This Article compares the material scope of several comprehensive consumer data privacy (or data protection) laws enacted recently in the United States, both with each other and with the European Union’s General Data Protection Regulation (“GDPR”). Our comparative analysis
covers five broad state consumer data privacy laws enacted and in effect as of the end of 2023, specifically those adopted in California, Virginia, Colorado, Utah, and Connecticut. We contrast these against each other and the GDPR. We compare how each of these laws define and scope their subject matter (e.g., what constitutes “personal data”), how they define data subjects, what amounts to data processing, and which entities are obligated to respect the data subjects’ rights provided by these laws. We demonstrate how the existing state laws are more limited in most respects than the GDPR, and how their framing as consumer protection laws significantly limits their applicability and restricts their ability to adequately address the broad range of data privacy problems that confront contemporary society. Drawing on neorepublican political philosophy, we argue that most of these laws generally fail to adequately constrain commercial data markets in many contexts and that they also fail to address the problem of law enforcement agencies acquiring personal information from the commercial sector—ultimately raising concerns about domination and the potential for uncontrolled interference by both corporate and state interests in the private lives of the data subjects who ostensibly acquire rights. In the end, most of these “comprehensive” consumer data privacy laws at the state level in the United States do little to reign in corporate and state power to collect and use personal data in many contexts and represent a missed opportunity to provide much more significant protections for individual data privacy rights in the United States.
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Introduction</strong></td>
<td>1058</td>
</tr>
<tr>
<td>I.</td>
<td>Republicism and Theories of Privacy: A Brief Overview</td>
<td>1065</td>
</tr>
<tr>
<td></td>
<td>a. Data, Privacy, and Uncontrolled Interference</td>
<td>1066</td>
</tr>
<tr>
<td>II.</td>
<td>Methodology and Case Selection</td>
<td>1072</td>
</tr>
<tr>
<td>III.</td>
<td>Comparing the Material Scope of U.S. and EU Data Privacy Laws</td>
<td>1075</td>
</tr>
<tr>
<td></td>
<td>a. Subject Matter</td>
<td>1078</td>
</tr>
<tr>
<td></td>
<td>i. Inputs</td>
<td>1084</td>
</tr>
<tr>
<td></td>
<td>ii. Links</td>
<td>1089</td>
</tr>
<tr>
<td></td>
<td>iii. Identification</td>
<td>1093</td>
</tr>
<tr>
<td></td>
<td>1. The “Identified” and/or “Identifiable” Data Subject</td>
<td>1093</td>
</tr>
<tr>
<td></td>
<td>2. De-Identified Data</td>
<td>1098</td>
</tr>
<tr>
<td></td>
<td>iv. Subjects</td>
<td>1103</td>
</tr>
<tr>
<td></td>
<td>v. Additional Exceptions</td>
<td>1106</td>
</tr>
<tr>
<td></td>
<td>1. Publicly Available Information</td>
<td>1106</td>
</tr>
<tr>
<td></td>
<td>2. Aggregated Consumer Data</td>
<td>1111</td>
</tr>
<tr>
<td></td>
<td>3. Other Exemptions</td>
<td>1112</td>
</tr>
<tr>
<td></td>
<td>b. Processing: Regulated Use of Data</td>
<td>1113</td>
</tr>
<tr>
<td></td>
<td>c. Covered Entities: Controllers and Processors</td>
<td>1118</td>
</tr>
<tr>
<td>IV.</td>
<td>Comparing the Law in Various Contexts</td>
<td>1125</td>
</tr>
<tr>
<td></td>
<td>a. Uber and Its Drivers: Automated Employment Decisions</td>
<td>1126</td>
</tr>
<tr>
<td></td>
<td>b. Paint Chips as Personal Data: The Scope of Personal Data</td>
<td>1127</td>
</tr>
<tr>
<td></td>
<td>c. Posting on a Website: Individuals Posting Personal Data on the Internet</td>
<td>1130</td>
</tr>
<tr>
<td></td>
<td>d. Post-Roe Surveillance After Dobbs</td>
<td>1132</td>
</tr>
<tr>
<td>V.</td>
<td>Conclusion</td>
<td>1134</td>
</tr>
</tbody>
</table>
INTRODUCTION

Information privacy is having a moment in the United States. Only a few decades ago, data was described as “the sludge of the information age—stuff that no one has yet thought very much about.” Now, scholars refer to the widespread commercial practice of “trafficking in human information.” Clearly, much has changed in the intervening years, yet only recently have any U.S. jurisdictions adopted broad—what might be called non-sectoral, comprehensive, or omnibus—data protection (or data privacy) laws. Recently, there has been a resurgence of legislative interest in data protection or data privacy laws in the United States. This “new wave” of data privacy laws began with enactment of the California Consumer Privacy Act in 2018 and subsequent revisions, including those promulgated in the California Privacy Rights Act of 2020 (referred to collectively hereinafter as “CCPA”). Subsequently, Virginia and Colorado enacted “comprehensive” consumer privacy laws in 2021, followed by Utah and Connecticut in 2022. In 2023, Delaware, Indiana, Iowa, Montana, Oregon, Tennessee, and Texas also enacted similar laws, followed by New Jersey in early 2024. The text of these laws often tracks language proposed, but not adopted, in Washington State over several previous legislative sessions. Legislatures around the United States, in state houses and houses of Congress, have considered adopting more comprehensive consumer privacy and data protection laws in the last few years. However, Congress has failed to order anything from the plentiful

5 Anupam Chander et al., Catalyzing Privacy Law, 105 U. Minn. L. Rev. 1733, 1734 (2021).
menu before it, although significant bipartisan steps have been taken, leaving U.S. data privacy law in the hands of the few state governments that have moved to adopt such a law.

