CAL. HEALTH CARE FOUND. BLOG, https://www.chcf.org/blog/crumbling-data-infrastructure-undermines-nations-pandemic-response/ [https://perma.cc/N4YP-WQJ6] (last visited Apr. 11, 2022).

and syndromic surveillance. Nevertheless, there is intense study on the potential of non-traditional data sources to promote population health.[55] These efforts include calls to promote investigation of new digital health applications—such as using data from health information technology, wearable devices, mobile applications, and other big data—to identifying challenges and opportunities to incorporate new data sources to supplement public health responses.[56] For example, Katsis et al. applied big data methods to identify the top determinants of life expectancy in San Diego, including data on the physical and built environment and consumer buying patterns, and successfully identified important factors (e.g., violent crime, parks, fast food density). However, their analysis had to contend with differentially aggregated datasets that could not be combined, in contrast to many private sector big data applications that utilize non-aggregated data that are highly linkable.[57] Widespread efforts to incorporate new data into public health applications, including occupational and environmental health, policymaking, and disaster response, are nascent and promising. However, their success will hinge on the existence of data protection laws that permit data to be used for these purposes.[58]

[55] *See generally* Yannis Katsis et al., Big Data Techniques for Public Health: A Case Study, 2017 IEEE/ACM INT'L CONF. ON CONNECTED HEALTH: APPLICATIONS, SYS. AND ENG'G TECH. (CHASE) 222, https://ieeexplore.ieee.org/document/8010636 [https://perma.cc/SV8B-2LQZ]; Sudip Bhattacharya et al., *Applications of m-Health and e-Health in Public Health Sector: The Challenges and Opportunities*, 8 INT'L J. MED. & PUB. HEALTH 56–57 (2018); Jennifer L. Chan & Hemant Purohit, *Challenges to Transforming Unconventional Social Media Data into Actionable Knowledge for Public Health Systems During Disasters*, 14 DISASTER MED. & PUB. HEALTH PREPAREDNESS 352–359 (2020), https://www.cambridge.org/core/journals/disaster-medicine-and-public-health-preparedness/article/abs/challenges-to-transforming-unconventional-social-media-data-into-actionable-knowledge-for-public-health-systems-during-disasters/8E422A5362F4D81F9C7BFE51531DEF6A [https://perma.cc/J3QH-55US]; David M. Stieb et al., *Promise and Pitfalls in the Application of Big Data to Occupational and Environmental Health*, 17 BMC PUB. HEALTH 1–4 (2017), https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-017-4286-8 [https://perma.cc/6VP7-62VP]; Michelina Mancuso et al., *Proof of Concept Paper: Non-Traditional Data Sources for Public Health Surveillance*, PROC. OF THE 6TH INT'L CONF. ON DIGIT. HEALTH CONF. 91 (2016) https://dl.acm.org/doi/10.1145/2896338.2896369 [https://perma.cc/J9AK-4Y2B]; Zachary H. Seeskin et al., *Uses of Alternative Data Sources for Public Health Statistics and Policymaking: Challenges and Opportunities*, 2018 JOINT STATISTICAL MEETINGS (2018) https://www.norc.org/PDFs/Publications/SeeskinZ_Uses%20of%20Alternative%20Data%20Sources_2018.pdf [https://perma.cc/R624-5FDL].

[56] Eric R. Buhi, *Digital Health and AJPH: The Time Has Come!*, 105 AM J. PUB. HEALTH 420 (2015). https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2015.302585 [https://perma.cc/US2E-XWLT]. *See generally* Chan & Purohit, *supra* note 55; Shawn Dolley, *Big Data's Role in Precision Public Health*, 6 FRONTIERS IN PUB. HEALTH 68 (2018).

[57] Katsis et al., *supra* note 55, at 226.

[58] *See* Panhuis et al., *supra* note 15, at 1-9 (noting the legal barriers to public health data use); Schmit et al., *supra* note 28, at 83–86.

B. American Data Protection and Privacy Law

In the United States, statutes typically govern personal data, if they do so at all, based on their substantive content. Many different federal laws do so, as do some state laws. Until 2021, only one state—California—had a comprehensive data privacy law. In that year, two more states—Virginia and Colorado—adopted statutes similar in many ways to each other and quite different from California's. Also in 2021, the Uniform Law Commissioners adopted the Uniform Personal Data Protection Act. This Section explains these developments.

1. The Current Patchwork of Law

Sectoral laws that define protected data by their substantive content are typical in the U.S. federal data protection framework. Most of them are *sui generis* approaches to specific types of information or specific regulated entities. Laws regulate health information,[59] education records,[60] substance use disorder records,[61] financial aid information,[62] financial transaction records,[63] video rental history,[64] children's internet activity,[65] government records,[66] laboratory data,[67] customer records,[68] scientific research data,[69] and social service data.[70] Many of these were enacted to address specific problems. For instance, the Genetic Information Nondiscrimination Act of 2008 (GINA)[71] was enacted to address fears that advancements in genomic science—specifically the discovery of genetic markers predictive of future health conditions—would enable discrimination by employers and insurers. Similarly, the Protection of Pupil

[59] Health Insurance Portability and Accountability Act of 1996 (HIPAA), H.R. 3103, 104th Cong. (1996); 45 C.F.R. §§ 164.102–164.534 (2021).

[60] Family Educational Rights and Privacy Act (FERPA), 20 U.S.C. § 1232g; 34 C.F.R. §§ 99.1–99.67 (2021).

[61] Confidentiality of Substance Use Disorder Patient Records, 42 C.F.R. §§ 2.1-2.67 (2021).

[62] Higher Education Act, 20 U.S.C § 1092b; 34 C.F.R. Part 5b (2021).

[63] Gramm-Leach-Bliley Act (GLBA), S. Res. 900, 106th Cong. (1999) (enacted); 16 C.F.R. § 313 (2021).

[64] Video Privacy Protection Act (VPPA), S. Res. 2361, 100th Cong. (1988) (enacted) (amended by H.R. Res. 6671, 112th Cong. (2013) (enacted)); 18 U.S.C. § 2710.

[65] Children's Online Privacy Protection Act (COPPA), S.R. 2326, 105th Cong. (2000) (enacted); 15 U.S.C. §§ 6501–6506.

[66] Privacy Act of 1974, S. Res. 3418, 93rd Cong. (1974) (enacted), 5 U.S.C. § 552.

[67] Clinical Laboratory Improvement Amendments of 1988, H.R. Res. 5471, 100th Cong. (1988) (enacted); 42 C.F.R. § 493 (2021).

[68] FTC Act, 15 U.S.C. §§ 41-58, as amended; Fair Credit Reporting Act (FCRA), H.R. Res. 15073, 91st Cong. (1970); 15 U.S.C. § 1681.

[69] Common Rule, *supra* note 9.

[70] *See* the confidentiality provisions of 7 U.S.C. Ch. 51; 7 C.F.R. § 246.26.

[71] H.R. Res. 493, 110th Cong. (2008) (enacted).

Rights Amendment (PPRA)[72] was enacted to address parents' concerns that school-based surveys would collect information from children that parents deemed inappropriate (e.g., politics, religion, sex, mental and behavioral health, income).

State data privacy laws also usually limit their scope to data records with certain kinds of information or regulated entities in certain industries.[73] And many states have long had comprehensive regulations regarding data records that *governments themselves* collect.[74] Here, too, many states have deliberated on comprehensive bills, but until California in 2018 and now Virginia and Colorado in 2021, none have been adopted.[75]

In public health, defining protected data records by the substantive content of the information makes sense where the risks of inappropriate information use or disclosure are sufficiently different than other data with different substantive content. For example, during the early years of the AIDS epidemic, there was enormous concern that AIDS and HIV records would be used to facilitate discrimination and social stigma.[76] In response, many states enacted special data laws regulating HIV data differently than other health data.[77] However, studies cast doubt on whether these additional privacy protections were efficacious for public health outcomes.[78] Nevertheless, HIV and AIDS information carry substantively different risks than other types of health information. Consequently, such sensitive information may appropriately be subjected to greater protections or restrictions than less sensitive information (e.g., phone book information).

Critically, differential data protection on data types has consequences. For example, health records can contain data that are regulated by different laws. The Health Insurance Portability and Accountability Act (HIPAA) governs health

[72] 20 U.S.C. § 1232(h).

[73] *See, e.g.,* DEL. CODE ANN. tit. 18, §§ 8601, 8602 (2021) (The Delaware Insurance Data Security Act, covering security breaches of data records with financial and health information retained by insurance licensees in the state).

[74] *See., e.g.,* MINN. STAT. § 13.02(7) (2020) (Minnesota Government Data Practices Act governing "all data collected, created, received, maintained or disseminated by any government entity").

[75] Anupam Chander, Margot E. Kaminski, & William McGeveran, *Catalyzing Privacy Law*, 105 MINN. L. REV. 1733, 1772–76 (2021); *see also* VCDPA, *supra* note 4; CPA, *supra* note 4.

[76] Matthew L. Levine, *Contact Tracing for HIV Infection: A Plea for Privacy*, 20 COLUM. HUM. RTS. L. REV. 157, 183 (1988); James M. Tesoriero et al., *The Effect of Name-Based Reporting and Partner Notification on HIV Testing in New York State*, 98 AM. J. PUB. HEALTH 728, 728 (2008).

[77] Laura Lin & Bryan A. Liang, *HIV and Health Law: Striking the Balance Between Legal Mandates and Medical Ethics*, 7 VIRTUAL MENTOR. 687, 687–89 (2005).

[78] *See* Tesoriero et al., *supra* note 76, at 732–34 (finding evidence that the benefits of name-based reporting outweigh any potential deterrent effect).

information collected or held by covered entities generally, but a health record could contain information about HIV status, which may be subject to state laws, or substance use disorder information, which is governed by the restrictive 42 CFR Part 2 regulations.[79] In 2015, researchers railed against a decision by the U.S. Centers for Medicare and Medicaid Services (CMS) to strip research datasets of all records containing substance use disorder codes to protect against Part 2 violations.[80] Researchers argued that the CMS application of Part 2 not only left researchers and public health practitioners flying blind during the opioid epidemic but also that the decision caused substantial harm by creating bias within the remaining data and specifically tainting HIV and Hepatitis C research.[81] Additionally, distinct legal protections on different data types limit opportunities to link datasets to discover important associations between various factors.[82] For instance, low education is one of the most significant causes of death in the United States, killing approximately the same number of people annually as heart attacks.[83] However, the research exception in the Family Educational Rights and Privacy Act of 1974 (FERPA) does not permit use of identifiable education records for health research, effectively hobbling data scientists' ability to understand this substantial cause of mortality.

Moreover, when datasets contain substantive information covered by different data protection laws, multiple laws might apply simultaneously. For example, up to six different data protection laws can apply to health records held by the U.S. Department of Veterans Affairs (VA).[84] Consequently, a legal analysis of a proposed VA health data use or disclosure requires an analysis of six different laws to determine which provisions of the laws are most stringent and should apply.[85] Public health data projects using data on different social,

[79] 45 C.F.R. Parts 160 and 164. (n.d.); 42 C.F.R. Part 2. (n.d.); See Pennsylvania's Confidentiality of HIV-Related Information Act, 35 Pa. Stat. Ann. § 7601, et al. (West).

[80] *See generally* Austin B. Frakt & Nicholas Bagley, *Protection or Harm? Suppressing Substance-Use Data*, 372 NEW ENG. J. MED. 1879–1881 (2015).

[81] *Id.* at 1881.

[82] Braveman & Gottlieb, *supra* note 14, at 27; SCHMIT ET AL., *supra* note 30, at 2–3.

[83] Galea et al., *supra* note 14, at 1462.

[84] The Freedom of Information Act, 5 U.S.C. § 552, implemented by 38 C.F.R. §§ 1.550–1.562; the Privacy Act, 5 U.S.C. § 552a, implemented by VA at 38 CFR 1.575-1.582; the VA Claims Confidentiality Statute; 38 U.S.C. § 5701, implemented by 38 CFR Section 1.500-1.527; Confidentiality of Drug Abuse, Alcoholism and Alcohol Abuse, Human Immunodeficiency Virus (HIV) Infection, and Sickle Cell Anemia Health Records, 38 U.S.C. § 7332, implemented by 38 CFR 1.460–1.496; HIPAA, 45 CFR Parts 160 and 164; Confidentiality of Medical Quality Assurance Review Records, 38 U.S.C. § 5705, implemented by 38 CFR 17.500–17.511.

[85] DEP'T OF VETERANS AFFS., VHA DIRECTIVE 1605.01: PRIVACY AND RELEASE OF INFORMATION 1, 3 (2016), https://www.va.gov/vhapublications/ViewPublication.asp?pub_ID=3233 [https://perma.cc/73Z2-N9ZB] (providing that "all six statutes will be applied simultaneously" and "the result will be the application of the more stringent provision for all uses or disclosures of VHA health care data").

economic, and environmental factors frequently face similar issues because these data are often covered by different laws within the U.S. patchwork.

2. Changes on the Horizon

Most of the laws we have discussed here are federal laws. There have been efforts to adopt a federal comprehensive data protection act, so far with no success. At least eleven bills that would have provided a comprehensive federal data protection regime were introduced in Congress between 2018 and 2020.[86] Hearings continue on new initiatives.[87] "The prospect for a comprehensive federal privacy law coming to the fore in 2022 is slim," however, thanks in part to it being an election year in a closely divided Congress.[88]

States are beginning to move into the gap. In 2018, California adopted the California Consumer Privacy Act (CCPA), which became operative on January 1, 2020.[89] Nevertheless, the voters considerably amended its provisions with a referendum adopted in the 2020 general election, titled the "California Privacy Rights Act of 2020," with provisions taking effect January 1, 2023.[90] While the older provisions of the CCPA remain in effect through December 31, 2022, we focus our attention in this Article on versions of the provisions that will be effective in 2023.

Other states have not remained entirely idle during this time. There were several failed attempts in various states to enact comprehensive privacy legislation,[91] but in 2021, two states succeeded where others failed: Virginia

[86] Chander et al., *supra* note 75, at 1734 n.6 (2021); *see also* Solow-Niederman, *supra* note 45, at 38–39 (noting the "116th Congress, which convened from January 2019 to January 2021 and featured a score of comprehensive (also sometimes called 'omnibus') information privacy statutes alongside a bevy of bills that emphasize a particular aspect of information privacy"); Julie E. Cohen, *How (Not) to Write a Privacy Law*, KNIGHT FIRST AMENDMENT INSTITUTE (Mar. 23, 2021), https://knightcolumbia.org/content/how-not-to-write-a-privacy-law [https://perma.cc/MPX6-22A8].

[87] *See, e.g.,* Cameron F. Kerry, *Senate Hearing Opens the Door to Individual Lawsuits in Privacy Legislation*, BROOKINGS TECHTANK (Oct. 8, 2021), https://www.brookings.edu/blog/techtank/2021/10/08/senate-hearing-opens-the-door-to-individual-lawsuits-in-privacy-legislation/ [https://perma.cc/A334-Q6EE].

[88] Jake Holland, *2022 Privacy Legislation Success Viable as Three States Lead Way*, BLOOMBERG LAW (Jan. 3, 2022, 4:00 AM), https://news.bloomberglaw.com/privacy-and-data-security/2022-privacy-legislation-success-viable-as-three-states-lead-way [https://perma.cc/TTG5-NRQ2].

[89] 2018 Cal. Legis. Serv. Ch. 55 (A.B. 375) (West). The legislature amended it two times within its first year of existence. 2018 Cal. Legis. Serv. Ch. 735 (S.B. 1121) (West); 2019 Cal. Legis. Serv. Ch. 757 (A.B. 1355) (West).

[90] 2020 Cal. Legis. Serv. Prop. 24 (West).

[91] *See* David Stauss, *Status of Proposed CCPA-Like State Privacy Legislation as of June 14, 2021*, BYTE BACK (June 13, 2021), https://www.bytebacklaw.com/2021/06/status-of-proposed-ccpa-like-state-privacy-legislation-as-of-june-14-2021/ [https://perma.cc/NGR8-Q4H8].

adopted its Consumer Data Protection Act,[92] with its terms slated to become effective January 1, 2023,[93] and Colorado followed suit when Governor Polis signed the Colorado Privacy Act,[94] with its terms taking effect July 1, 2023.[95] Most recently, Utah and Connecticut became the fourth and fifth states to enact a comprehensive privacy law, borrowing elements from the California, Virginia, and Colorado statutes.[96]

Meanwhile, the Uniform Law Commissioners had decided to consider a uniform statute, authorizing a drafting committee for the UPDPA in summer 2019 and adopting a final version of it in July 2021.[97] ULC was formed to promote consistency among state laws,[98] and its uniform statutes have often been met with great success. For example, the 2015 Revised Uniform Fiduciary Access to Digital Assets Act (RUFADAA) has been adopted in forty-five states (along with the District of Columbia and the Virgin Islands).[99] It provides means for fiduciaries like executors of estates, trustees, and attorneys-in-fact to gain access to a principal's intangible digital assets—including websites and domains

[92] 2021 Va. Legis. Serv. 1st Sp. Sess., Ch. 36 (S.B. 1392).

[93] *Id.* § 4.

[94] 2021 Colo. Legis. Serv. Ch. 21-190 (West).

[95] *Id.* § 7.

[96] Utah Consumer Privacy Act, SB 227 (2022); Connecticut Data Privacy Act, S.B. 6, (2022). Unfortunately, we were unable to incorporate these most recent developments into our analysis due to its proximity to publication.

[97] Katie Robinson, *New Drafting and Study Committees to be Appointed*, UNIF. L. COMM'N (July 24, 2019, 4:37 PM), https://www.uniformlaws.org/committees/community-home/digestviewer/viewthread?MessageKey=bc3e157b-399e-4490-9c5c-608ec5caabcc&CommunityKey=d4b8f588-4c2f-4db1-90e9-48b1184ca39a&tab=digestviewer; UPDPA, *supra* note 4 (see title page) [https://perma.cc/EQ74-4HS6].

[98] UNIF. L. COMM'N, *supra* note 3. State governments appoint ULC commissioners, all of whom are members of the bar—some practicing lawyers and some legal scholars. ULC is not the only national organization promoting uniform or model privacy legislation, though. In 2017, the National Association of Insurance Commissioners, which represents state insurance regulators, promulgated a state "Insurance Data Security Model Law." INSURANCE DATA SECURITY MODEL LAW (NAT'L ASS'N OF INS. COMM'RS (2017), https://content.naic.org/sites/default/files/inline-files/MDL-668.pdf. [https://perma.cc/5BJ5-RL8L]. As of June 2020, NAIC claimed eleven states had adopted the act. NAT'L ASS'N OF INS. COMM'RS & THE CTR. FOR INS. POL'Y & RSCH., STATE LEGISLATIVE BRIEF (2020), https://www.naic.org/documents/cmte_legislative_liaison_brief_data_security_model_law.pdf [https://perma.cc/8BHX-W9KQ]; *See, e.g.*, DEL. CODE ANN. tit. 18, § 8601 (2021).

[99] *Fiduciary Access to Digital Assets Act, Revised,* UNIF. L. COMM'N, https://www.uniformlaws.org/committees/community-home?CommunityKey=f7237fc4-74c2-4728-81c6-b39a91ecdf22 [https://perma.cc/L5RN-JSL3]; *see also Summary*, UNIF. L. COMM'N, https://www.uniformlaws.org/acts/ucc [https://perma.cc/256S-F3DV] (noting that all fifty U.S. states have adopted the Uniform Commercial Code, which governs common commercial transactions, such as sales of goods, negotiable instruments, and secured transactions); UNIF. ANATOMICAL GIFT ACT (UNIF. L. COMM'N 2007) (governing organ donations and adopted in every state—along with the District of Columbia, Puerto Rico, and the Virgin Islands—except Delaware, Florida, New York, and Pennsylvania).

and computer files in the cloud—just as such fiduciaries have been able to access tangible assets (like cars, real estate, etc.) to carry out the wishes of the principal.[100] Not all of ULC's uniform statutes have been so widely adopted.[101] Nor should the reader be misled by the "uniform" in each of these statutes' names, because each jurisdiction may adopt the act with variations.[102] As a consequence of these limitations, it's difficult to know whether, when, and how provisions of the UPDPA will become the law in states.

Nevertheless, the interest that some populous states have shown in privacy legislation and the speed with which the RUFADAA (and its revised version) have been widely adopted suggest that the UPDPA may be on many legislatures' agendas in spring 2022.[103] Indeed, within six months of ULC's adoption of the UPDPA, three jurisdictions had introduced it for deliberation.[104] In addition to the UPDPA, the California, Virginia, and Colorado laws are serving as alternative templates to states exploring comprehensive privacy legislation.[105]

The existing complexity in U.S. privacy law supports an argument for comprehensive federal data privacy legislation that would preempt state acts: Additional inconsistent privacy laws adopted state by state could further complicate efforts to monitor public health issues across jurisdictions. We do not

[100] FIDUCIARY ACCESS TO DIGITAL ASSETS ACT, REVISED (UNIF. L. COMM'N 2015).

[101] For example, only eight jurisdictions (seven states and D.C.) have adopted 2007's Limited Cooperative Association Act. *See Limited Cooperative Association Act*, UNIF. L. COMM'N, https://www.uniformlaws.org/committees/community-home?CommunityKey=22f0235d-9d23-4fe0-ba9e-10f02ae0bfd0 [https://perma.cc/SXA7-FZ5C]. And so far, only four states (as of April 28, 2022) have enacted 2019's Registration of Canadian Money Judgments Act. Two more have introduced legislation to adopt it. *Registration of Canadian Money Judgments Act*, UNIF. L. COMM'N, https://www.uniformlaws.org/committees/community-home?CommunityKey=49ecb2a9-a8b7-4041-8eba-e9d6f7293ea5 [https://perma.cc/5N42-J2BE].

[102] I. Richard Ploss, *Estate Planning for Digital Assets: Understanding the Revised Uniform Fiduciary Access to Digital Assets Act and Its Implications for Planners and Clients*, J. FIN. PLANNING Apr. 2018 (noting that "state legislatures are free to pick and choose which sections [of a uniform act] they wish to enact . . . ." so though "the RUFADAA defines a 'fiduciary' to include a court-appointed conservator, New Jersey's version of the RUFADAA specifically excludes a conservator from the definition of a fiduciary").

[103] Stauss, *supra* note 91 (summarizing 2021 legislative initiatives from June 2021 and identifying more than twenty states where bills had been introduced, of which only Virginia's and Colorado's were adopted); *see also* CS/CS/HB 969 (2021) - Consumer Data Privacy, FLA. H. REP., https://www.myfloridahouse.gov/Sections/Bills/billsdetail.aspx?BillId=72062 [https://perma.cc/ZC3A-E3XP] (showing that this Florida bill failed to be adopted).

[104] Personal Data Protection Act, UNIFORM LAW COMMISSION, https://www.uniformlaws.org/committees/community-home?CommunityKey=28443329-e343-4cbc-8c72-60b12fd18477 [https://perma.cc/H4B6-YVUQ].

[105] As of April 7, 2022, fifteen U.S. states had at least one legislative proposal introduced in both legislative houses, and Utah had adopted a statute. Taylor Kay Lively, *US State Privacy Legislation Tracker*, IAAP.COM (Apr. 7, 2022), https://iapp.org/resources/article/us-state-privacy-legislation-tracker/ [https://perma.cc/MS5Q-2RPY].

have the space here to analyze all the potential preemption issues relating to the UPDPA and the CAVACO statutes. We can note, however, as Professors Chander, Kaminski, and McGeveran have done, that the new comprehensive state laws are not likely preempted by any existing federal law under the Dormant Commerce Clause.[106] And a new comprehensive federal privacy law, when enacted, might provide only a floor that state law could build on—much as the previous sectoral federal laws have done—rather than a preemptive ceiling.[107] Public health advocates on the whole view preemption with skepticism, however, because such legislation has sometimes been proposed as a tool to suppress innovative public health measures by local governments (e.g., taxes on sugar-sweetened beverages, menu labeling).[108] Nevertheless, within public health informatics, variation in data protection laws stands as a barrier to public health practice in and of itself.[109] For similar reasons, data privacy advocates—and even some members of the ULC—suggest that a comprehensive and preempting federal privacy law is a preferred approach to the current U.S. patchwork.[110]

Legal scholars have not been silent regarding these developments, both from the perspective of privacy law and of public health. Many of their commentaries focus on normative concerns generally and particularly at the boundaries of these two disciplines.

C. Normative Concerns at the Boundaries

Professors Daniel Solove and Paul Schwartz conceive of privacy as "a constitutive element of civil society."[111] Professor Solove further identifies nearly a dozen bases upon which privacy is therefore valuable.[112] Deliberations on bills covering data protection and data privacy occur against a backdrop of legal

[106] Chander et al., *supra* note 75, at 1794–96.

[107] *Id.* at 1797–99.

[108] Policy Statement, AM. PUB. HEALTH ASS'N, *Impact of Preemptive Laws on Public Health, Policy Number: 201511* (Nov. 03 2015), https://www.apha.org/policies-and-advocacy/public-health-policy-statements/policy-database/2016/01/11/11/08/impact-of-preemptive-laws-on-public-health [https://perma.cc/59J8-65GL].

[109] Schmit et al., *supra* note 28, at 84.

[110] Joseph Duball, *Uniform Law Commission Takes Up Privacy Law Endeavor*, IAPP (Feb. 25, 2020), https://iapp.org/news/a/uniform-law-commission-takes-up-privacy-law-endeavor [https://perma.cc/CKN8-MMV3].

[111] Daniel J. Solove, *The Myth of the Privacy Paradox*, 89 GEO. WASH. L. REV. 1, 28 (2020) (quoting Paul M. Schwartz, *Privacy and Democracy in Cyberspace*, 52 VAND. L. REV. 1609, 1613 (1999)).

[112] *Id.* at 31–33 (identifying them as limiting government and company power, respecting individuals, allowing reputation management, maintenance of appropriate social boundaries, trust, "control over one's life," "freedom of thought and speech," "freedom of social and political activities," the opportunity to "change and have second chances," "protection of intimacy, body, and sexuality," and "not having to explain or justify oneself").

scholarship that theorizes the paradigm exhibited most in existing U.S. statutes as the "notice and choice" or "consumer protection" paradigm.[113] Its central tenet is that those who gather and process data should be able to use it as they please, so long as data subjects are able to decide whether to share data for primary and secondary uses after being given notice of the intended uses. Much recent scholarship has criticized this paradigm, including work that has noted weaknesses in "notice and choice" on its own terms and work that has proposed instead paradigms focused on other interests. We discuss them briefly here, identifying normative concerns, especially as they relate to public health. We will assess those concerns in relation to the UPDPA and CAVACO statutes in Part III.

1. *Is "Notice and Choice" Possible?*

Consumers' attitudes reflect a preference for limiting the collection of their personal information and a skepticism of sharing of their information with third parties.[114] Of course, consumer privacy attitudes vary considerably within populations. For example, research has measured differences in privacy concerns and behaviors between different age groups on social-network websites.[115] Additionally, consumer experience can affect privacy concerns. For example, individuals with more positive healthcare experiences were less concerned with the privacy of their health records.[116] Consumer privacy concerns are also frequently a topic in national news coverage of data breaches, or novel data uses, increasing public awareness and concerns.[117]

[113] Solow-Niederman, *supra* note 45, at 17 (asserting that the California Act "remains focused on individual rights and attempts to empower individuals by providing opportunities to opt-out of data collection"); Cohen, *supra* note 86 (arguing that almost all current congressional approaches "adopt a basic structure that is indebted to property thinking").

[114] CISCO CYBERSECURITY, CONSUMER PRIV. SERIES 3, 3–7, 11–12 (Nov. 2019) https://www.cisco.com/c/dam/global/en_uk/products/collateral/security/cybersecurity-series-2019-cps.pdf [https://perma.cc/LL99-M5S2]; H. Jeff Smith, Sandra J. Milberg & Sandra J. Burke, *Information Privacy: Measuring Individuals' Concerns about Organizational Practices*, 20 MIS Q. 167, 189, 195 (1996). *See generally* Timothy R. Graeff & Susan Harmon, *Collecting and Using Personal Data: Consumers' Awareness and Concerns*, 19 J. CONSUMER MKTG. 302 (2002); Mary J. Culnan, *"How Did They Get My Name?": An Exploratory Investigation of Consumer Attitudes Toward Secondary Information Use*, 17 MIS Q. 341, 345 (1993).

[115] *See generally* Murat Kezer, et al., *Age Differences in Privacy Attitudes, Literacy and Privacy Management on Facebook*, 10 J. PSYCH. RSCH. CYBERSPACE CYBERPSYCHOLOGY (2016).

[116] Vaishali Patel, et al., *The Role of Health Care Experience and Consumer Information Efficacy in Shaping Privacy and Security Perceptions of Medical Records: National Consumer Survey Results*, 3 JMIR MED. INFORMATICS 12–13 (2015).

[117] Rob Copeland, *Google's 'Project Nightingale' Gathers Personal Health Data on Millions of Americans*, WALL ST. J. (Nov. 11, 2019), https://www.wsj.com/articles/google-s-secret-project-nightingale-gathers-personal-health-data-on-millions-of-americans-11573496790 [https://perma.cc/VCW5-QDMK]; Kevin Granville, *Facebook and Cambridge Analytica: What*

While consumers often demand notice and choice rights, a growing body of literature suggests that the sense of control they provide may be illusory. As Alicia Solow-Niederman has noted, "individual rights to opt into or out of data collection or subsequent uses won't help if there are flaws in the individual control model to begin with."[118]

For example, there is a well-documented disconnect between consumers' stated privacy attitudes and consumers' privacy behaviors. The literature on this "privacy paradox" describes a phenomenon where individuals who express strong privacy concerns often will casually give personal information to businesses or organizations that request it, receiving in return only a *de minimis* benefit.[119] Professor Daniel Solove has proposed to dissolve the privacy paradox by noting that consumers' abstract privacy preferences and their personal practices in particular contexts are conceptually distinct.[120] In his view, it is quite consistent on the one hand for consumers to have privacy-enhancing preferences in the abstract and on the other hand, for them to fail to protect their own privacy when faced with a plethora of privacy policies and terms of use. The problem lies in the structural implementation and context where notice and choice rights are provided to consumers.

Unquestionably, the cost in time to assess each individual privacy option a consumer has, what Solve calls "privacy self-management," is great.[121] Even carefully designed interfaces intended to help consumers understand their choices better[122] are of little help if the consumer confronts hundreds of them during a

You Need to Know as Fallout Widens, N.Y. TIMES (Mar. 19, 2018), https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html [https://perma.cc/HLT5-W3KS].

[118] Solow-Niederman, *supra* note 45, at 7.

[119] Patricia A. Norberg, Daniel R. Horne & David A. Horne, *The Privacy Paradox: Personal Information Disclosure Intentions Versus Behaviors*, 41 J. CONSUMER AFFS. 100, 118 (2007); Patricia A. Norberg & Daniel R. Horne, *Privacy Attitudes and Privacy-Related Behavior*, 24 PSYCH. & MKTG. 829, 830 (2007); Susanne Barth & Menno D.T. de Jong, *The Privacy Paradox: Investigating Discrepancies Between Expressed Privacy Concerns and Actual Online Behavior*, 34 TELEMATICS & INFORMATICS 1038, 1039 (2017); Spyros Kokolakis, *Privacy Attitudes and Privacy Behaviour: A Review of Current Research on the Privacy Paradox Phenomenon*, 64 COMPUTS. & SEC. 122, 131 (2017).

[120] Solove, *supra* note 111, at 4 (stating that "behavior involves risk decisions within specific contexts," while "[a]ttitudes are more general views about value and can exist beyond specific contexts").

[121] *Id.* at 5 ("Managing one's privacy is a vast, complex, and never-ending project that does not scale; it becomes virtually impossible to do comprehensively.").

[122] *See, e.g.*, Patrick Gage Kelley et al., *Standardizing Privacy Notices: An Online Study of the Nutrition Label Approach*, CHI '10: PROC. SIGCHI CONF. ON HUMAN FACTORS IN COMPUTING SYS. 1573 (Apr. 2010) (describing development and testing of a graphical interface to facilitate consumer privacy choices).

year.[123] Other legal scholars have also questioned whether consumers have the capacity to understand the implications of their consent when increasingly sophisticated algorithms are being developed to make predictions or inferences about them or persons like them.[124] These and other concerns raise legitimate questions on whether notice and choice rights provide consumers meaningful protections.

2. *Is "Notice and Choice" Desirable?*

Many scholars have challenged the "notice and choice" paradigm on the grounds that it starts with the wrong assumptions. These include scholars who propose that there are interests at stake in data privacy and protection other than those of the data subjects and those who collect and process the data; others advocate for a model of "information fiduciaries." There exists debate, too, as to the extent that the European Union's General Data Protection Regulation (GDPR) should be a model for American regulation. Professor Julie Cohen has noted that "[c]urrent approaches to crafting privacy legislation are heavily influenced by the antiquated private law ideal of bottom-up governance via assertion of individual rights, and that approach, in turn, systematically undermines prospects for effective governance of networked processes that operate at scale."[125] The individual rights approach may fail in terms of being both over- and underprotective of individual interests.

The individual-rights paradigm is underprotective when it fails to account for the ways that data may be used about consenting and non-consenting data subjects. As Solow-Niederman has noted, "[i]t's difficult to imagine that a social media user who consented to a platform's terms of service imagined that disclosure in that context would permit . . . emergent profiling. When any bit of data might be relevant in any range of future contexts, it becomes impossible for an individual to conceptualize the risks of releasing data."[126] This is especially true when data are processed by "downstream" recipients who have no direct

[123] *See generally* Jacob Leon Kröger, Otto Hans-Martin Lutz & Stefan Ullrich, The Myth of Individual Control: Mapping the Limitations of Privacy Self-Management (July 15, 2021) (unpublished manuscript), https://dx.doi.org/10.2139/ssrn.3881776 [https://perma.cc/FS7G-9TKS]. *See also* Cohen, *supra* note 86, at 4 ("The continuing optimism about consent-based approaches to privacy governance is mystifying, because the deficiencies of such approaches are well known and relatively intractable."); *id.* at 5 ("The issues that users must navigate to understand the significance of consent are too complex and the conditions surrounding consent too easy to manipulate.").

[124] Solow-Niederman, *supra* note 45, at 24 ("Machine learning analytics make it practically impossible for an individual to determine how data might or might not be significant or sensitive in a future setting."); Cohen, *supra* note 86, at 5, n.8–9.

[125] Cohen, *supra* note 86, at 3.

[126] Solow-Niederman, *supra* note 45, at 26.

relationship with data subjects.[127] The individual-rights paradigm also fails to account for the ways that publicly available information about data subjects may be combined using complex and opaque machine learning to profile persons who have not consented to being profiled, a long-standing concern in the privacy literature.[128]

The individual-rights paradigm is overprotective when it prevents data uses that would produce significant public benefits. As Professors Jane Bambauer and Brian Ray have noted, efforts to use technology to track the spread of COVID-19 were hampered by "state and federal governments (as well as influential private firms) . . . prioritizing a fetishized notion of individual privacy over collective public health."[129] The focus on individual privacy above all else led to poor designs, destined to fail.[130] They contrasted the efforts of the South Korean government, which used "multiple independent sources of information—geolocation, credit card data, closed-circuit television, facial recognition, and old-fashioned interviews—to better trace contacts and predict the risk of transmission for each person."[131] Bambauer and others have noted that "it doesn't make sense, given the particular characteristics of [COVID-19], to treat each individual's privacy choices as a matter for individual control. As with lockdowns, the decision must be made at a collective level. A user choice conception of privacy must give way to other societal interests."[132] Likewise, Professor Alan Rozenshtein offered a full-throated defense of the principle that mandatory "digital disease surveillance" is valuable but nevertheless refused to endorse the idea, saying it is "conceivable . . . that digital disease surveillance is never the right option; even well-designed digital disease surveillance presents many dangers to privacy, liberty, and equality, and there is no guarantee that such surveillance will be well designed."[133]

Importantly, "notice and choice" is used to promote the ethical principle of "respect for persons," but it is not the only mechanism to do so. The foundational

[127] Solow-Niederman, *supra* note 45, at 47.

[128] *See* Brian N. Larson & Genelle I. Belmas, *Second Class for the Second Time: How the Commercial Speech Doctrine Stigmatizes Commercial Use of Aggregated Public Records*, 58 S.C. L. REV. 1, 23–29 (and sources cited therein).

[129] Jane Bambauer & Brian Ray, *COVID-19 Apps are Terrible—They Didn't Have to Be* 2 THE DIGITAL SOCIAL CONTRACT: A LAWFARE PAPER SERIES (NOV. 2020), https://www.lawfareblog.com/covid-19-apps-are-terrible-they-didnt-have-be [https://perma.cc/2EA4-8XDT].

[130] *Id.*

[131] *Id.* at 7.

[132] Jane Bambauer et al., *It's Time to Get Real About COVID Apps*, MEDIUM (May 14, 2020), https://medium.com/@DataVersusCovid/its-time-to-get-real-about-covid-apps-dd82e08895f2 [https://perma.cc/H9UD-Z7CP].

[133] Alan Z. Rozenshtein, *Digital Disease Surveillance*, 70 AM. UNIV. L. REV. 1511, 1517 (2021).

declarations of bioethics—including the Declaration of Helsinki and the Belmont Report[134]—established the central tenets of bioethics and placed a special importance on the principle of respect for persons. In clinical research contexts, this often required taking steps to enable the autonomy of research subjects who were seen as particularly vulnerable to abuse given the significant knowledge gaps and power dynamics between researchers and their subjects. Informed consent (i.e. "notice and choice") became the primary tool to promote autonomy and, by extension, respect for persons. In the context of established researcher-subject relationships, where a duty of care exists (i.e., nonmaleficence), "notice and choice" requirements can be powerful protections.

However, this bioethical approach to respect for persons is not well-suited for all contexts. For example, in 1991 the Council for International Organizations of Medical Sciences noted that traditional bioethical guidance did not adequately cover the special features of epidemiological research, which concerns itself with groups of people rather than individual research subjects.[135] In the context of public health surveillance, "notice and choice" protections can be problematic because nonparticipation of a relative few can bias results and impede community benefits.[136] Consequently, public health ethicists recommend different approaches to the "respect for persons" principle. Instead of relying on "notice and consent," public health ethicists recommend involving communities in the decision-making process for population-level interventions.[137] Like public health, big data applications also must reckon with the unique ethical challenges associated with population-scale activities as opposed to just the ethical

[134] WORLD MED. ASS'N, *Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects* (1964); NAT'L COMM'N FOR THE PROT. OF HUM. SUBJECTS OF BIOMED. & BEHAV. RSCH., *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research* (1979) [hereinafter BELMONT REPORT].