The importance of data privacy has also received additional attention after the Supreme Court decided *Dobbs v. Jackson Women’s Health Organization* on June 24, 2022. In that case, the Court overruled *Roe v. Wade* and *Planned Parenthood v. Casey*, holding that the U.S. Constitution does not provide a federal right (of decisional privacy) to choose to abort a pregnancy. Importantly, for purposes of privacy law, Justice Alito’s majority decision sparked concerns amongst abortion rights advocates that commercial data collection and surveillance practices would be harnessed by law enforcement to investigate individuals seeking information about and access to abortions in states where abortion is now illegal or at

---


least highly regulated. Of course, privacy and data protection concerns are broad and varied, but this post-Roe America fervor has only put increased pressure on legislatures to act in meaningful ways to protect individual privacy—from both private commercial actors as well as from law enforcement. Indeed, the need to do something about data privacy is only underscored by the fact that, “[u]ntil the CCPA, most American law permitted [non-governmental] entities to collect and use personal data however they wished by default, absent a specific legal rule forbidding a particular practice.” According to Determann, this default presumption to allow data collection and use absent some specific legal prohibition is one of the hallmarks that set data (or information) privacy law apart from data protection law. The second defining characteristic of data protection law is that the focus is on “protecting information concerning persons” rather than persons themselves (personal data essentially operates as a legal proxy where the aim is to protect persons but the focus is on personal data).

Although many countries around the world have enacted broad data protection laws, the most frequent reference is to the European Union (“EU”)’s General Data Protection Regulation (“GDPR”). But the GDPR—which generally applies to personal data processing for non-law enforcement purposes—is not the only

---


12 Chander et al., supra note 5, at 1748.


14 Id. at xv.


16 See Greenleaf & Cottier, supra note 15, at 29 (noting that many scholars have linked the rise of privacy laws around the world to the GDPR); General Data Protection Regulation 2016/679, 2016 O.J. (L 119) 1 [hereinafter GDPR].
important data protection tool in the EU. The GDPR was adopted alongside Directive 2016/680, commonly referred to as the Law Enforcement Directive or “LED.” The Law Enforcement Directive specifically regulates public law enforcement-related processing of personal data, applying when personal data is collected or processed “by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the safeguarding against and the prevention of threats to public security . . . .” Additionally, Article 8 of the Charter of Fundamental Rights of the European Union (“EU Charter”) makes the protection of personal data a fundamental rights concern within the EU. These data protection rights also exist against the backdrop of a broad fundamental right to a private and family life, home, and communications, as enshrined in Article 7 of the EU Charter and Article 8 of the European Convention on Human Rights (“ECHR”).

In the United States, no comparable fundamental rights to privacy, federal data protection, or law enforcement-directed data protection law exist. The state-level data protection laws analyzed


20 See Charter of Fundamental Rights of the European Union 326/02, art. 8, 2012 O.J. (C 326) 391, 397 (Regarding “[p]rotection of personal data,” the article provides, in part, that “[e]veryone has the right to the protection of personal data concerning him or her.”).

21 See id. (“Everyone has the right to respect for his or her private and family life, home and communications”).

22 Convention for the Protection of Human Rights and Fundamental Freedoms, art. 8, (Nov. 4, 1950) (the first provision of which reads, “[e]veryone has the right to respect for his private and family life, his home and his correspondence”).

23 See Dobbs, 597 U.S. at 215. Even though the preamble to the Colorado Privacy Act proclaims that “[t]he people of Colorado regard their privacy as a fundamental right and an essential element of their individual freedom,” the law itself does not treat privacy as a fundamental right but as a consumer protection issue.
in this Article represent the closest things to the GDPR—or to broad data protection law generally—in the United States at present.\textsuperscript{24} Information gathering practices of law enforcement are primarily regulated by state and federal search and seizure laws, including the Fourth Amendment to the U.S. Constitution\textsuperscript{25} and other constitutional provisions, as well as communication privacy laws such as the Stored Communications Act,\textsuperscript{26} but those provisions are simply not comparable to the LED in purpose or scope, and have failed to limit warrantless law enforcement access to commercial databases in several high-profile contexts.\textsuperscript{27} As such, members of Congress have introduced bills such as the proposed Fourth Amendment is Not For Sale Act,\textsuperscript{28} which would require law enforcement to obtain a warrant prior to accessing certain personal data in commercial databases.