[135] *Preamble*, COUNCIL FOR INTERNATIONAL ORGANIZATIONS OF MEDICAL SCIENCES, INTERNATIONAL GUIDELINES FOR ETHICAL REVIEW OF EPIDEMIOLOGICAL STUDIES (1991).

[136] One can argue that a right of "consent" has a countervailing "right to be counted." For example, the residents of Love Canal, N.Y., fought for a community-wide assessment of the health effects of a nearby toxic waste dump. The empirical evidence showing a connection between the waste and the community's health empowered the community to force a governmental response. Jordan Kleiman, *Love Canal: A Brief History*, SUNY GENESEO, https://www.geneseo.edu/history/love_canal_history [https://perma.cc/LZ5M-9ZFN]. The "right to be counted" asserts that what isn't counted, doesn't count, implying that assessing public health and social problems is an essential step to correcting them. *See* Amy L. Fairchild, Ronald Bayer, & James Colgrove, *Searching Eyes : Privacy, the State, and Disease Surveillance in America*, 14 EMERGING INFECTIOUS DISEASES 1826 (2008), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2630762/ [https://perma.cc/LK6J-DLVA].

[138] WORLD HEALTH ORG., *Who Guidelines on Ethical Issues in Public Health Surveillance* (2017), http://apps.who.int/iris/bitstream/10665/255721/1/9789241512657-eng.pdf. [https://perma.cc/G7YE-H3ZF]. We choose "notice and choice" as our default term for this paradigm, but when quoting the work of others, we use "notice and consent" if they do so.

challenges typical of researcher-participant relationships.[138]

Another emerging alternative to the "notice and choice" paradigm uses the concept of "information fiduciaries." Professor Jack Balkin casts the information fiduciary model as a "movement to viewing privacy in relational terms of trust and trustworthiness."[139] For Balkin, fiduciary obligations are borne "out of social relationships, and the power and vulnerability inherent in these relationships," whether those relationships are with a doctor, lawyer, or Facebook. Balkin argues that the model is needed to respond to the vulnerability and dependence created by information capitalism.[140] Under this model, Balkin argues that digital companies that collect and use end-user data should have three duties: care, confidentiality, and loyalty. He argues that the duties of "confidentiality and care require digital companies to keep their customers' data confidential and secure" and that these must "run with the data" (imposing a duty to "vet" partners and downstream data processors).[141] For Balkin, the duty of loyalty "means that digital companies may not manipulate end users or betray their trust."[142]

Interestingly, for Balkin, the duty of loyalty and to act in the interest of the data subject extends beyond the individual to the public more broadly. He argues that "large platforms like Facebook, Google, and Amazon have so many end users that a requirement that they must act in the interests of their end users *effectively requires them to act in the interests of the public as a whole.*"[143] This last point suggests the fiduciary model—which appears consumer-focused when described as a relationship between a data subject and a data controller—could function as a public-benefit model when applied to big data across many data subjects or the whole population. From a public health perspective, a "best interests" analysis could take into account community benefits from uses for public health that result perhaps only in small marginal benefits to the individuals to whom the data refer or only indirect benefits in the form of positive externalities. Balkin's fiduciary approach could be more consistent with a bioethical (or even public ethics) approach to data protection given that fiduciary obligations implicate other ethical principles beyond "respect for persons" and because traditional "notice and consent" practices fall short of these

[139] Lisa M. Lee, *Public Health Ethics Theory: Review and Path to Convergence*, 40 J. LAW MED. & ETHICS 85–98 (2012).

[140] Jack M. Balkin, *The Fiduciary Model of Privacy*, 134 HARV. L. REV. F. 11 (2020).

[141] For example, he argues that to "live without interacting with any of these services means greatly constricting one's life and opportunities," making the explicit point that "dependencies will increase over time" and the implicit point that notice and choice models are quasi-illusory because withholding consent has adverse consequences for an individual. *Id.* at 13.

141 *Id.* at 14.

142 *Id.*

143 *Id.* at 18 (emphasis added).

considerations.[144] The information fiduciary model is subject to continued debate,[145] and we do not have the space here to explore it fully.

Finally, there is debate about whether U.S. jurisdictions should shift away from the consumer-focused data privacy model traditionally used in U.S. laws and toward a more European data protection framework. Professors Chander, Kaminski, and McGeveran argue that the traditional consumer-focused U.S. approach to data privacy relies on the tenuous ability of "notice and choice" to adequately protect consumers, assuming consumers get the benefit of their bargain with data-collecting businesses. In contrast, they argue that a data protection regime like the GDPR has protections that "follow the data" and establishes the "default in Europe . . . that personal information cannot be collected or processed unless there is a specific legal justification for doing so."[146] Professors Chander, Kaminski, and McGeveran argue that the California act "shares the presumption of most other American privacy law that personal data may be collected, used, or disclosed unless a specific legal rule forbids these activities."[147] Moreover, based on their analysis of an early draft of the UPDPA and several state and federal privacy bills, they posit the idea that California is driving comprehensive privacy regulation in American jurisdictions as opposed to Europe.[148] They conclude that California is poised to catalyze comprehensive privacy regulation in American jurisdictions.[149] We conclude below that the

[144] The Belmont Report describes the "respect for persons" as having two primary considerations: First, actions that promote an individual's autonomy (i.e., informed consent); second, protection of vulnerable persons. BELMONT REPORT, *supra* note 134. Balkin's information fiduciary model, in many respects, promotes the latter respect for persons principle in that it creates a duty to act in the best interests of data subjects who might not fully understand the risks and benefits associated with certain big data applications. *See also* Solow-Niederman, *supra* note 45.

[145] *See generally id.*; Lina M. Khan & David E. Pozen, *A Skeptical View of Information Fiduciaries*, 133 HARV. L. REV. 497 (2019); Neil M. Richards & Woodrow Hartzog, *A Duty of Loyalty for Privacy Law*, 99 WASH. U. L. REV. (forthcoming) https://doi.org/10.2139/SSRN.3642217 [https://perma.cc/74PL-QE8F].

[146] Chander et al., *supra* note 75, at 1747–48.

[147] *Id.* at 1756.

[148] *Id.* at 1771, 1772–76.

[149] *Id.* at 1771, 1772–76. We note that Chander, Kaminski, and McGeveran discussed only state legislative proposals that were not enacted and not the bills eventually enacted in Colorado and Virginia. Chander et al., *supra* note 75, at 1772-76. This is no surprise as their article came out about the time of these enactments. The timing also makes it likely that the version of the UPDPA they analyzed was a draft from summer 2020, which looked radically different than the draft eventually adopted in 2021. *Compare* Collection and Use of Personally Identifiable Data Act [draft for discussion only] (Apr. 24, 2020), https://www.uniformlaws.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=f897ee80-6e47-13cd-1370-2f8c395bdde6&forceDialog=0 [https://perma.cc/65L4-R22J], *with* UPDPA, *supra* note 4. One report of an empirical study of privacy policies since the GDPR and CCPA sought evidence of the effect of these statutes on companies behavior. Jens Frankenreiter, *The Missing 'California Effect' in Data Privacy Law*, 39 YALE J. REGUL., manuscript at 8-9

UPDPA and CAVACO statutes chart a different route.[150]

Though, as we explain in the next subsection, we adopt the "notice and choice" framework as our own normative paradigm, we do so with some modifications reflecting this literature, and we will attempt to touch in Part III on points where these other frameworks may be valuable.

3. Defaults Should Play an Important Role

Given that "notice and choice" is the predominant paradigm in existing law in the United States, both at federal and (as we shall see) state levels, the normative framework we adopt here is grounded in that paradigm. A common theme of many justifications for privacy is autonomy or agency of citizens; in this case, data subjects.[151] This aligns well with a foundational document on research ethics well known among public health researchers and practitioners, the Belmont Report.[152] The Belmont Report values "respect for persons," the principle "that individuals should be treated as autonomous agents."[153] "An autonomous person is an individual capable of deliberation about personal goals and of acting under the direction of such deliberation."[154] Thus, "[t]o show lack of respect for an autonomous agent is to repudiate that person's considered judgments, to deny an individual the freedom to act on those considered judgments, or to withhold information necessary to make a considered judgment . . . ."[155]

Our view is that for this autonomy to be possible, the data subject must know how a controller will use their personal data—what we will call transparency—and have a meaningful opportunity to deliberate on whether to enter the relationship that involves the controller's data practices. As we noted in the previous subsection, such deliberation may be impossible or unlikely, and in that event, regulators should set "defaults" in line with collective expectations about data privacy. Within our framework, this means that most public health data uses, whether primary or secondary, should be disclosed to data subjects but either not subject to their consent or subject only to an opt-out, what we call

(forthcoming 2022) https://dx.doi.org/10.2139/ssrn.3883728 [https://perma.cc/M3ZU-6DK4] (finding "the impact of EU data privacy law on the relationship between U.S. businesses and their U.S. customers might be more limited than is commonly assumed"); *id.* at 9–10 ("cast[ing] doubt" on the "expectation that the [sic] California's new data privacy law (the CCPA) will have nationwide effects").

[150] *See infra* Part III(A).

[151] *See, e.g.,* Solove, *supra* note 111, at 39–41.

[152] BELMONT REPORT., *supra* note 134, pt. B(1).

[153] *Id.*

[154] *Id.*

[155] *Id.*

"passive consent."[156]

Normatively, regulators should prefer that data practices that do not require the consent of data subjects be disclosed wherever possible, even if they involve data practices in which a data custodian or "controller"[157] would be forced to engage. For example, a privacy policy should inform data subjects that the controller may disclose their personal data in response to a court order. Even if all controllers acknowledge this data practice, leaving consumers with little choice among them, it permits the (admittedly rare) consumer who is a privacy hawk to choose to withhold their personal data from all such controllers.

Defaults play a different role, because they have an outsized impact on what consumers will select.[158] Requiring only passive consent (allowing for an opt-out)[159] may be appropriate for data practices that data subjects would accept in principle or that serve public policy goals; by default, the data subject consents to them. Active consent (requiring an opt-in)[160] may be appropriate for those practices that data subjects typically reject or doubt in principle or that undermine public policy goals; by default, the data subject does not consent. This does not address all the concerns, as controllers may use a variety of other techniques to pressure data subjects into actively consenting.[161] Nevertheless, as we see below, such a default approach has a critical role to play for public health matters.[162] Absent regulatory defaults, data controllers will likely adopt the most self-serving approach, often at the expense of or risk to data subjects.

Of course, accepting that defaults are a good idea and knowing what they should be are two very different things. Despite some notable differences in privacy attitudes within the broader population, there is a growing body of literature showing broad support for the use of data for research purposes.[163] The public is generally comfortable sharing their personal information if they believe

[156] *See infra* Part II(E)(2).

[157] *See infra* Part II (defining terms).

[158] *See infra* Part II(E)(2).

[159] *Infra* Part II(E)(2).

[160] *Id.*

[161] *Id.*

[162] *See infra* Part I(C)(3).

[163] *See, e.g.*, Mhairi Aitken et al., *Public Responses to the Sharing and Linkage of Health Data for Research Purposes: A Systematic Review and Thematic Synthesis of Qualitative Studies*, 17 BMC MED. ETHICS 1 2, 4–5 (Nov. 10, 2016); Laura J. Damschroder, et al., *Patients, Privacy and Trust: Patients' Willingness to Allow Researchers to Access Their Medical Records,* 64 SOC. SCI. & MED. 223, 224 (2007); S.B. Haga & J. O'Daniel, *Public Perspectives Regarding Data-Sharing Practices in Genomics Research*, PUB. HEALTH GENOMICS 319, 321–22 (Apr. 27, 2010); Emily C. O'Brien, et al., *Patient Perspectives on the Linkage of Health Data for Research: Insights from an Online Patient Community Questionnaire*, 136 INT'L J. MED. INFORMATICS 9, 12–15 (2019); Donald J. Willison, et al., *Patients' Consent Preferences for Research Uses of Information in Electronic Medical Records: Interview and Survey Data*, 326 B. MED. J. 1, 3 (Feb. 15 2003).

that their information will contribute to the furtherance of scientific knowledge. This is particularly true for health research where participants may believe that sharing their personal health information may confer some indirect benefit in the form of new discoveries or improved treatments for their health conditions.[164]

Further evidence of the public's attitudes is provided by a series of studies that two of the authors (Schmit and Kum) have been performing with others.[165] In February 2020, they conducted a survey of 504 adults in the United States who were fluent in English and recruited by a consumer research company hired to identify a representative national sample.[166] The respondents were balanced for gender, race/ethnicity, age, education, income, and census region. Their health insurance coverage was also similar to the national distribution in data published by the U.S. Census Bureau. Researchers sought consumers' relative preferences among scenarios that varied based on the source of identifiable data, who would be using it, and the proposed data use (taking into account both legal restrictions and exceptions for data use or disclosure). The fractional factorial design the researchers used in the study allowed them to test seventy-two different data-use scenarios to determine consumers' relative preferences among them and to assess the weight that each variable had in the consumers' decisions. Through this design, the researchers were able to test whether consumer preferences aligned with the patchwork approach to U.S. privacy laws by using scenarios that varied according to the purpose for which their data would be used, the persons or entities using the data, and the type of data used. Use of these methods by the researchers allowed them to assess comparative weighting for various features in a manner not typically pursued in the research literature.

For these consumers' preferences, information about the purpose for which the data would be used was the highest priority, the identity of the user of second-greatest importance, and the nature of the data used of least importance. First, consumers supported uses for promoting population health and for research leading to scientific knowledge; they disfavored uses for identifying criminal activity, marketing and recruitment, and, most significantly, undifferentiated profit-driven activities. Second, consumers preferred data uses by university researchers, followed by non-profit organizations; they disfavored government and business users. Finally, consumers were most tolerant of uses of educational and health records and less tolerant of data from government sources and data relating to consumers' economic activity or customer behavior. The four sources

[164] Aitken et al., *supra* note 163, at 12.

[165] *See generally* Cason D. Schmit et al., *US Privacy Laws Go Against Public Preferences: Impeding Public Health and Research: Survey Study*, 23 J. MED. INTERNET RES. 1 (July 5, 2021). Another study, looking at changes to responses nine months into the COVID-19 pandemic, is in preparation.

[166] *Id*.

of data, however, were fairly close to being neutral in consumers' assessments.

When Schmit, Kum, and their colleagues combined the factors in the scenarios, they found that the top ten most acceptable scenarios all involved use by a university researcher or non-profit for scientific research or public health. Represented among the top ten were all four data sources: education, health, government-program related, and economic or customer activity. The five most *disfavored* scenarios involved for-profit businesses using data for profit-driven or marketing activities—regardless of the nature of the consumer data used. Rounding out the bottom ten least-favored uses were those involving for-profit or government uses to market programs or products and to identify criminal activity.

The researchers noted the inconsistency between consumer preferences and existing privacy laws: "Ironically, our data indicate that the U.S. public's most preferred data re-use scenario is currently prohibited under FERPA while the U.S. public's least preferred data re-use is completely legal and ubiquitous under the permissive FTC Act."[167]

The true picture of the public's preferences is of course far more complex. Public support for some data uses and for privacy frequently does not square with the fact that data privacy and data utility are competing interests. Data controllers can substantially increase data privacy, but these efforts will often make the data more difficult (or impossible) to use for certain purposes. Alternatively, fewer privacy restrictions make data more useful, but they increase the privacy risks for data subjects. For example, data can be deidentified to protect the identity of data subjects, but without identifiers, these data can no longer be linked to other databases to answer otherwise unsolvable problems. Similarly, individual privacy preferences can be incongruent. For example, some patients want their information used for research to be deidentified, and they also want to be asked before their information is reused for new research projects.[168] These wishes are incompatible: Researchers have no way to notify a deidentified data subject, much less ask for their consent to subsequent data uses. Consequently, policy and good data governance practices, grounded in data subjects' preferences and interests, are critical tools to balance the competing interests of privacy and data utility.

Trust, transparency, and individual control are critical factors for sharing data for research purposes.[169] The absence of any one of these can swiftly undermine public support in research data uses. For example, Google and the Ascension health system partnered to develop and test new big-data tools for

[167] *Id.*

[168] Aitken et al., *supra* note 163, at 12.

[169] *Id.* at 12–14.

healthcare applications.[170] This partnership was not publicly transparent, and patients were not notified or asked to opt in to the research partnership.[171] The absence of a consent undermined Ascension's patients' sense of control. The lack of transparency of the partnership with the commercial entity Google raised suspicions and undermined trust in the endeavor. As a result, the partnership faced substantial backlash.

In summary, privacy is popular with consumers in principle, but their conduct seems often to run counter to their expressed preferences. A resolution of this *privacy paradox* requires transparency from controllers and action from regulators to set the *defaults* of consumer consent, defaults that reduce social harms and promote social benefits. Informing those defaults should be our developing knowledge of consumers' preferences and an awareness of the tension between data privacy and data utility, recognizing that public health practices receive considerably more support from consumers than profit-driven activities.

Effective public health responses sometimes require balancing the rights of individuals and their autonomy with the needs of the community. It may be necessary for the community's well-being to use personal data without data subjects' opportunity to deliberate and to choose to participate.[172] Decisions to do so should not be taken lightly, however.[173]

In Part II, we will examine the three state comprehensive statutes adopted so far and the new uniform data privacy act to assess their substantive provisions, particularly those related to public health. In Part III, we will assess them against these normative frameworks and propose next steps for public health researchers and professionals.

II. ANALYSIS OF THE UPDPA AND CAVACO STATUTES

The descriptive task of this Article is somewhat daunting, and it may seem that we are getting quite far down "into the weeds," but for the reader interested in making a comparative assessment of the UPDPA and California, Virginia, and Colorado acts—what we have called the "CAVACO statutes"—a thorough doctrinal description is necessary before a normative evaluation. Those readers who are legislators or planning to take part in legislative deliberation, lobbying, etc., over similar acts will likely benefit from the detailed analysis in this Part.

[170] Copeland, *supra* note 117.

[171] Nevertheless, this project was likely compliant with HIPAA's requirements. The Google and Ascension had a signed business associate agreement, and the development of software tools likely falls within the HIPAA allowance for use and disclosure for healthcare operations or under HIPAA's generous research exception. 45 CFR § 164.501, 502. 512(i); Copeland, *supra* note 117.

[172] Bambauer et al., *supra* note 132; Rozenshtein, *supra* note 133, at 1517.

[173] Rozenshtein, *supra* note 133, at 1517.

Other readers may prefer to skip to Part III, our normative assessment of these statutes, referring back to this Part only for details of interest.

Here, we lay out a conceptual framework, which allows us to define terms to use as representational devices in a discussion of the subject matter. We intend it as a vocabulary where the definitions are stipulated but expected to be consistent with a layperson's intuitions about what they mean and how they are used. This framework could prove useful for other efforts to compare privacy paradigms and statutes.[174] The CAVACO statutes and the UPDPA have some common requirements and some that differ. This Part examines the UPDPA in more detail, setting out its basic requirements; scope; favored, restricted, and prohibited data practices; and enforcement and penalties, noting its differences from the CAVACO statutes and their differences from each other. Along the way, we will point out interesting features and address terms that will be of interest to public health professionals and researchers.

For our conceptual framework, we have drawn from the European Union's GDPR,[175] the American Law Institute's 2019 statement of the principles of data privacy law,[176] and the legislative enactments we analyze below when we have found them conceptually sound.

As a preliminary matter, a distinction between "information" and "data" is tenable on grounds that the data that are recorded may or may not accurately represent the information about the individual or the world. We can think of "information" as the truth about the world and "data" as what's collected.[177] We'll refer to a "data record" to denote data that are stored in some readable form.[178] "Personal data" is any data "relating to an identified or identifiable natural person . . . ."[179] "[A]n identifiable natural person is one who can be identified, directly or indirectly."[180] A "personal data record" is thus a data record containing personal data. The individual about whom a personal data record purports to record information is a "data subject."[181] We will refer to a "data-record practice," or just "data practice" for short, as "collection, recording,

[174] *See, e.g.,* Chander et al., *supra* note 75, at 1749–62 (comparing CCPA, GDPR, and proposed state legislation).

[175] *See generally* 2016 O.J. (L 119).

[176] *See generally* PRINCIPLES OF DATA PRIVACY, *supra* note 1.

[177] *Cf.* PRINCIPLES OF DATA PRIVACY, *supra* note 1, § 2(a) (2019). Of course, one can make a statement about data, i.e., offer information about data. But here we are generally concerned with information about and data relating to human beings.

[178] *See, e.g.*, Privacy Act of 1974, 5 U.S.C. § 552a(a)(4) ("[T]he term 'record' means any item, collection, or grouping of information about an individual that is maintained by an agency.").

[179] 2016 O.J. (L 119), art. 4(1). *See also* PRINCIPLES OF DATA PRIVACY, *supra* note 1, § 2(b) (2019).

[180] 2016 O.J. (L 119), art. 4(1).

[181] 2016 O.J. (L 119), art. 4(1). This is also the language the UPDPA uses. UPDPA, *supra* note 4, § 4. *See also* PRINCIPLES OF DATA PRIVACY, *supra* note 1, § 2(C) (2019).

organi[z]ation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction" of personal data records.[182]

Some individual or entity must engage in a data practice for there to be a legal question. We define a "data controller" as a natural person or entity that "determines the purposes and means" of a data practice,[183] and a "processor" as the natural person or entity that actually performs a data practice.[184] If the same entity both decides what data practices to undertake and also performs them, it is both a controller and a processor regarding that data practice.[185] Because of their power to decide, "data controllers have greater responsibilities than data processors."[186] Not all controllers are created equal, however. Acting together or with others, one controller "collects personal data directly from a data subject"[187]—it is the "collecting controller." As a controller, the collecting controller "determines the purpose and means of processing" of the data records,[188] but it may also make the data records available to another controller, a "third-party controller."[189]

Many uses of personal data are "secondary uses" or "secondary data practices," where data collected for one purpose is re-used for a different purpose. These secondary uses often require dissemination by the collecting controller to some other controller. For example, consumers might consent to having their local dry cleaner share records about their dry-cleaning purchases with a university researcher, who might then process the records for purposes of

[182] 2016 O.J. (L 119), Art. 4(2). This is the definition that the GDPR provides for "processing," and is quite similar to the activities that the Privacy Act of 1974 defines as "maintaining" a record. 5 U.S.C. § 552a(a)(3) ("[T]he term 'maintain' includes maintain, collect, use, or disseminate."). The UPDPA defines "maintain" more narrowly. UPDPA, *supra* note 4, § 2(8) ("'Maintains,' with respect to personal data, means to retain, hold, store, or preserve personal data as a system of records used to retrieve records about individual data subjects for the purpose of individualized communication or decisional treatment."). *See also* PRINCIPLES OF DATA PRIVACY, *supra* note 1, § 2(d) (listing "collection," "access," "retention," "use," "sharing," and "destruction" as "personal data activities").

[183] *Compare* 2016 O.J. (L 119), Art. 4(7) *with* PRINCIPLES OF DATA PRIVACY, *supra* note 1, § 2(e).

[184] *Compare* 2016 O.J. (L 119), Art. 4(8) *with* PRINCIPLES OF DATA PRIVACY, *supra* note 1, § 2(f). *But see* Solow-Niederman, *supra* note 45, at 48 (taking "controller" to mean collecting controller and "processor" to include third-party controllers).

[185] The UPDPA takes a different tack, seeming to make "controller" and "processor" mutually exclusive. UPDPA, *supra* note 4, § 2(12) (defining "processor" as one "that processes personal data *on behalf of a controller*" (emphasis added)).

[186] PRINCIPLES OF DATA PRIVACY, *supra* note 1, § 2 cmt. g.

[187] UPDPA, *supra* note 4, § 2(1).

[188] *Id.* § 2(3).

[189] *Id.* § 2(21).

research. In this example, the dry cleaner is a collecting controller, the university researcher is a third-party controller, and their research practices are secondary data practices.

Along this pipeline, any controller may use one or more processors. Controllers need not use external processors, in which case they would engage in the processing in-house. Thus, a collecting controller may be the only stop in a pipeline that it builds and maintains. The dry cleaner in the example above, for example, might use its own customer data records to market related services to its customers. It is then the sole collecting controller of the data records, and there are no other processors. Much more elaborate pipelines are, however, possible.

Given this basic vocabulary, we can consider several components that a conceptual framework for data protection must have. A critical one—and thus the first we address—is the definition of which data records are subject to the regulation. Second, we take up some considerations relating to controllers and processors. Third, we discuss common data practices that are subject to regulation. Fourth, we consider matters of the scope and jurisdiction of data privacy law. Finally, we will briefly mention enforcement mechanisms and penalties for violating the data privacy laws.[190]

A. Substantive Information Content

The UPDPA and CAVACO statutes are comprehensive personal data protection laws. Like the European Union's General Data Protection Regulation, the CAVACO statutes and the UPDPA include within their scope all personal data; importantly, though, they carve out a variety of exceptions and exemptions. Other U.S. federal and state data protection laws define protected data records using some form of description of the substantive content of the information they purport to represent or the nature of the controllers or processors.[191] We discuss the normative consequences of those choices in Part III.[192]

Subject to the UPDPA are "personal data" that relate to a "data subject" that a "collecting controller" collects and of which the controller maintains a "record."[193] Personal data under the UPDPA is "a record that identifies or describes a data subject by a direct identifier or is pseudonymized data," tracking the CAVACO statutes fairly closely.[194] UPDPA and the CAVACO statutes

[190] Because our principal focus is on public health activities, we assume that the actors involved will avoid violating the laws' requirements and may therefore be less concerned about enforcement. Readers attempting to assess risks for private actors under UPDPA and the CAVACO statutes should review those provisions of the acts and advise clients accordingly.

[191] *Supra* Section I(B).

[192] *Infra* Section III(B).

[193] UPDPA, *supra* note 4, §§ 2(1), 2(4), 2(10).

[194] *Id.* § 2(10); CCPA, *supra* note 4, § 140(o)(1) ("'Personal information' means information

exclude some data from "personal data" based on their identifiability or sensitivity, discussed further below. There are also some substantive categories of data excluded: For example, these acts do not cover personal data "processed or maintained in the course of a data subject's employment or application for employment."[195]

The UPDPA and CAVACO statutes take slightly different approaches to an exemption for personal data "processed or disclosed as required or permitted by a warrant, subpoena, or court order or rule, or otherwise as specifically required by law."[196] The UPDPA exempts these practices from its own application, but we argue it would protect data subjects better if it covered these data while permitting their disclosure only to the extent required by law, categorizing such disclosures as favored or "compatible" data practices, leaving them subject to the act.[197] The CAVACO statutes take the latter approach, not exempting these types of data from coverage but expressly not limiting a controller or processor's ability to respond to the situations described in this paragraph.[198]

that identifies, relates to, describes, is reasonably capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household."); VCDPA, *supra* note 4, § 59.1-575 ("any information that is linked or reasonably linkable to an identified or identifiable natural person"); CPA, *supra* note 4, § 6-1-1303(17)(a) (identical to VCDPA).

[195] UPDPA, *supra* note 4, § 3(c)(5). Though the official comment does not explain this exclusion, it would be reasonable to conclude that it has been excepted here because of the significantly different nature of the employment relationship and because state laws presently offer varied protections for data relating to employees. *See also* CCPA, *supra* note 4, § 145(m)(1) (excluding a variety of employment-related activities); VCDPA, *supra* note 4, § 59.1-575 (excluding from definition of "consumer," VCDPA's counterpart to data subject, "a natural person acting in a commercial or employment context"); *id* § 59.1-575(c)(14) (excluding employment-related data from application under the act); CPA, *supra* note 4, § 6-1-1304(k) (excluding "data maintained for employment records purposes"). Such a limitation in UPDPA is not without its likely critics. Elizabeth A. Brown, *The Fitbit Fault Line: Two Proposals to Protect Health and Fitness Data at Work*, 16 YALE J. HEALTH POL. L. & ETHICS 1, 14 (2016) (detailing employer uses of surveillance data); *id.* at 24 (asserting that HIPAA does not cover them); *id.* at 46–47 (proposing that HIPAA's definition of covered entities include employers, fitness-app developers, and wearable-device manufacturers).

[196] UPDPA, *supra* note 4, § 3(c)(3). This is peculiar, and possibly a drafting error, in part because personal data relating to a data subject, even sensitive data, would be taken out of protection of UPDPA in the event the controller or processor had to disclose it in litigation with a third party. Thanks to this exemption, it appears the third party would be under no restriction where further processing and disclosure of the data are involved. The controller or processor might reasonably seek a protective order when disclosing the data. Perhaps the act should require this.

[197] In fact, UPDPA elsewhere implies that type of disclosure is a compatible data practice. *See* UPDPA, *supra* note 4, § 7(b)(2), (7), (9) (defining compatible data practices to include processing "reasonably necessary to comply with a legal obligation or regulatory oversight of the controller," processing in a manner that "is reasonably necessary to prevent, detect, investigate, report on, prosecute, or remediate an actual or potential" crime, and processing that "is reasonably necessary to comply with or defend a legal claim").

[198] CCPA, *supra* note 4, § 145(a); VCDPA, *supra* note 4, § 59.1-582(A); CPA, *supra* note 4, § 6-1-1304(3)(a).

A second exemption from the UPDPA of interest here relates to research: the UPDPA does not apply to personal data "processed or maintained *solely* as part of human-subjects research conducted in compliance with legal requirements for the protection of human subjects."[199] This appears broadly to support the use of personal data for research purposes subject to the Common Rule and potentially other regimes for research ethics. Personal data collected, analyzed, and used in accord with such a research protocol would thus entirely escape the application of the UPDPA. The "solely" in the UPDPA is important, however. Data "processing" under the UPDPA includes collecting data.[200] This exemption, applying only to personal data *collected solely* for research, probably does not cover disclosures by controllers and processors to secondary data researchers. For example, if Amazon were to provide personal data about its customers' transactions (identifying customers) to a researcher solely so that the researcher could do IRB-approved research, this does not appear to be processing "*solely* as part of human-subject research" because the data was initially collected for a non-research purpose (i.e., commercial transaction). This data would be useful to public health researchers because consumer behavior data can be used to infer and predict health status. Similarly, these data would enable researchers to determine whether there is a connection between using certain products and certain health outcomes.

Getting such data from companies like Amazon is a boon for researchers because it removes the cost of recruiting survey participants from the public and provides a complete picture of the population (at least of Amazon users). But the researchers do their processing, limited by the IRB protocol, solely as part of human-subjects research, while Amazon, the collecting controller of the personal data, collects and processes the data for other reasons. As the UPDPA covers these data, researchers would instead have to determine whether the data practice is permitted under it.[201]

Slightly less strict is the Virginia Act, which broadly exempts data records in research conducted according to applicable ethical standards.[202] But it goes further and exempts information used "only for public health activities and purposes as authorized by HIPAA,"[203] which includes disclosures to a "public health authority that is authorized by law to collect or receive such information for the purpose of preventing or controlling disease, injury, or disability, including, but not limited to, the reporting of disease, injury, vital events such as birth or death, and the conduct of public health surveillance, public health

[199] UPDPA, *supra* note 4, § 3(c)(2) (emphasis added).

[200] *Id.* § 2(11).

[201] *See infra* Part II(E).

[202] VCDPA, *supra* note 4, § 59.1-576(C)(4).

[203] *Id.* § 59.1-576(C)(9).

investigations, and public health interventions."[204] This exemption, however, affects disclosures only by "covered entities," which are a "health plan," "health care clearinghouse," or "health care provider who transmits any health information in electronic form in connection with a transaction covered by" the act.[205] And the "only" in the operative Virginia provision again prevents the secondary uses contemplated in the Amazon example.

More relaxed still are the California and Colorado Acts. The California statute starts with a somewhat similar approach to the UPDPA, exempting from its application personal data that are either (a) deidentified as provided in the Code of Federal Regulations and "derived from patient information that was originally collected, created, transmitted, or maintained by an entity regulated by [HIPAA], the Confidentiality Of Medical Information Act, or . . . the Common Rule;"[206] or (b) "collected, used, or disclosed in research, as defined in [45 C.F.R. § 164.501] . . . and that is conducted in accordance . . . the Common Rule" or similar regulations."[207] But the California statute exempts use *and disclosure* in research. Colorado's statute also exempts data records *collected* in IRB-approved research, but like California's, it goes further in exempting "personal data *used or shared* in research.[208] Either statute would allow our hypothetical researcher to get access to the hypothetical Amazon data discussed in the previous paragraph, arguing it is not covered by the applicable statute.

B. Data Identifiability

U.S. data protection laws predominantly protect only identified or identifiable data records.[209] Consequently, how identifiability is defined in a law is essential to determine whether the law protects a data record. Such definitions often include one or more of three factors: The presence of direct identifiers, the presence of indirect identifiers, and the likelihood of identification through inference. In some cases, identifiability definitions are difficult to apply, so some laws include legal standards for taking identified data and rendering it pseudonymous or deidentified by law. A law may then provide different levels of protection for these levels of identifiability, or it may exclude one or more of

[204] 45 C.F.R. § 164.512.

[205] 45 C.F.R. § 160.103.

[206] CCPA, *supra* note 4, § 146(a)(4)(A).

[207] *Id.* § 146(a)(5).

[208] CPA, *supra* note 4, § 6-1-1304(2)(d) (emphasis added).

[209] There are some notable exceptions of laws that protect information based on its content. For example, trade secret laws protect information that can be identifiable (e.g., customer lists) or non-identifiable (e.g., marketing strategies). *See* TEX. CIV. PRAC. & REM. CODE ANN. § 134A.002(6) (West, 2021). Similarly, the Freedom of Information Act excludes certain sensitive government records from its disclosure requirements. 5 U.S.C. § 552.

them from its application. This Section describes these different degrees of identifiability—direct identifiers, indirect identifiers, and inferences—and explains deidentification and pseudonymized data.

Direct identifiers are data that can in theory be used by themselves to identify a specific individual. Common examples of direct identifiers include names, social security numbers, home addresses, email addresses, and phone numbers. Most direct identifiers are insufficient by themselves, however, to identify a specific individual with certainty. For example, the name "John Smith" is common and does not differentiate one John Smith from another, and even social security numbers are not always unique to an individual.[210] Still, these data can practically identify many individuals. Consequently, direct identifiers are often a core part of legal definitions of identifiability.[211]

Indirect identifiers can identify an individual, but only in combination with other data. For example, a million or more Americans may share a birthday, excluding the year—an indirect identifier—so date of birth cannot, by itself, identify an individual. However, knowing the date of birth of John Smith might enable someone to distinguish one "John Smith" from another. Similarly, postal (ZIP) codes, race, and gender information are indirect identifiers that, together with other data, can help identify a data subject.[212]

Laws that define identifiable personal data as including indirect identifiers can impede socially beneficial secondary data practices. For example, health, economic, and social outcomes can vary considerably depending on an individual's race or where they live, and data about them are often essential to research on public health. If a data processor strips data of all indirect identifiers to free it from a law's restrictions, the secondary use of the data records for research can be severely limited.

Some laws define identifiability by the possibility that an individual might determine the identity of a particular data subject by inference rather than by the presence of specific direct or indirect identifiers, for example, where "there is a reasonable basis to believe the information can be used to identify the individual,"[213] or where there is information "alone or in combination" that "would allow a reasonable person in the . . . community, who does not have

[210] Frank Hayes, *Not So Unique*, COMPUTERWORLD (Aug. 6, 2007, 12:00 AM), https://www.computerworld.com/article/2552992/not-so-unique.html [https://perma.cc/2T6S-26CC].

[211] GDPR, for example, gives the following examples of direct identifiers: "a name, an identification number, location data, [or] an online identifier . . . of [a] natural person." 2016 O.J. (L 119), art. 4(1).

[212] GDPR gives the following examples of indirect identifiers: "one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person." 2016 O.J. (L 119), art. 4(1).

[213] 45 C.F.R. § 160.103.

personal knowledge of the relevant circumstances, to identify the [data subject] with reasonable certainty."[214] All of these approaches to defining legally identifiable data ask data processors to consider the possibility that someone else could identify a data subject of a data record.[215]

Therein lies a critical problem: when data pertain to individual data subjects, often it is mathematically possible to identify at least some data subjects within a dataset.[216] Quantitatively minded data processors are of course keenly aware that without substantial redaction or data manipulation, there will always be a lingering possibility that a data subject may be reidentified if a disclosed dataset is combined with external information.[217] Consequently, absent clear safe-harbor provisions, laws that define identifiability using the possibility, foreseeability, or reasonable belief that a data subject may be reidentified using inference will always create uncertainties due to persistent possibilities of reidentification.

Perhaps because of ambiguities in legal definitions of identifiability, some laws include standards for deidentifying data. Deidentified data are data once protected by a data protection law that have been modified or redacted in such a way that they have much-diminished or even no protection under the law. Deidentification standards are particularly important for laws with broad or ambiguous definitions for identifiable data because persistent uncertainties about a law's applicability may prevent a data processor from disclosing data for socially desirable purposes. For example, HIPAA defines protected data as that which "identifies an individual" or where there is a reasonable belief that it can identify an individual. Absent a specific deidentification standard, it is difficult to know what data elements need to be redacted or modified so the data no longer meets this definition. Fortunately, HIPAA regulations contain standards that permit data processors to render data legally deidentified.[218]

Some data protection laws define a middle ground between identifiable data and deidentified data. Data in this middle ground are sometimes called

[214] 34 C.F.R. § 99.3 (2021).

[215] Contrast the Common Rule, which draws the boundary here: "identity of the subject is or may readily be ascertained by the investigator." 45 C.F.R. § 46.102. This is narrower and more easily determined than the other tests. *See also* PRINCIPLES OF DATA PRIVACY, *supra* note 1, § 2(b) (including in definitions whether "there is a moderate probability" or "low probability" that data "could be linked to a specific natural person").

[216] Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, UCLA L. REV. 1701, 1713 (2010). *But see*, Victor Janmey & Peter L. Elkin, *Re-Identification Risk in HIPAA De-Identified Datasets: The MVA Attack*, AMIA ANN. SYMP. PROC. 1329, 1329 (2018); Kathleen Benitez & Bradley Malin, *Evaluating Re-identification Risks with Respect to the HIPAA Privacy Rule*, 17 J. AM. MED. INFORMATICS ASS'N 169, 169 (2010).

[217] Hye-Chung Kum et al., *Social Genome: Putting Big Data to Work for Population Informatics*, 47 COMPUT. 56, 61–63 (2014); Benitez & Malin, *supra* note 216; *see also* Ohm, *supra* note 216.

[218] 45 C.F.R. § 164.514(b).

"pseudonymized," "coded," or "limited" data. We will use the first of these terms. For example, GDPR defines pseudonymous data as personal data that "can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organi[z]ational measures to ensure that the personal data are not attributed to an identified or identifiable natural person."[219] Other laws define it as data that is partially deidentified (or less identifiable) but does not have a key or code that connects a pseudonym with data subject identifiers.[220]

Pseudonymized data is an important category for public health research and public health population interventions. Often, research proceeds with protocols in place to replace direct identifiers in the data, such as participants' names and email addresses, with codes that allow data about a single participant to be examined in the aggregate without identifying the participant. Often researchers will keep a "key" that would allow reidentification.

By incorporating reduced restrictions for less identifiable data, laws implicitly recognize the tradeoff between privacy and data utility. Provisions that give additional flexibility for less-identifiable data enable greater data use than would typically be permitted under an all-or-nothing approach where data are either identifiable and fully protected or not identifiable and not protected. Data in these categories often receive a lower level of protection under the data protection laws. Laws that have special provisions for pseudonymized data often require some information redaction or modification (usually the removal of enumerated direct or indirect identifiers), but not so much as to render the data fully deidentified. For example, HIPAA allows for the disclosure of limited datasets. In contrast to fully deidentified datasets, limited datasets can include much more geographic information, including city, county, and ZIP code. These data permit analyses that would not be possible under fully deidentified data; however, limited datasets are often still viewed as "identifiable" data and HIPAA rules still apply.[221] Similarly, the Common Rule permits an exemption from some requirements where researchers record otherwise identifiable data in such a manner that data subjects cannot be identified.[222] Other laws, like GDPR, do not expressly provide less restrictive provisions for less identifiable data, but instead cite pseudonymization as a method to meet legal requirements for use, disclosure, or secure maintenance of data.[223]

Turning to the UPDPA and CAVACO statutes, the UPDPA's three

[219] 2016 O.J. (L 119), art. 4(5).

[220] *See, e.g.,* 45 C.F.R. § 164.514(e).

[221] 45 C.F.R. § 164.514(e).

[222] 45 C.F.R. § 46.104(d)(4).

[223] *See, e.g.,* 2016 O.J. (L 119), art. 89 (citing pseudonymization as an example for a data safeguard that can be used when disclosing information for research or public interest purposes).

categories of data identifiability are personal data, deidentified data, and non-identified data. Personal data are the central focus of the Act. A data record is "personal data" if it "direct[ly] identif[ies]"[224] the data subject or if it has been "pseudonymized," meaning that it does not directly identify the subject but "can be reasonably linked to a data subject's identity or is maintained to allow individualized communication with, or treatment of, the data subject."[225] The three CAVACO statutes define "personal data" in ways similar, but not quite identical, to the UPDPA.[226] All include pseudonymized data in personal data.

In practice, the UPDPA employs the term "pseudonymized" in only three places: eliminating the controller's responsibility to provide the data subject a copy of data if the data are "pseudonymized and not maintained with sensitive data";[227] defining the creation of pseudonymized data as a compatible data practice;[228] and prohibiting reidentification of pseudonymized data unless certain conditions are met.[229]

The CAVACO statutes introduce an additional requirement to the definition of pseudonymized data: "that the additional information is kept separately and is subject to technical and organizational measures to ensure that the personal information is not attributed to an identified or identifiable consumer."[230] Given that IRBs typically expect researchers to explain how they will achieve these very tasks, the UPDPA and CAVACO statute definitions of pseudonymized data do not appear more stringent than current research practices, though the UPDPA might be less so.

"Deidentified data"—"personal data that is modified to remove all direct identifiers and to reasonably ensure that the record cannot be linked to an identified data subject by a person that does not have personal knowledge or special access to the data subject's information"[231]—is subject to some

[224] "'Direct identifier' means information that is commonly used to identify a data subject, including name, physical address, email address, recognizable photograph, and telephone number." UPDPA, *supra* note 4, § 2(6).

[225] "The term [pseudonymized] includes a record without a direct identifier if the record contains an internet protocol address, a browser, software, or hardware identification code, a persistent unique code, or other data related to a particular device. The term does not include deidentified data." UPDPA, *supra* note 4, § 2(14).

[226] CCPA, *supra* note 4, § 140(v)(1)(K); VCDPA, *supra* note 4, § 59.1-575; CPA, *supra* note 4, § 6-1-1303(17) (identical to VCDPA).

[227] UPDPA, *supra* note 4, § 5(a). To do otherwise would be exceptionally difficult because the pseudonymization makes it difficult to know whose record belongs to who or whose needs correction; and may actually compromise privacy more through the reidentification process.

[228] UPDPA, *supra* note 4, § 7(b)(5).

[229] *Id.* § 9(b).

[230] CCPA, *supra* note 4, § 140(aa). *Accord* VCDPA, *supra* note 4, § 59.1-575; CPA, *supra* note 4, § 6-1-1303(22). This language mirrors the GDPR. *See supra* note 219.

[231] UPDPA, *supra* note 4, § 2(5).

restrictions under the UPDPA but is not its focus.[232] Because deidentified data are personal data that are modified, we can also think of them as "personal data, but for the fact that they've been deidentified." The California statute defines "deidentified data" similarly to the UPDPA.[233] Virginia and Colorado's statutes narrow the definition slightly, considering data to be deidentified only if it cannot be linked to the data subject or "a device linked to" the data subject.[234] These acts probably thus consider indirect identifiers, such as IP and MAC addresses on computers, sufficient to identify a data subject through a device linked to them. The Colorado and California acts also require—in very similar language—controllers and processors of deidentified data to take certain steps to keep it from being reidentified.[235]

As noted above, deidentified data are practically difficult to keep that way. In theory, statutes could specify standards for deidentification to resolve just this issue, but neither the UPDPA nor the CAVACO statutes do so.[236]

The third data category of identifiability, one not actually named or described in the UPDPA or CAVACO statutes, can be defined by elimination and consists of data about entities other than human data subjects. These acts do not regulate use of such "non-personal data."

C. Data Sensitivity

Assuming that data records are identifiable, there is still a question of how sensitive they are. The extant privacy acts appear to recognize at least three levels of data record sensitivity: "sensitive" personal data, publicly available personal data, and everything else, what we'll call "general personal data." Publicly available data includes public government records and information "available to the general public in widely distributed media," including most widely available websites, directories, media programs, and news media.[237] "Sensitive data" is

[232] *Id.* § 9(b) provides it is a "prohibited data practice to collect or create personal data by reidentifying or causing the reidentification of pseudonymized or deidentified data." The same section provides some technical exceptions to that rule. *Id.*

[233] CCPA, *supra* note 4, § 140(m).

[234] VCDPA, *supra* note 4, § 59.1-575; CPA, *supra* note 4, § 6-1-1303(11).

[235] CCPA, *supra* note 4, § 140(m); CPA, *supra* note 4, § 6-1-1301(11). The UPDPA practically includes similar provisions, but it does not tie them to the definition of "deidentified data." *See* UPDPA, *supra* note 4, § 9(b) (making it a prohibited practice for any regulated entity to "collect or create personal data by reidentifying . . . deidentified data"); § 6(a) (requiring disclosure in the controller's privacy policy of uses); and § 4 (requiring controllers and processors to comply with instructions of, and obligations laid on, collecting controllers).

[236] Oddly, the Colorado statute, which already limits the duties of controllers and processors where deidentified data are concerned, places data deidentified under the standards in 45 C.F.R. 164 entirely outside its application. § 6-1-1304(2)(g).

[237] UPDPA, *supra* note 4, § 2(15).

information in categories defined by the statute that are usually subject to greater protections or more processing restrictions.[238] General personal data is a catch-all category that consists of personal data that is neither publicly available nor sensitive.

The UPDPA recognizes these three levels of personal-data sensitivity. It defines "publicly available information" to include public government records; information "available to the general public in widely distributed media," including most widely available websites, directories, media programs, and news media; information made available to the public lawfully; and observations of the data subject made "from a publicly accessible location."[239] The UPDPA excludes such data entirely from its protection, not considering them part of "personal data."[240] Though the CAVACO statutes vary in their terms from the UPDPA, they appear practically to have similar meanings, and they also exclude publicly available information from their coverage.[241]

The UPDPA defines "sensitive data" as "personal data that reveals" any information in a broad range of categories: "racial or ethnic origin, religious belief, gender, sexual orientation, citizenship, or immigration status"; "a credit or debit card number or financial account number"; most government-issued identification numbers, including SSN, taxpayer ID, etc.; present geolocation coordinates; "diagnosis or treatment for a disease or health condition" or "genetic sequencing information"; criminal records; and any "information about a data subject the controller knows or has reason to know is under 13 years of age."[242] It also includes a subject's ID and password for services to be accessed remotely.[243] Of these, criminal record and income are unique to the UPDPA. There are other variations between the UPDPA and the CAVACO statutes and among them that are interesting, but mostly minor.[244]

[238] ALI's principles do not define sensitive data categories, but the drafters nevertheless claim that the principles are adaptable to concerns about sensitive data. PRINCIPLES OF DATA PRIVACY, *supra* note 1, § 2 cmt. e. For a list of data categories considered sensitive under the UPDPA and CAVACO statutes, see *infra* Section II(C).

[239] UPDPA, *supra* note 4, § 2(15).

[240] UPDPA, *supra* note 4, § 3(c).

[241] CCPA, *supra* note 4, § 140(v)(2); VCPDA § 59.1-575; CPA, *supra* note 4, § 6-1-1303(17)(b). Note that Solow-Niederman expresses concern about the negative externalities of processing of publicly-available data. Solow-Niederman, *supra* note 45, at 5, 31-38.

[242] UPDPA, *supra* note 4, § 2(17).

[243] *Id.* ("credentials sufficient to access an account remotely").

[244] The Virginia and Colorado statutes use almost identical language and are the least expansive in covering sensitive data, not including account credentials; financial accounts and credit and debit card numbers; Social security, taxpayer ID, driver's license, or military identification number; or geolocation. VCPDA § 59.1-575; CPA, *supra* note 4, § 6-1-1303(24). California and Colorado cover "sex life," while Virginia does not. CCPA, *supra* note 4, § 140(ae)(2)(c); VCPDA § § 59.1-575; CPA, *supra* note 4, § 6-1-1303(24). California alone covers philosophical beliefs, union membership and "contents of a consumer's mail, email, and text

The “sensitive data” category varies in its importance in the statutes, as well. Its key role in the UPDPA is to differentiate between cases where the data subject must opt in to restricted data practices (called “incompatible data practices” in the Act) involving sensitive data via “express consent in a signed record for each practice.”[245] The controller need only provide notice and the opportunity to opt-out of incompatible data practices using non-sensitive data.[246] The significant effect of the “sensitive” category under the California statute is that data subjects have certain rights to restrict their use, though the statute expresses this in a confused jumble of limitations and exceptions.[247] The California act also provides for specific means for the data subject to opt out of disclosure and distribution of their sensitive data.[248] Virginia and Colorado require consent for any data practice involving sensitive data.[249] Each also requires that controllers and processors perform a “data protection assessment” for processing where sensitive data are concerned.[250]

The third, catch-all category of data sensitivity, what we call “general personal data,” is not named or defined in the UPDPA or CAVACO statutes, but consists of personal data that is neither publicly available nor sensitive data.

D. Regulated Entity

Central to many data protection laws is a delineation of particular types of data controllers or processors subject to the law, in other words, the regulated entities. In comprehensive data protection laws, the definition of the regulated entity is often broad. GDPR applies to processing of personal data by controllers and processors established within the European Union—the location of the regulated entity—and “personal data of data subjects who are in the Union by a

messages unless the business is the intended recipient of the communication.” CCPA, *supra* note 4, § 140(ae)(1)(D)-(E). It also allows the listed to be extended by regulation. CCPA, *supra* note 4, § 185(a)(1). There are some variations in the identification of geolocation, biometric, and genetic data among the statutes. UPDPA, *supra* note 4, § 2(17); CCPA, *supra* note 4, § 140(ae)(1)(c), (1)(f), (2)(a); VCPDA § 59.1-575; CPA, *supra* note 4, § 6-1-1303(24).

[245] UPDPA, *supra* note 4, § 8(c).

[246] The notice must be sufficient for the “data subject to understand the nature of the incompatible data processing.” UPDPA, *supra* note 4, § 8(b). The UPDPA also affects the data subject’s right to request a copy of data from a controller. *Id.* § 5(a).

[247] *See* CCPA § 121. California also subjects a controller to greater disclosure obligations to the data subject regarding the collection of sensitive data. CCPA, *supra* note 4, § 100(a)(2)-(3).

[248] CCPA, *supra* note 4, § 135.

[249] VCDPA, *supra* note 4, § 59.1-578(A)(5); CPA, *supra* note 4, § 6-1-1308(7).

[250] CPA, *supra* note 4, § 6-1-1309(2) (categorizing the processing of any sensitive data as “processing that presents a heightened risk of harm to a consumer”); *id.* § 6-1-1309(1) (requiring data protection assessments for practices that present a heightened risk); VCDPA, *supra* note 4, § 59.1-580(A)(4) (requiring data protection assesments for practies that involve the “processing of senstivie data”)..

controller or processor not established in the Union"—the location of the data subject at the time of the processing.[251] GDPR also defines some entities that are not regulated (e.g., natural persons engaged with personal or household activities).[252] In existing U.S. federal laws, the limited scope of separate statutes results in the sectorial "patchwork" of regulation, which is not particularly analytically useful with the comprehensive state statutes discussed here. The newer statutes do a more thorough job of conceptually identifying various controllers and processors in the "pipeline" of data processing.[253]

Importantly, U.S. data protection laws are not mutually exclusive when it comes to the defined regulated entities. For example, most entities regulated as substance-abuse treatment programs are also HIPAA-covered entities. Consequently, they have to comply with HIPAA and the 42 CFR Part 2 regulations. This also creates complexities between federal and state regulatory approaches. For example, health information exchange organizations are regulated under HIPAA as business associates of covered entities,[254] but in 2016, thirty-one states had privacy laws specifically regulating health information exchanges.[255] When different data protection laws overlap on a single regulated entity, it can be especially difficult to determine which legal provisions apply and which policies to implement to ensure compliant data practices.

Turning to the UPDPA, at its broadest level, it applies to any person—whether individual or legal entity[256]—that is a controller or processor of personal data, provided the controller or processor "conducts business in [the adopting] state or produces products or provides services purposefully directed to residents of" the adopting state.[257] Like the CAVACO statutes, the UPDPA excludes from its effect the adopting state and any "agency or instrumentality . . . or a political subdivision" of it.[258] Not-for-profit enterprises may or may not be covered,

[251] 2016 O.J. (L 119), art. 3.

[252] *Id.* art. 1–2, 18.

[253] *See* text accompanying notes 183–190.

[254] 45 C.F.R. § 160.103 (2021).

[255] Cason D. Schmit, Sarah A. Wetter & Bita A. Kash, *Falling Short: How State Laws Can Address Health Information Exchange Barriers and Enablers*, 25 J. AM. MED. INFORMATICS ASS'N 635, 635–644 (2018).

[256] The definition of "person" includes both individuals and entities but excludes any "public corporation or government or governmental subdivision, agency, or instrumentality." UPDPA, *supra* note 4, § 2(9).

[257] UPDPA, *supra* note 4, § 3(a).

[258] *Id.* § 3(b); *see* CCPA, *supra* note 4, § 140(d)(1) (defining "business"—the entities regulated under the act—as any "sole proprietorship, partnership, limited liability company, corporation, association, or other legal entity that is organized or operated for the profit or financial benefit of its shareholders or other owners," thus implicitly excluding government entities); VCDPA, *supra* note 4, § 59.1-576(B) (withholding application from "body, authority, board, bureau, commission, district, or agency of the Commonwealth or of any political subdivision of the Commonwealth"); COLO. REV. STAT. ANN. § 6-1-102 (West 2021) (defining, for purposes of CPA,

depending on state-law determinations about what counts as "conducting business." The Colorado act is silent on that matter. California, meanwhile, defines the businesses to which CCPA applies as those "organized or operated for the profit or financial benefit of its shareholders or other owners," seemingly excluding non-profits.[259] Virginia's act expressly excludes from its application any non-profit organization[260] or "institution of higher education."[261]

Like the CAVACO statutes, the UPDPA has certain size thresholds for regulated entities. A controller or processor that "maintains personal data about more than [50,000] data subjects who are residents of this state"[262] or that "earns more than [50] percent of its gross annual revenue during a calendar year from maintaining personal data as a controller or processor" is fully subject to the UPDPA.[263] It's up to each enacting state to fill in the bracketed thresholds.[264] Similarly, the California Consumer Privacy Act applies to a smaller entity if it "[d]erives 50 percent or more of its annual revenues from selling or sharing consumers' personal information."[265] The Virginia Consumer Data Protection Act and Colorado Privacy Act never apply to smaller controllers or processors.[266]

"person" as "an individual, corporation, business trust, estate, trust, partnership, unincorporated association . . . , or any other legal or commercial entity," again implicitly excluding government entities).

[259] CCPA, *supra* note 4, § 140(d)(1).

[260] Defined as "any corporation organized under the Virginia Nonstock Corporation Act . . . or any organization exempt from taxation under § 501(c)(3), 501(c)(6), or 501 (c)(12) of the Internal Revenue Code." VA. CODE ANN. § 59.1-575 (West 2021).

[261] VA. CODE ANN. § 59.1-576(B)(iv)-(v) (West 2021).

[262] UPDPA, *supra* note 4, § 3(a)(1) ("excluding data subjects whose data is collected or maintained solely to complete a payment transaction"). Note that the square brackets in the quoted language in the original. Whether a data subject is protected by a state's adoption of the UPDPA appears to be unrelated to whether the data subject is a resident of the adopting state. This is because the definition of regulated entities noted above relates to whether the controller or processor does business in the adopting state or purposefully directs its services to the state's residents and not whether any breach involves data records of a resident of the adopting state. *See* the discussion, *infra* Section H, for implications in enforcement.

[263] UPDPA, *supra* note 4, § 3(a)(2). A processor working for a controller or processor that meets either of these size requirements is also held to be in this category. UPDPA, *supra* note 4, § 3(a)(3).

[264] "The threshold numbers are in brackets [so] each State can determine the proper level of applicability." UPDPA, *supra* note 4, § 3 cmt.

[265] CCPA, *supra* note 4, § 140(d)(1)(C). Otherwise, CCPA governs only larger controllers and processors, those that have "annual gross revenues in excess of twenty-five million dollars ($25,000,000) in the preceding calendar year" or that "annually buy[], sell[], or share[] the personal information of 100,000 or more consumers or households." *Id.* § 140(d)(1)(A)–(B).

[266] CPA, *supra* note 4, § 6-1-1304(1) (applying only to a controller or processor that "controls or processes the personal data of one hundred thousand consumers or more . . . [or] derives revenue . . . from the sale of personal data and processes or controls the personal data of twenty-five thousand consumers or more"); VCDPA, *supra* note 4, § 59.1-576(A) (processors and controllers that "control or process personal data of at least 100,000 consumers or . . . control or

Normatively, these acts are practically equivalent on the issue of covered entities, but one concern under the UPDPA is its coverage of smaller players. A controller or processor *of any size* is subject to the UPDPA if it engages in any of the "restricted" or "incompatible" data practices described below.[267] On the one hand, it's unclear how much expense smaller players will have to incur to educate themselves about the Act so that they understand what they may do without becoming subject to *all* of the UPDPA's requirements. The result might be widespread confusion, and a catastrophic implementation of the Act in a state that affects small-business owners could sour legislators on the act in general. On the other hand, exempting small controllers and processors—who likely make up a large proportion of the players in this space—could leave much data entirely unprotected, much as they are by the CAVACO statutes.

E. Data Practices

Our framework recognizes three types of data practices in which controllers and processors may engage: *favored*, *restricted*, and *prohibited* data practices. Favored and restricted data practices each have two subcategories. Those that are favored may be disclosed or undisclosed and do not require data subject's consent; those that are restricted require the data subject's consent, passively through an opt-out or actively through an opt-in mechanism. Thus, permitted data practices represent a continuum from those that least constrain the controller, undisclosed favored; to those that most constrain it, active-consent restricted. All other data practices are prohibited.

1. Favored Data Practices

Generally, data protection laws will permit the use of collected data for enumerated purposes without any consent from the data subjects other than their choice to enter a relationship with the controller. These favored practices will almost always include the primary data use, or the use for which the data was collected. This "purpose limitation" often intends that "personal information should be collected only for a specified purpose and not further processed in a manner incompatible" with it.[268] For example, HIPAA permits covered entities to use protected information for treatment, payment, and healthcare operations. Similarly, FERPA permits educational entities to use protected education records for legitimate educational interests. These purposes align with reasonable data-subject expectations for the use of collected data.

process personal data of at least 25,000 consumers and derive over 50 percent of gross revenue from the sale of personal data").

[267] UPDPA, *supra* note 4, § 3(a)(4).

[268] PRINCIPLES OF DATA PRIVACY, *supra* note 1, at 3.

Data protection laws may also permit some secondary data uses—data collected for one purpose but reused for another purpose—without a data subject's consent. Secondary data uses may be favored data practices if they advance government interests, data subjects' interests, or social interests. A secondary data use could advance a government interest if it facilitates government oversight or enforcement (e.g., fraud detection). Similarly, a secondary data use could promote the data subject's interest, as, for example, when federal public assistance programs permit program data to be used to assess a beneficiary's eligibility for additional benefits. Finally, some laws permit some secondary uses without consent to advance social interests, as when they permit data to be used for research or public health purposes.[269] All these favored uses can be either *disclosed*, meaning that the collecting controller discloses—usually in a privacy policy—that it will engage in the data practice, or *undisclosed*, meaning that the controller does not disclose them.

The basic regime of the UPDPA is to permit what it calls "compatible data practices" without consumer consent, though the collecting controller must disclose those favored data practices in which it routinely engages in its privacy policy. These are thus disclosed favored practices in our framework. There are three bases upon which a data practice can be a compatible data practice under the UPDPA. The most straightforward basis is for the practice to fall within an enumerated list of compatible practices: section 7(b)–(c) of the Act. This includes managing transactions between controller and data subject and managing controller's business—both part of the primary purposes for which the data are collected—and permitting oversight of controller's data practices, preventing or investigating crime, complying with legal requirements, and defending against legal claims—data practices that the drafters regarded as sufficiently integral to the primary purposes of the data collection to warrant this status.[270]

The second basis upon which a data practice may be classified as compatible under the UPDPA is if it entails "processing [that (1)] is consistent with the ordinary expectations of data subjects or [(2)] is likely to benefit data subjects substantially."[271] Note that elements (1) and (2) here are disjunctive, so either will do. The Act offers six factors for assessing whether a particular data practice would satisfy this requirement.[272]

[269] Hulkower, *supra* note 15, 150–60; *see generally* Tara Ramanathan, Cason Schmit, Akshara Menon & Chanelle Fox, *The Role of Law in Supporting Secondary Uses of Electronic Health Information*, 43 J. L. MED. & ETHICS 48 (2021).

[270] UPDPA, *supra* note 4, § 7(b).

[271] *Id.* § 7(a).

[272] *Id.* ((1) the data subject's relationship with the controller; (2) the type of transaction in which the personal data was collected; (3) the type and nature of the personal data that would be processed; (4) the risk of a negative consequence on the data subject by the use or disclosure of the personal data; (5) the effectiveness of a safeguard against unauthorized use or disclosure of the

The third basis under the UPDPA for classifying a data practice as compatible is in accordance with a voluntary consensus standard (VCS). This is a formal standard that a controller or processor can adopt, developed (probably) by an industry group in consultation with consumers and others, and approved by the attorney general (or other privacy official designated by the enacting state). As the VCS is a significant innovation of the UPDPA that provides value to public health researchers and professionals, we treat it in more detail below.[273]

Under the UPDPA, the collecting controller must disclose in its privacy policy any compatible data practices it or its authorized processors "appl[y] routinely to personal data."[274] The UPDPA's use of the word "routinely" seems unnecessarily vague here. For example, a controller may disclose personal data that provides evidence of criminal activity to a law enforcement agency without listing this practice" if "this type of disclosure is unusual."[275] There is no definition of "routinely" in the UPDPA, and it does not appear in other uniform acts of the ULC. Even *Black's* struggles to define "routine practice" without appeal to the synonym "regular": "A customary action or procedure that is regularly followed; a habitual method adhered to as a matter of regularity."[276]

The California act does not require specific consent for data practices "reasonably necessary and proportionate to achieve the purposes for which the personal information was collected or processed, or for another disclosed purpose that is compatible with the context in which the personal information was collected."[277] It does require that the collecting controller disclose the categories of personal information (including sensitive data), its expected uses, and the duration of its retention.[278] Virginia and Colorado also require these disclosures[279] and do not require consent for "collection of personal data to what is adequate, relevant, and reasonably necessary in relation to the purposes for which such data is processed, as disclosed to the consumer"[280] or processing for those purposes or for purposes "compatible" with them,[281] provided the data are

personal data; and (6) the extent to which the practice advances the economic, health, or other interests of the data subject.)

[273] *See infra* Section II.G.

[274] UPDPA, *supra* note 4, § 6(a)(3).

[275] *Id.* § 6 cmt.

[276] *Routine Practice*, BLACK'S LAW DICTIONARY (11th ed. 2019).

[277] CCPA, *supra* note 4, § 100(c).

[278] *Id.* § 100(a).

[279] VCDPA, *supra* note 4, § 59.1-578(C); CPA, *supra* note 4, § 6-1-1308(1) (using language very similar to Virginia's).

[280] VCDPA, *supra* note 4, § 59.1-578(A)(1); *see also* CPA, *supra* note 4, § 6-1-1308(3) (using very similar language).

[281] VCDPA, *supra* note 4, § 59.1-578(A)(2): *see also* CPA, *supra* note 4, § 6-1-1308(4) (using very similar language).

not sensitive.[282]

The UPDPA and the CAVACO statutes differ from each other somewhat in their overt treatment of public health. The UPDPA classifies as a compatible data practice—a disclosed favored practice—one that "permits analysis . . . to discover insights related to public health, public policy, or other matters of general public interest and does not include use of personal data to make a prediction or determination about a particular data subject."[283] This provision also appears to permit public health surveillance and development of population interventions to protect public health, but it specifically excludes individualized interventions.[284] California establishes a narrow undisclosed favored practice for public health: Reidentification of deidentified records for public health purposes and for research subject to the Common Rule.[285] Colorado, on the other hand, offers a broad permission for public health practices, providing that the act does not "restrict a controller's or processor's ability . . . to process personal data for reasons of public interest in the area of public health, but solely to the extent that the processing . . . (a) is subject to suitable and specific measures to safeguard the rights of the consumer whose personal data are processed; and (b) is under the responsibility of a professional subject to confidentiality obligations under federal, state, or local law."[286] This is also an undisclosed favored practice in our framework. The California and Virginia acts treat public health practices as restricted data practices, thus requiring consent, though the consent need only be passive (opt out) in California's case but must be active (opt in) in Virginia's. See the next subsection for further discussion.

2. *Restricted Data Practices*

Restricted data practices are those that require the data subject's consent. There are two subsets of restricted data practices: *passive consent* and *active consent*. They represent default states for data practices. In passive consent, the data subject is presumed to consent unless they opt out; in active consent, the data subject is presumed not to consent unless they opt in. There may also be heightened requirements for notice and more formal requirements for consent for

[282] *See* VCDPA, *supra* note 4, § 59.1-578(A)(5).

[283] UPDPA, *supra* note 4, § 7(b)(6)(A). In fact, the controller has to disclose the data use only if it is "routine."

[284] UPDPA, *supra* note 4, § 7 cmt. (A compatible practice "would include the use of personal data to initially train an AI or machine learning algorithm. However, subsequent use of such an AI or machine learning algorithm in order to make a prediction or decision about a data subject . . . must comply with this act through another provision.").

[285] CCPA, *supra* note 4, § 148(a)(2), (3).

[286] *Id.* § 6-1-1304(3)(a)(xi).

some restricted data practices.[287]

The UPDPA refers to restricted data practices as "incompatible data practices."[288] Despite their name, the UPDPA does not prohibit them, instead merely requiring the data subject's consent. There is considerable variation in the acts' determinations of which restricted data practices are passive-consent, permitting data subjects to opt out, and active-consent, requiring data subjects to opt in. The UPDPA and California require active consent in the smallest class of cases, while Virginia and Colorado appear to require active consent in a broad class of cases.

Considering passive consent first, when the data controller collects data for an incompatible data practice under the UPDPA, the subject must be informed and have a chance to opt out.[289] The California act provides a data subject an opt-out right to "to direct a business that sells or shares personal information about the consumer to third parties" without regard to the reason for which the controller is selling or sharing data.[290] Similarly, uses of sensitive data outside those that are favored give rise to a data subject's right to opt out in California.[291] Virginia and Colorado provide that data subjects may opt out of "(i) targeted advertising, (ii) the sale of personal data, or (iii) profiling in furtherance of decisions that produce legal or similarly significant effects concerning the consumer."[292] Based on these provisions, a controller will have to provide at the

[287] And there may be a variety of kinds of consent. As background, 2017 revisions to the Common Rule introduced a new type of consent, called "broad consent." *Revised Common Rule FAQs*, HHS.GOV OFFICE FOR HUMAN RESEARCH PROTECTIONS, https://www.hhs.gov/ohrp/education-and-outreach/revised-common-rule/revised-common-rule-q-and-a/index.html [https://perma.cc/ZXE9-LCHC] (last visited Feb. 12, 2022). This new provision allows researchers to solicit consent that covers a broad range of potential research applications. *Id.* Rather than seeking specific consent for each new research project. Anecdotally, we believe that IRBs are struggling to practically implement a "broad consenting" process and that it is consequently an underutilized legal tool. It may be that "consent" in most commercial settings—click-through privacy policies—is a lot like a broad consent but without the rigor of IRB review.

[288] UPDPA, *supra* note 4, § 8(a) (defining the term by process of elimination, labeling data practices that are not compatible or prohibited "incompatible," and also including violations of a privacy policy).

[289] *Id.* § 8(b); *see also* UPDPA, *supra* note 4, § 6 (requiring a collecting controller to have a privacy policy that identifies categories and purpose of data it maintains and distributes to others and identifies all incompatible data practices it will apply unless the consumer opts out).

[290] CCPA, *supra* note 4, §§ 120(a), 115(d). The act's authorization of regulations, however, suggests that the reasons might be spelled out. § 185(a)(19)(A)(vi). *See also* CCPA, *supra* note 4, § 120(b) (requiring a controller to disclose any selling or sharing of data in which it engages).

[291] CCPA, *supra* note 4, § 121(a). *See also id.* § 135 (detailing methods for providing this opt out).

[292] VCDPA, *supra* note 4, § 59.1-577(A)(5); CPA, *supra* note 4, § 6-1-1306(1)(a)(i) (using identical language). "'Targeted advertising' means displaying advertisements to a consumer where the advertisement is selected based on personal data obtained from that consumer's activities over time and across nonaffiliated websites or online applications to predict such consumer's

least notice and an opportunity to opt out before providing data for public health practices or research if they cannot be considered favored practices but are instead restricted practices. Practically speaking, this is not much more of an impediment than that imposed for disclosed favored practices: With passive consent, the default is participation, and harried data subjects are unlikely even to notice that they may opt out. However, in contrast to disclosed favored practices, data controllers seeking to share passive-consent data for public health have implementation costs to develop systems and workflows to collect, manage, and enforce opt-out preferences.

But the UPDPA and the California and Virginia acts include some data practices that require active consent. The Virginia statute provides that all data practices beyond the favored ones described above, and any processing involving sensitive data, are subject to the data subject's consent.[293] As it defines consent as "a clear affirmative act signifying a consumer's . . . agreement to process personal data relating to the consumer,"[294] this appears to be an opt-in form of consent. The Colorado statute's requirements are similar, but it classifies public health activities as favored practices that do not require consent. In California, a very small class of cases—where the controller wants to enroll the data subject in "into a financial incentive program"[295]—are subject to active consent. Under the UPDPA, only where sensitive data[296] are concerned must the data subject consent specifically to each incompatible data practice.[297]

preferences or interests." VCDPA, *supra* note 4, § 59.1-575; *see also* CPA, *supra* note 4, § 6-1-1301(25) (adopting very similar language. "'Profiling' means any form of automated processing performed on personal data to evaluate, analyze, or predict personal aspects related to an identified or identifiable natural person's economic situation, health, personal preferences, interests, reliability, behavior, location, or movements." VCDPA, *supra* note 4, § 59.1-575; *see also* CPA, *supra* note 4, § 6-1-1301(20) (adopting nearly identical language).

[293] VCDPA, *supra* note 4, § 59.1-578(A)(2), (5); *see also* CPA, *supra* note 4, § 6-1-1308(4), (7) (using very similar language).

[294] VCDPA, *supra* note 4, § 59.1-575; *see also* CPA, *supra* note 4, § 6-1-1303(5) (using very similar language). California appears at first to define consent more broadly as "any freely given, specific, informed, and unambiguous indication of the consumer's wishes by which the consumer . . . , including by a statement or by a clear affirmative action, signifies agreement to the processing of personal information relating to the consumer for a narrowly defined particular purpose." CCPA, *supra* note 4, § 140(h). The "including" before "by a statement or by a clear affirmative action" suggests there are other possibility. The section proceeds to provide that "[a]cceptance of a general or broad terms of use, or similar document, that contains descriptions of personal information processing along with other, unrelated information, does not constitute consent. Hovering over, muting, pausing, or closing a given piece of content does not constitute consent. Likewise, agreement obtained through use of dark patterns does not constitute consent." *Id. See generally* Jamie Luguri & Lior Jacob Strahilevitz, *Shining a Light on Dark Patterns*, 13 J. LEGAL ANALYSIS 43 (2021) (providing an explanation and analysis of dark patterns).

[295] CCPA, *supra* note 4, § 1798.125(b)(3).

[296] *See supra* Section II(C).

[297] UPDPA, *supra* note 4, § 8(c).

3. *Prohibited Data Practices*

Prohibited data practices are those practices that are never permitted. The CAVACO statutes do not define prohibited data practices, except to the extent that prohibition arises from going beyond what is permitted in favored and restricted practices.[298] In contrast, the UPDPA expressly describes several prohibited data practices.[299] As a preliminary matter, the UPDPA makes it a prohibited practice to reidentify deidentified data, subject to certain exceptions.[300] This Section thus brings deidentified data within the UPDPA's scope, but only to the extent that a processor attempts to reidentify it. The UPDPA inventories other categories of prohibited data practices into three groups: breaking rules elsewhere, personal harms, and security harms. The Act prohibits data processing if the processor engages in processing that would otherwise be a restricted ("incompatible") data practice and fails to get the data subject's consent.[301]

The UPDPA also makes it a prohibited data practice to process personal data in a manner that would "constitute a violation of other law, including federal or state law against discrimination."[302] The Virginia and Colorado acts contain similar prohibitions.[303]

The personal harms against which the UPDPA protects data subjects arise from data practices likely to "subject a data subject to specific and significant: (A) financial, physical, or reputational harm; (B) embarrassment, ridicule, intimidation, or harassment; or (C) physical or other intrusion on solitude or seclusion."[304] These UPDPA strictures could have effect on some public health practices.[305] For example, individualized public health interventions might under certain circumstances have the negative effects described in the UPDPA. The CAVACO statutes do not call out these particular harms as relating to prohibited data practices, again, because they do not specifically define prohibited practices.

The security harms against which the UPDPA protects data subjects arise from data practices likely to "result in misappropriation of personal data to assume another's identity," or "fail to provide reasonable data-security measures."[306] The CAVACO statutes imply similar requirements in their overall use limitations and in their requirements for risk assessments.[307]

[298] *See, e.g.,* CCPA, *supra* note 4, §§ 100(a), 100(c), 120(d), 121(b).

[299] UPDPA, *supra* note 4, § 9(a).

[300] *Id.* § 9(b).

[301] *Id.* § 9(a)(5).

[302] *Id.* § 9(a)(3).

[303] VCDPA, *supra* note 4, § 59.1-578(A)(4); CPA, *supra* note 4, § 6-1-1308(6).

[304] UPDPA, *supra* note 4, § 9(a)(1).

[305] *See infra* Part III.

[306] UPDPA, *supra* note 4, § 9(a)(2), (4).

[307] CCPA, *supra* note 4, § 1798.185(a)(15); VCDPA, *supra* note 4, § 59.1-580; CPA, *supra*

F. Other Requirements of Controllers and Processors

Recall that a smaller data controller or processor that engages only in compatible data practices is not bound to meet any other requirements under the UPDPA.[308] As for the larger controller or processor, or the smaller one that wishes to engage in incompatible data practices, the UPDPA's key requirements are to engage in incompatible data practices only with the data subject's consent (opt-in or opt-out, depending on data-content sensitivity) and not to engage in prohibited data practices. The UPDPA imposes other obligations on these data controllers and processors. They fall into three categories: offering a public privacy policy, responding to data subject's requests, and performing data risk assessments.

The UPDPA requires that a controller make its privacy policy available in two ways: First, it must be "reasonably available to a data subject at the time personal data is collected about the subject," and second, the controller must post its privacy policy on its website, if it has one.[309] The CAVACO statutes do not impose the latter requirement. As for the contents of privacy policies, they fall into two categories, one relating to the controller's data practices and the other to the procedures and laws under which it operates. The UPDPA and the CAVACO statutes have similar requirements for privacy policies regarding data practices, discussed above.[310] Where procedures and laws are concerned, the UPDPA and the CAVACO statutes require that the privacy policy provide "the procedure for a data subject to exercise a right" requiring the controller's response.[311] Under the UPDPA, the controller must also identify "federal, state, or international privacy laws or frameworks with which the controller complies," and explain whether the controller has adopted "any voluntary consensus standard."[312]

The second major category of responsibilities for data controllers under the UPDPA and CAVACO statutes involves responding to requests from data subjects, including requests for copies of data, for correcting data, and for deleting data. The collecting controller is principally responsible here because it has (or had) a relationship with the data subject at the time of collection. The collecting controller is responsible for providing to a data subject a copy of their personal data and correcting errors in the data.[313] The data controller is responsible for coordinating activities of processors and downstream controllers

note 4, § 6-1-1309.

[308] UPDPA, *supra* note 4, § 3(a)(1)–(4).

[309] UPDPA, *supra* note 4, § 6(b)-(c).

[310] *Supra* Section II.F.

[311] UPDPA, *supra* note 4, § 6(a)(5); accord VCDPA, *supra* note 4, § 59.1-578(C); CPA, *supra* note 4, § 6-1-1308(1)(a)(iii).

[312] UPDPA, *supra* note 4, § 6(a)(5)-(7).

[313] *Id.* §§ 4(a)(1)-(2), 5(a).

to comply with these requirements, and those processors and controllers are bound to cooperate.[314] The controller may not retaliate against a data subject for making any of these requests.[315] California, Virginia, and Colorado all provide that the controller must comply with a data subject request to delete personal data.[316] The UPDPA does *not* provide a right for the data subject to request the deletion of personal data.[317] Nevertheless, all four statutes provide some individual rights that persist throughout the data processing lifecycle, which some legal scholars argue is characteristic of the European GDPR.

The CAVACO statutes provide for a duty of care "to implement and maintain reasonable security procedures and practices appropriate to the nature of the information."[318] The UPDPA makes it a prohibited data practice to "fail to provide reasonable data-security measures, including appropriate administrative, technical, and physical safeguards to prevent unauthorized access."[319]

The final major category of responsibility for data controllers and processors under the UPDPA is that they must "conduct and maintain . . . a data privacy and security risk assessment" that addresses risks, their characteristics, and efforts taken to mitigate them.[320] The California statute provides for regulations addressing risk assessments, but regulations promulgated under the previous version of the California Consumer Privacy Act do not address them, despite the statutory requirement that they do so.[321] Colorado and Virginia require assessments for processing of sensitive data and some other data practices.[322] Neither the UPDPA nor the CAVACO statutes directly require periodic updates of risk assessments. Under the UPDPA, the controller or processor must update

[314] *Id.* § 5(b).

[315] *Id.* § 5(c). There are some special cases where the controller can change its relationship with the data subject after changing data at the subject's request or if the subject withholds consent from an incompatible data practice. *Id.* §§ 5(c), 7(c), 8(c).

[316] CCPA, *supra* note 4, § 105(a); VCDPA, *supra* note 4, § 59.1-577; CPA, *supra* note 4, § 6-1-1306(1)(d). *But see* CCPA, *supra* note 4, § 105(d)(6) (providing that a controller need not delete data records at the data subject's request if the data are being processed for research to which the subject consented and "deletion of the information is likely to render impossible or seriously impair the ability to complete" the research).

[317] UPDPA, *supra* note 4, § 4, official comment.

[318] CCPA, *supra* note 4, § 150(a)(1). *Accord* VCDPA, *supra* note 4, § 59.1-578(A)(3); CPA, *supra* note 4, § 6-1-1308(5).

[319] UPDPA, *supra* note 4, § 9(a)(4).

[320] *Id.* § 10(a).

[321] CCPA, *supra* note 4, § 185(a)(15)(B) (requiring the California Attorney General to adopt regulations by July 1, 2020, "requiring businesses whose processing of consumers' personal information presents significant risk to consumers' privacy or security, to . . . [s]ubmit to the California Privacy Protection Agency on a regular basis a risk assessment with respect to their processing of personal information"). As of this writing, no such regulations appear to have been promulgated.

[322] CPA, *supra* note 4, § 6-1-1309(2); VCDPA, *supra* note 4, § 59.1-580(A).

the assessment if "there is a change in the risk environment or in a data practice that may materially affect the privacy or security of the personal data."[323] Language of the CAVACO statutes might be construed to require a new assessment when similar changes occur.[324]

Among these provisions, only the right to deletion raises concerns for public health, and then only if a significant proportion of data subjects request it.

G. The UPDPA Voluntary Consensus Standards

A marked innovation in the UPDPA is its use of VCSs. As one official comment on the Act notes: "[H]ow these obligations are implemented may depend on the particular business sector . . . . [a]nd consumers have vastly different expectations about the use of their personal information depending on the underlying transaction for which their data is sought."[325] According to the UPDPA reporter, "[p]roviding an opportunity for industry sectors, in collaboration with stakeholders including data subjects, to agree on methods of implementing privacy obligations provides the flexibility any privacy legislation will require."[326] The comment notes the apparent success of such standards under the Children's Online Privacy Protection Act (COPPA).[327]

In the UPDPA, the result is a process for groups of stakeholders to gather and set baselines for particular industries or types of project. Such stakeholders could include industry groups and public health researchers and professionals. In brief, a group of "stakeholders"[328] gathers to adopt a set of baselines relating to various requirements of the Act, those not spelled out in the Act itself. For example, what counts as a compatible data practice in a particular industry?[329] The Act categorizes data practices by a controller or processor subject to a VCS as "compatible data practices" if the VCS defines them so.[330] How must a controller obtain consent from data subjects when it is required?[331] What are industry-standard practices for responding to a consumer request for access to

[323] UPDPA, *supra* note 4, § 10(a)-(b).

[324] *See* VCDPA, *supra* note 4, § 59.1-580; CPA, *supra* note 4, § 6-1-1309.

[325] UPDPA, *supra* note 4, § 12, official comment.

[326] *Id.*

[327] *Id. See also* BBB NAT'L PROGRAMS, INC., TWENTY YEARS OF SUCCESSFUL CO-REGULATION UNDER COPPA:
A MODEL FOR FOSTERING CONSUMER PRIVACY (Oct. 2019), https://bbbnp-bbbp-stf-use1-01.s3.amazonaws.com/docs/default-source/whitepapers/bbb-np-report---20-years-of-coppa-self-regulation---10-15-2019.pdf [https://perma.cc/4CBW-ULEM].

[328] UPDPA, *supra* note 4, § 2(19).

[329] *Id.* § 13(1).

[330] *Id.* § 7(d).

[331] *Id.* § 13(2).

and correction of data?[332] A controller must announce in its privacy policy that it is complying with a VCS.[333] A controller that adopts and complies with a VCS setting out those standards is compliant with the UPDPA.[334] This approach offers a frank acknowledgment that data privacy is not a matter of *one size fits all*.[335]

Four sections of the UPDPA's twenty sections, and a considerable proportion of its word count, are dedicated to explaining the effect of VCSs, what they contain, how they are developed, and how they are recognized by the attorney general (or other privacy officer).[336] Key for developing a VCS is that the process must be open and deliberative in a way similar to ULC's own deliberative process, with "stakeholders representing a diverse range of industry, consumer, and public interests," and must give effort to hearing, responding to, and resolving stakeholder concerns.[337] The result does not have to be unanimous, and stakeholders can file "statement[s] of dissent."[338] The attorney general must be satisfied that the group adopted and followed a set of procedures to "provide adequate notice of meetings and standards development."[339] The attorney general evaluates requests to recognize a VCS according to rules the attorney general adopts for the requests.[340] If the attorney general recognizes the VCS, it becomes a public record and thus usable by any regulated entity.[341] The attorney general can later withdraw recognition, if they determine the VCS "or its implementation is not consistent with" the act.[342]

Practically speaking, there is nothing like VCSs in the CAVACO statutes. There are provisions that enable some change and development, however. For example, California's act provides authority for the state's privacy authority to issue and maintain regulations that address changes in technology and providing for many details of the relationship between controller and data subject.[343] It neither expressly permits nor forbids the industry-specific approach that the VCSs contemplate. The Colorado act provides its attorney general a one-time grant of authority to "adopt rules that govern the process of issuing opinion letters and interpretive guidance to develop an operational framework for

[332] *Id.* § 13(3).

[333] *Id.* § 6(a)(7).

[334] *Id.* § 12.

[335] PRINCIPLES OF DATA PRIVACY, *supra* note 1, at 3 (noting that "uniformity and specificity is not always desirable in light of the necessity for contextual shaping of [fair information practices] in different areas of data use").

[336] UPDPA, *supra* note 4, §§ 12–15.

[337] *Id.* § 14(1).

[338] *Id.* § 14(1), (5).

[339] *Id.* § 14(4).

[340] *Id.* § 15(b).

[341] *Id.* § 15(I).

[342] *Id.* § 15(d).

[343] CCPA, *supra* note 4, § 1798.185.

business that includes a good faith reliance defense of an action that may otherwise constitute a violation" of the act.[344] Virginia provides no such mechanisms.

For public health researchers and professionals, a VCS might prove a very valuable way to identify as many of their data practices as possible as being either exempt from the UPDPA or as being disclosed favored practices, what the UPDPA calls "compatible data practices."

H. Enforcement and Penalties

Typically, the remedies and penalties under a statute and who can enforce it are determined by the statute. Professor Cohen describes—and criticizes—conventional enforcement strategies broadly as "private remedial litigation initiated by affected individuals and public enforcement action initiated by agencies." In practice, these penalties can consist of civil damages, civil penalties, injunctions, and criminal penalties. Professor Cohen proposes three alternatives to these conventional approaches that she argues could lead to more impactful enforcement of privacy violations: 1) deputizing online intermediaries to discipline actors within their information ecosystems, 2) disgorgement of profits that accrue from privacy violations, and 3) permitting senior executives to be held personally liable for privacy violations. However, none of Professor Cohen's alternatives—or criminal penalties for that matter—play a significant role in the statutes we discuss in in this Article.

The UPDPA assumes that the adopting state's attorney general (or the state data privacy officer that the adopting state substitutes for the attorney general in the Uniform Act) will have a significant role in enforcement of the Act and adoption of VCSs.[345] As for enforcement authority, though, that depends on the adopting state's consumer protection act, which the UPDPA cross-references for "enforcement authority, remedies, and penalties" under the Act.[346] In some states, this may mean that only the state attorney general may enforce the act, that only the attorney general and local district attorneys may enforce the Act, or that affected data subjects might have their own private rights of action against controllers and processors. Similar variability exists regarding remedies and penalties.

The CAVACO statutes do not take a single approach, either. California provides for a private civil right of action, with actual damages or statutory damages between $100 and $750 per consumer per incident,[347] and power for its

[344] CPA, *supra* note 4, § 6-1-1313(3).

[345] UPDPA, *supra* note 4, § 16.

[346] *Id.* § 16(a).

[347] CCPA, *supra* note 4, § 150(a).

privacy authority to enforce the act administratively, with penalties of $2,500 per incident[348] Both the privacy authority and private litigants can seek injunctions.[349] Virginia allows only its attorney general to enforce its act, seeking injunction, civil penalties up to $7,500 per violation, or both.[350] Colorado provides that its attorney general and district attorneys can bring actions, with remedies the same as Colorado's statute governing deceptive trade practices.[351]

I. Interaction with Other Statutes

The UPDPA and CAVACO statutes have certain exclusions from their coverage grounded in federal laws, while the UPDPA takes an unusual approach to other states' laws. The UPDPA takes a different approach to federal privacy laws than the CAVACO statutes. the UPDPA provides that a "controller or processor complies with [the Act] with regard to processing" if they are compliant with any of six federal statutes: HIPAA, Fair Credit Reporting Act (FCRA), Gramm-Leach-Bliley Act (GLBA), Driver's Privacy Protection Act (DPPA), FERPA, and COPPA, all of which we discussed above.[352] In the patchwork metaphor, The UPDPA is the blanket laid behind the patches that these federal laws represent. In this "two-ply" protection, if a controller or processor complies with the applicable federal law, it is also complying with the UPDPA. If it violates the federal law, it may also violate the UPDPA.[353]

The CAVACO statutes take different—dare we say "patchwork"?—approaches to the federal laws. California carves out several exceptions, some of them relating to controllers and processors, some relating to types of personal data, and some relating to particular data practices. It excludes controllers and processors that are "provider[s] of health care" and medical information subject to HIPAA;[354] it excludes personal data that are "collected, processed, sold, or disclosed" pursuant or subject to GLBA and DPPA;[355] and it excludes data practices governed by the FCRA.[356] Similarly, Virginia carves out entities and data subject to GLBA and HIPAA;[357] data subject to the DPPA, FERPA, and the Farm Credit Act;[358] and data practices subject to FCRA.[359] And Colorado

[348] *Id.* § 155(b).
[349] *Id.* §§ 155(b), 199.90(a).
[350] VCDPA, *supra* note 4, § 59.1-584(A), (C).
[351] CPA, *supra* note 4, § 6-1-1311(1).
[352] *See* UPDPA, *supra* note 4, § 11(a), (b). Virginia takes the same approach with COPPA.
[353] Subject to a pre-emption analysis.
[354] CCPA, *supra* note 4, § 145(a)(1-2).
[355] *Id.* § 145(e), (f).
[356] CCPA, *supra* note 4, § 145(d).
[357] VCDPA, *supra* note 4, § 59.1-576(B)(ii), (B)(iii), (C)(1)
[358] *Id.* § 59.1-576(C)(11)–(13).
[359] *Id.* § 59.1-576(C)(10).

excludes some healthcare information and data subject to HIPAA[360] and data subject to GLB, DPPA, COPPA, and FERPA;[361] data practices subject to FCRA;[362] and controllers subject to GLBA.[363] In the CAVACO states, these personal data, processors, and practices are simply not covered by their statutes: They rely entirely on the cited federal acts to govern these types of data practices, in contrast to the UPDPA in enacting states, which provides the two-ply protection mentioned above. Neither the UPDPA nor the CAVACO statutes give a pass to controllers and processors complying with privacy provisions of other federal laws not named here.

The UPDPA is different from the CAVACO statutes in another way: It is attentive to the laws of other states. The UPDPA expressly directs courts "applying and construing" the Act that they should "consider the promotion of uniformity of the law among jurisdictions that enact it."[364] The UPDPA also includes a bootstrap provision that allows a controller or processor to seek from the adopting state's attorney general (or designated privacy officer) a determination that complying with another jurisdiction's privacy law provides equal or greater protections than the adopting state's UPDPA.[365] Thus, a controller working in California and the adopting state might ask the attorney general in the adopting state to conclude that its compliance with the California Consumer Privacy Act of 2018 and California Privacy Rights Act of 2020 is sufficient to meet the requirements of the adopting state's implementation of the UPDPA.[366] The CAVACO statutes are silent on the laws of other states.

[360] CPA, *supra* note 4, § 6-1-1304(2)(I(e).

[361] *Id.* § 6-1-1304(2)(j).

[362] *Id.* § 6-1-1304(2)(i).

[363] *Id.* § 6-1-1304(2)(q).

[364] UPDPA, *supra* note 4, § 18. It also requires the attorney general (or other privacy officer) to "consider the need to promote predictability and uniformity among the states and give appropriate deference to a voluntary consensus standard developed . . . and recognized by a privacy-enforcement agency in another state," *id.* § 15(c), and to "consider the need to promote predictability for data subjects, controllers, and processors, and uniformity among the states" when considering adopting rules under the act, *id.* § 16(c).

[365] UPDPA, *supra* note 4, § 11(a).

[366] This is an arguable contention on the data controller or processor's part because, as this Article has shown, there are respects in which the CCPA does not cover personal data, regulated entities, or data practices quite the same way as UPDPA. The attorney general may set a fee for providing that this determination "reflect[s] the cost reasonably expected to be incurred . . . to determine" whether the other jurisdiction's law is good enough." *Id.* The UPDPA's drafters conclude that the attorney general would then be able to enforce the other jurisdiction's law against any controller or processor that had asserted another jurisdiction's privacy regime as a "substitute" for the adopting state's UPDPA. UPDPA, *supra* note 4, § 11, official comment ("Adoption of this act confers on the state attorney general, or other privacy data enforcement agency, authority not only to enforce the provisions of this act but also to enforce the provisions of any other privacy regime that a company asserts . . . as a substitute for compliance with this act.").

For public health researchers and professionals, the UPDPA's goal of uniformity is critically valuable. Though there are certainly public health projects based in single states, many research projects and interventions seek to operate across the country. If a state-by-state patchwork of non-uniform privacy laws supplements the substantive patchwork of federal privacy laws, public health researchers and professionals face the very real challenge of complying with an ever-larger number of regulatory regimes.[367]

III. EVALUATION AND INTERVENTIONS

We have so far provided a conceptual framework for data protection and analyzed how the enacted CAVACO statutes and the proposed UPDPA fit into that framework. This Part first briefly considers how these statutes relate to some of the normative assertions in the privacy-law literature.[368] It then evaluates how these statutes' provisions advance and impede public health work within our normative framework[369] and suggests ways that public health researchers and professionals should intervene to improve the situation in the coming months and years.

A. The UPDPA and the CAVACO Statutes vs. Normative Privacy Frames

As we noted above, the copious literature relating to data protection and privacy law in the United States casts a critical eye on the existing patchwork of laws. As a preliminary matter, we do not see evidence in the UPDPA and the Colorado and Virginia statutes that they have adopted the GDPR as their model, but neither do we see them adopting the California statute as a model, as Professors Chander, Kaminski, and McGeveran suggested they would. Among other things, Professors Chander, Kaminski, and McGeveran made much of the facts that the GDPR and California statutes differ greatly in length, with a "paperback of the GDPR run[ning] some 130 pages" and the CCPA being "around 25 pages";[370] that the CCPA "affords individuals little control" compared to the GDPR's "data protection" model;[371] that the CCPA does not provide private rights of action for individuals, while the GDPR did;[372] that the GDPR spelled out broad principles, while the CCPA provided much more specific enforcement mechanisms;[373] and that "the backdrop against which these

[367] *See supra* Section I.B.

[368] *See supra* Section I.C.

[369] *See Id.*

[370] Chander et al., *supra* note 75, at 1746.

[371] *Id.* at 1757.

[372] *Id.* at 1759.

[373] *Id.* at 1760.

two privacy laws were enacted, or . . . their legal setting, differs significantly," particularly as a result of First and Fourth Amendment jurisprudence.[374]

Taking at face value the differences that Professors Chander, Kaminski, and McGeveran identified between the GDPR and CCPA, the Colorado and Virginia statutes and the UPDPA appear to exhibit as much difference from the CCPA as CCPA does from the GDPR. Of course, all the American acts arose in a similar "legal setting." As for length, however, the California act (after the 2020 referendum amendments) weighs in at more than 24,000 words, while Virginia's is around 6,000 words, Colorado's is under 8,300, and the UPDPA comes in under 4,800.[375] We have noted[376] a considerable number of differences between California on the one hand and Virginia and Colorado on the other, including several places where Colorado's statutes followed Virginia's verbatim. Nevertheless, we have also noted that the UPDPA departs from approaches that the CAVACO states use, both some on which the CAVACO states agree and some on which they differ. As we also noted above,[377] California does provide a private right of action, though only for breaches of data security,[378] but Virginia and Colorado do not provide any private right of action at all.[379] The UPDPA, on the other hand, defers to the adopting state's consumer protection act, which the UPDPA cross-references for "enforcement authority, remedies, and penalties" under the act,[380] and which may or may not provide a private right of action.

Chander et al. concluded that "GDPR's vagueness is arguably deliberate," and that "EU authorities wanted to allow companies and sectors to fill in details of how to comply with the law over time, whether formally by establishing codes of conduct or certification mechanism . . . or informally through self-regulation . . . ."[381] Our description above[382] of the voluntary consensus standard that is integral to the UPDPA sounds more like the GDPR than the CCPA here, as VCSs allow for industry groups to build customized substantive and procedural regimes under the UPDPA that differ from each other.

In summary, we don't have space here fully to explore the question, but we expect that there is a new set of practical norms coalescing around discussions associated with the Virginia, Colorado, and Uniform Law Commission statutes,

[374] *Id.* at 1761.

[375] Indeed, even the difference between the CCPA and GDPR may not be as great as Chander et al. suggested, as the GDPR's operative provisions are under 31,000 words, with a considerable portion of its length consisting of more than 24,000 words of recitals.

[376] *Supra* Section II.E.

[377] *Supra* Section II.E.

[378] CCPA, *supra* note 4, § 1798.150(a).

[379] VCDPA, *supra* note 4, § 59.1-584(A), (C); CPA, *supra* note 4, § 6-1-1311(1).

[380] UPDPA, *supra* note 4, § 16(a).

[381] Chander et al., *supra* note 75, at 1760.

[382] *Supra* Section II.G.

as they were being developed at the same time in 2020 and 2021. In any event, the new practical normative model of the UPDPA, if it is new, clearly still embraces the "notice and choice" model already at the heart of the U.S. patchwork of sectoral privacy laws, as opposed to an information fiduciary model, for example. The CAVACO statutes and the UPDPA in some ways do exhibit the "follow the data" data-protection characteristic of GDPR that Chander et al. use when distinguishing it from the California model: The UPDPA regulates "data practices" and includes some as "prohibited," which cannot be consented to. In any event, these acts are thus unlikely to satisfy the expectations of scholars who are asking for more. Though, as we shall see, these acts set some defaults in a way that favors public goods—namely public health—some other defaults they set generally favor commercial uses of the kind that we found consumers comparatively disfavor.

Colorado and Virginia come closest to requiring opt-in, active consent for the data practices that consumers appear to disfavor.[383] As we noted above,[384] these statutes do not require consent for "collection of personal data to what is adequate, relevant, and reasonably necessary in relation to the purposes for which such data is processed, as disclosed to the consumer"[385] or processing for those purposes or for purposes "compatible" with them,[386] provided the data are not sensitive.[387] They require passive consent for certain uses, including "(i) targeted advertising, (ii) the sale of personal data, or (iii) profiling in furtherance of decisions that produce legal or similarly significant effects concerning the consumer."[388] But they require active consent, an opt-in, for almost all other data practices. The California and the UPDPA laws require active consent in the smallest number of cases: In California, only where the controller wants to enroll the data subject in "into a financial incentive program";[389] and under the UPDPA, only where sensitive data are concerned.[390] Given the default choices for consumers under these acts, they do little to address the concerns we identified above.[391]

These acts also do nothing to address the use of publicly available

[383] *Supra* Section I.C.3.

[384] *Supra* Section II.E.

[385] VCDPA, *supra* note 4, § 59.1-578(A)(1); *see also* CPA, *supra* note 4, § 6-1-1308(3) (using very similar language).

[386] VCDPA, *supra* note 4, § 59.1-578(A)(2): *see also* CPA, *supra* note 4, § 6-1-1308(4) (using very similar language).

[387] *See* VCDPA, *supra* note 4, § 59.1-578(A)(5).

[388] *Id.* § 59.1-577(A)(5); CPA, *supra* note 4, § 6-1-1306(1)(a)(i). *See also supra* Section II.E (discussing favored, restricted, and prohibited data practices).

[389] CCPA, *supra* note 4, § 1798.125(b)(3).

[390] UPDPA, *supra* note 4, § 8(c).

[391] *Supra* Section I.C.

information, which critics have noted can function to profile data subjects in ways they could not expect and to which they would likely not consent.[392] As for personal data that are covered, the UPDPA and California acts do provide some implied and express limitations on inferential data practices, which some have argued are not adequately addressed in current laws.[393] Neither the UPDPA nor the CAVACO statutes heed Professor Cohen's call for updated enforcement mechanisms. Finally, none of these acts overtly establishes an "information fiduciary" model, though they do take some steps to manage the information pipeline that begins with the collecting controller. For example, as we noted above,[394] each act requires the collecting controller to provide copies of data to subjects, to correct errors in the data, and to employ reasonable security measures; and the collecting controller is responsible for imposing those requirements on processors and third-party controllers downstream. The CAVACO acts, but not the UPDPA, also give the data subject a right to have data deleted.

Given our goal of addressing public health concerns under these statutes, we turn now to an evaluation of them from that perspective, providing recommendations for public health researchers and practitioners to intervene.

B. Helping and Hindering Public Health Activities

This Section considers whether the UPDPA and CAVACO statutes help or hinder public health activities. After giving a brief overview, it examines the real and perceived barriers that data privacy laws can create and then examines the effects of these statues on data practices for research and on public health practices. As a preliminary matter, public health researchers and professionals must claim a seat at the table during deliberations on comprehensive data privacy or protection statutes, whether at the state or federal level. Legislators in general are not experts in public health and are not well situated to evaluate the effects of legislative proposals on public health. Other private and public interest groups are generally very skilled at advancing their objectives with legislatures, but those objectives may not fully support public health activities. Interventions by

[392] *See supra* Section I.C.2.

[393] *See Id.* CCPA covers includes in covered personal data any "[i]nferences drawn from . . . [personal information] to create a profile about a consumer reflecting the consumer's preferences, characteristics, psychological trends, predispositions, behavior, attitudes, intelligence, abilities, and aptitudes." CCPA, *supra* note 4, § 1798.140(o)(1)(K). The UPDPA may attempt to address this by limiting the use of personal data "to make a prediction or determination about a particular data subject," making it one of the factors used to determine whether a data practice is a favored (i.e., compatible) data practice. Nevertheless, this UPDPA provision likely does not go as far as Professor Solow-Niederman might like, as the UPDPA applies to only to identifiable and pseudonymized data.

[394] *Supra* Section II.F.

public health researchers and professionals matter. For example, in June 2021, the authors wrote a letter to the ULC committee developing the UPDPA—effectively in the eleventh hour of the committee's work, as it planned to introduce the final UPDPA to the full Commission in July—urging changes to support public health.[395] The committee made some of those changes, and the committee's reporter credited the letter for prompting them.[396]

From a normative perspective, transparency and autonomy for data subjects are probably well-protected under all four statutes for IRB-approved research where the data are collected for the primary purpose of research, as such research protocols typically require voluntary participation and consent or similar protections. The UPDPA and Virginia acts cover data in research that makes secondary use of data and in some public health practices.[397] From the data subject's perspective, this may be desirable, but it may create impediments to public health practice and research by bringing them within the purview of the acts. The California and Colorado acts exempt the greatest swaths of data, diminishing to some extent the data subjects' autonomy but removing barriers to public health research that makes secondary use of data and to public health practice. Exempting public health practice and research from the coverage of the UPDPA and the CAVACO statutes is only one way the acts might encourage public health, as we discuss below.

A legislator or lobby proposing legislation for data protection or privacy will most likely model it on one of the existing acts, the UPDPA or one of the CAVACO statutes. In that event, we have specific recommendations for changes, based on our normative model. Table 1 summarizes key characteristics of the UPDPA and CAVACO statutes as they affect public health practices and research; the entries in it that are highlighted in ***bold italic text*** are those that raise concerns according to our normative frameworks.

1. Real and Perceived Barriers to Data Use for Public Health Practice and Research

Evaluating the impact of a data protection law on secondary data use requires acknowledging that both real and perceived data-use barriers exist. Data protection laws impose real barriers when the text of the laws prohibits or impedes (i.e., through complicated requirements or procedures) the use of data.

[395] Letter from Cason Schmit et al., Faculty, Texas A&M University, to Harvey Perlman, Chair, Drafting Committee, Collection and Use of Personally Identifiable Data Act, Uniform Law Commission (June 6, 2021) (on file with authors).

[396] Letter from Jane Bambauer, Professor of Law, University of Arizona, to Cason Schmit et al., Faculty, Texas A&M University (July 6, 2021) (on file with authors).

[397] Schmit et al., *supra* note 28, 83–86.

For example, the UPDPA creates real data-sharing barriers for prohibited data practices because the law expressly prohibits those activities.[398] Similarly, although FERPA technically does permit some public health uses of education data by permitting the use of aggregate data or the use of personal data with the express consent of all individuals,[399] the utility deficiencies of aggregate data and the practical difficulties associated with consent in big-data applications effectively mean that FERPA poses real data sharing barriers to public health data practices.[400]

Perceived data-sharing barriers are different because the language of the law does not actually create a real barrier to secondary data practices. Instead, barriers exist when controllers or processors believe a barrier does, or could, exist. These perceived barriers are most likely to exist when data protection laws are complex, lack specific language, or carry substantial penalties that encourage hyper-conservative organizational practices. For example, HIPAA is often cited as a data-sharing barrier when, in fact, it contains generous provisions permitting research and public health activities.[401]

The vague definitions of protected data in these acts could also introduce perceived barriers. The CAVACO statutes and the UPDPA all use reasonableness to define protected data, which creates uncertainty for data controllers that wish to share data for public health practice or research. With this uncertainly, controllers will likely consider legal deidentification exceptionally difficult to practically accomplish without a clear safe-harbor exception (i.e., like the HIPAA regulations).

2. *Data Uses for Research*

Provisions in data protection laws that permit data gathering where the primary use is human subjects research regarding public health are beneficial to

[398] UPDPA, *supra* note 4, § 9.

[399] 34 C.F.R. § 99.30 (2021).

[400] ASS'N OF STATE & TERRITORIAL HEALTH OFFS., PUBLIC HEALTH AND SCHOOLS TOOLKIT, FAMILY EDUCATIONAL RIGHTS AND PRIVACY ACT, https://www.astho.org/advocacy/state-health-policy/legal-preparedness-series/public-health-and-schools-toolkit/? [https://perma.cc/R5PM-WDPH].

[401] 45 CFR 164.512 (b), (i) (2021); Steve Alder, *Do HIPAA Rules Create Barriers That Prevent Information Sharing?*, HIPAA JOURNAL (Nov. 19, 2018), **Error! Hyperlink reference not valid.**https://www.hipaajournal.com/hipaa-rules-barriers-to-information-sharing/ [https://perma.cc/F4ZK-BYVG]; *see also* 21st Century Cures Act, Pub. L. No. 114–255, 130 Stat. 1033 (2016), where Congress made "information blocking" illegal for certain health data applications to address restrictive organizational and technological practices that interfere with legitimate data sharing. Had there been real legal barriers, Congress likely would have needed to create or expand HIPAA data use provisions.

the public.[402] Data gathered particularly for public health research, including health records, environmental conditions, and consumer behavior data, can help public health professionals understand the causes of poor health and investigate interventions that promote well-being.

All four statutes provide some protections for transparaency and autonomy in research contexts. But there are also some impediments research makes secondary uses of data.[403] As Table 1 shows, the UPDPA and CAVACO statutes exempt from their application data gathered for public health research according to contemporary ethical principles. These provisions are beneficial from a public health perspective because they do not add additional requirements on top of the existing regulatory framework established by the already expansive federal Common Rule.[404]

Where human subject research relies on secondary data, however, there are some variations among these acts. As the second row of Table 1 shows, the UPDPA and the California and Colorado statutes generally permit such uses. The Virginia act, however, requires active consent before a data controller discloses data for the secondary purpose of research. At a minimum, public health researchers and professionals should seek to have research that is subject to the Common Rule classified as disclosed favored data practices or as passive-consent restricted data practices. The default on consent here is critical to ensuring that data a controller provides to researchers is representative. Of course, researchers will also have to satisfy IRBs that they are taking appropriate steps to protect data subjects from harms associated with research. For states considering the UPDPA, they should propose that Common Rule research be a "compatible data practice" under the Act. If a state data privacy act entirely exempts research from its application, controllers could in theory provide data to researchers without disclosing the fact to data subjects at all; and that would prevent data subjects having the right to opt out, either of the data practice or of a relationship with the controller altogether. Given that active consent is a poor default where obtaining consent for research is required, we urge public health researchers and professionals to oppose such requirements in acts in other states, and we suggest those in Virginia may want to seek an amendment to the Virginia act to correct this default.

[402] *See generally* Ramanathan et al., *supra* note 269.

[403] Public health practices and data disclosures entirely within and among government agencies are not covered. *See supra* Section II.D.

[404] Common Rule, *supra* note 9.

Table 1: Status of data practices relevant for public health under each act (matters of concern for public health in *highlighted text*). "Favored" means that data may be used for the purpose without consent; "restricted" that data may be used only with active consent (opt-in) or with passive consent (chance to opt-out).

Human subjects research (HSR)[405]				
	UPDPA	**California**	**Virginia**	**Colorado**
—HSR is primary use	Act does not cover activity if data are *collected solely* for HSR	Act does not cover activity if data are *collected* for HSR	Act does not cover activity if data are *collected* for HSR	Act does not cover activity if data are *collected* for HSR
—HSR is secondary use	Act favors activity: no consent required but must be disclosed if "routine"	Act does not cover activity if data are *disclosed* for HSR	***Act restricts activity: permitted only with active consent (opt-in required)***	Act does not cover activity if data are *used or shared* for HSR
Other public health activities				
	UPDPA	**California**	**Virginia**	**Colorado**
Public health surveillance	Act favors activity: no consent required but must be disclosed if "routine"	Act restricts activity: permitted with passive consent (opt-out offered).	***Generally, act restricts activity: permitted only with active consent (opt-in required)*** (Exception: If HIPAA permits the activity by covered entities for public health and public health is the data's primary use, Virginia act does not cover it.)	Act favors activity: no consent required; ***no disclosure required*** (subject to certain conditions)
Public health population interventions				
Public health individual interventions	Act restricts activity: permitted with ***active consent (opt-in required) for sensitive data***; passive consent for all others			

[405] Subject to IRB/Common Rule, *supra* note 9.

3. *Data Uses for Public Health Practice*

Provisions in data protection laws that permit secondary use of data for public health practices are also beneficial to the public. Many factors beyond biology, including social, environmental, and economic factors, determine an individual's health status.[406] Traditional public health data sources consist mainly of health records and surveillance data, such as reports of infectious diseases, but the myriad of data protection laws have created both real and perceived barriers to access data on many social, environmental, and economic factors.[407] These data are essential to fully leverage data to promote population well-being.[408] Moreover, research data-use exemptions are often not sufficient for public health activities that require swift action, such as surveillance for outbreak investigations.

As Table 1 shows, the California and Colorado acts broadly support data practices, primary and secondary, for all three categories of public health activity: surveillance, population interventions, and individual interventions. California restricts these activities, requiring notice and choice, but the choice is via passive consent and thus opt-out. The Colorado statute provides broad permission for data practices for "reasons of public interest in the area of public health."[409] In Colorado, these activities are favored, requiring no consent or disclosure to the data subject, provided those performing the activities meet the statute's requirements. Though supportive of public health, these provisions raise concerns on normative grounds that they deny data subjects transparency and autonomy. As a normative matter, we would prefer to see disclosure, which would allow data subjects either to opt out of the data practice or choose not to disclose data to the collecting controller in the first place. Public health professionals in Colorado might seek a revision to that act to address this concern.

The Virginia statute may have grave effects on the use of personal data for public health practices, and the UPDPA may have such effects on the use of sensitive personal data for public health practices. The Virginia act provides significant impediments to all public health activities, as it permits them generally only with active consent, requiring notice and opt in. (There is an exception for HIPAA-covered entities using data for the primary purpose of public health, but this is a narrow category.) The UPDPA favors data practices,

[406] *See* Frieden, *supra* note 38; Galea et al., *supra* note 14.
[407] Schmit et al., *supra* note 28, at 83–86; Braveman & Gottlieb, *supra* note 14, at 19–31.
[408] Kum et al., *supra* note 217.
[409] CPA, *supra* note 4, § 6-1-1304(3)(a)(xi) (2021).

primary and secondary, for public health surveillance and population interventions, requiring no consent but disclosure to the data subject. The UPDPA requires active consent, however, for individualized interventions involving sensitive data. When the default is to require active consent from the data subject—an opt in—subjects are much less likely to agree to participate, likely leaving public health efforts with spotty data that may be severely skewed based on which data subjects do decide to opt-in. As sensitive data under the UPDPA include sex, gender, etc., this problem may be particularly acute in adopting states. Though these provisions value personal autonomy, they do so at considerable danger to public health. Public health professionals in Virginia should seek to modify its act to align it more closely with the other CAVACO statutes and the UPDPA. They should also seek to modify the requirement for active consent for public health uses of sensitive data so that they require only passive consent.

Public health professionals may also seek a voluntary consensus standard[410] to clarify that such interventions *for public health* are indeed compatible data practices that do not require complicated consent. For example, one of the factors that can be weighed when determining whether an activity is a compatible data practice is whether the activity advances "the economic, health, or other interests of the data subject."[411] Given that this is the goal of many public health interventions, it is possible that many public health activities—even individual interventions—could be permitted under the UPDPA's factor-based definition for compatible data practices. On the other hand, the UPDPA's flexible, factor-based approach to compatible data practices creates substantial uncertainty about public health interventions that target specific individuals because of the absence of express permissive language. There are thus opportunities to improve or clarify the UPDPA rules to maximize data practices to promote population health. Some public health data practices, particularly if they result in individualized interventions and involve sensitive data, could be seen as restricted practices that require active consent.

Public health professionals would thus be wise to seek provisions in a VCS for practices that they want to be classified as favored. Such a VCS would greatly facilitate the work of public health professionals and researchers. But development of a VCS requires a critical mass of experts from the field, representatives of consumer groups, and others. It will take time and money. On the bright side, because the UPDPA calls for states to respect each other's judgments when approving VCSs, public health professionals need not create a VCS for only one state. Rather, they can collaborate to develop a national

[410] *See supra* Section II.G.

[411] UPDPA, *supra* note 4, § 7(a)(6).

standard with a hope that most or all the UPDPA states will adopt it and that the CAVACO states and others modeling their legislation on CAVACO statutes would amend their acts to come into conformity with the VCS.

Of course, the next step after developing a VCS is getting it accepted in the UPDPA states. We suggest that public health professionals focus their efforts on states with larger populations whose adoption will function to influence attorneys general more strongly in other states to accept it. A strategic effort to seek early adoption of the VCS in states with diverse political climates (e.g., some strongly Democratic and some strongly Republican) may also make it easier to obtain wider adoption by avoiding any apparent taint of partisanship.

The work of public health professionals is not over when the statutes and VCSs are adopted. Key for public health professionals in California and the UPDPA states is making sure that collecting controllers disclose proposed public health uses of data. They need to persuade private-sector controllers who may be their partners to provide notice in their privacy policies indicating they are engaging in these activities. This is probably not a burdensome requirement where private controllers are concerned, as the public health researchers and professionals must generally form relationships with them to obtain data anyway. In California and Virginia, collecting controllers that partner with public health researchers and practitioners may need to add the means for consumers to opt out or in for various proposed data practices. Similarly, if the UPDPA is adopted without modification from ULC's model, uses of sensitive data in individual public health interventions will also require an opt-in mechanism. Public health researchers and professionals in those jurisdictions would have to work with private-sector partners to provide disclosure and probably some kind of incentive to for data subjects to opt in.

In the UPDPA states where a VCS is adopted, controllers and processors will still need to indicate that they are complying with the VCS in their privacy policies so that public health uses can be considered "compatible" data practices under the UPDPA. In states that model their statutes on the CAVACO acts, further work may be necessary to ensure that private-sector controllers and processors can comply with requests from public health researchers and professionals to work with them.

CONCLUSION

Ideally, data privacy laws create restrictions to protect against risky or harmful data practices while permitting socially desirable data practices. Governmental and public interest in new privacy regulations is a reaction against the existing U.S. privacy approach to this balance. To a great extent, the advent of the UPDPA and CAVACO statutes may help to create the blanket of data privacy protection many have called for in recent years. For the most part, they appear to cover those areas left uncovered by the long-standing patchwork of data privacy protections. However, allowances for socially beneficial data uses in new privacy regulations are just as critical. There are great opportunities in these laws to extend and support public health practices and research under these blankets. Public health professionals should be alert to legislative and regulatory efforts, however, and engage with them to prevent restrictions that prevent public health work for the public good.