Conventional wisdom suggests that the growing interest in omnibus data protection laws in the U.S. has been driven by adoption of the GDPR in the EU.\textsuperscript{29} An alternative theory of catalysis posits that American data privacy laws differ significantly from the GDPR and are more likely inspired by the CCPA adopted in California in 2018, just a month after the GDPR went into effect in Europe.\textsuperscript{30} There is also a semantic difference in basic terminology, as American commentators often refer to these new laws as data privacy laws instead of data protection laws.\textsuperscript{31} However, to the

\textsuperscript{24} See Chander et al., supra note 5, at 1749-50 (arguing that the CCPA “has some data protection characteristics,” even though it is primarily a consumer protection-based “data privacy” law).

\textsuperscript{25} U.S. CONST. amend. IV.

\textsuperscript{26} 18 U.S.C. § 2701 et seq.


\textsuperscript{28} H.R. 2738/S.1265, 117th Cong. (2021).

\textsuperscript{29} See, e.g., Paul M. Schwartz, Global Data Privacy: The EU Way, 94 N.Y.U. L. REV. 771, 818 (2019); Chander et al., supra note 5, at 1735-36.

\textsuperscript{30} Chander et al., supra note 5, at 1733-34.

\textsuperscript{31} See id. at 1747 (noting “a fundamental difference in approach,” in which “‘data protection’ is universal in Europe, while most American law focuses on ‘consumer protection’ and ‘privacy’”). Chander et al. also make the case for why a
extent they focus on data rather than personal privacy and set baseline, default prohibitions on data processing, these laws resemble data protection laws, albeit limited to a consumer protection framing that is narrower than the fundamental rights foundation of the GDPR. Despite resembling data protection law, Chander, Kaminski, and McGeveran, argue that the CCPA, “differs significantly—and consciously—from the European model . . . [offering] a fundamentally different regime for data privacy,”32 one that lacks “major structural elements of the GDPR . . . .”33 Yet again, while either might have been the impetus, the current swath of laws being adopted and proposed in the last few years often appears to follow what has been called the “Virginia Model”34—copying and replicating many aspects of Virginia’s Consumer Data Protection Act of 2021 (“VaCDPA”) which was based on legislative language proposed previously in Washington State—even if their impetus was the CCPA.35 There is also evidence that corporate interests have been lobbying these states to create business-friendly, uniform laws that minimize variation across jurisdictions, including such lobbying in Virginia in support of the VaCDPA.36

Minimal variation is, of course, better for business. However, such uniformity at the state level would gut the possibility for legislative experimentation in the various states, and with high levels of corporate involvement and lobbying at the state level, the likely outcome is watered-down, business-friendly legislation that

32 Id. at 1736 (alteration in original).
33 Id. at 1746.
35 See Chander et al., supra note 5, at 1764 (proposing that states copied the CCPA when drafting proposed data privacy legislation).
lacks much real innovation in establishing strong data protection rules in ways that promote privacy interests. In that vein, Utah’s law, one of the latest enacted at the time of writing, has also been described as taking “a lighter, more business-friendly approach to consumer privacy” than other existing data protection laws, and it appears the trend may be moving away from more privacy-focused laws like the GDPR and CCPA. Additionally, at the federal level, movement in Congress toward potentially adopting the American Data Privacy and Protection Act (“ADPPA”) recent legislative sessions became enmeshed in debates about federal preemption of state data privacy, especially since the ADPPA was seen by some as less robust than the CCPA in California.

In this Article, we examine and compare the material scope of these new American data privacy laws. Understanding the material scope of the law is important, as it forms the foundation upon which the rights and obligations rest. We compare the laws in the first five states to have enacted broad data protection laws—each of which have also gone into effect as of the end of 2023—with each other and with the GDPR. Specifically, we look at how these laws define and conceptualize their subject matter, focusing on the key concepts of 1) personal data and similar terms that determine what is subject to regulation; 2) “persons,” “consumers,” and other terms that determine to whom and in which contexts the law provides rights (data subjects); 3) which forms of “data processing” are subject to regulation; and 4) which entities acquire obligations (and potential liability) arising from such rights. Importantly, we do not examine the rights enshrined in these laws themselves, or the full range of


40 Our analysis includes the California Consumer Privacy Act (as amended by the California Privacy Rights Act of 2020), Virginia Consumer Data Protection Act, Colorado Privacy Act, Utah Consumer Privacy Act, and Connecticut Data Privacy Act.
other provisions contained in the statutes, although future research ought to compare those provisions as well.

Examining the affordances that these new information privacy laws provide to commercial data controllers, including to share data with state institutions, highlights how certain legislative choices allow the continued consolidation of informational power within corporations and state organs, resulting in the potential for domination and the loss of individual and collective freedom. Drawing from neorepublican political philosophy, we analyze how well these laws protect user (consumer, data subject) privacy and ensure some measure of what Philip Pettit calls “antipower” — that is, the power to resist the possibility of arbitrary or uncontrolled interference by others. This analysis is informed by Julie Cohen’s notion of “semantic discontinuity,” and we show how developing more robust laws that regulate data practices with the aim of preserving or enforcing interstitial complexity can promote antipower and reduce possibility of informatic domination. This, in turn, can better protect privacy as a fundamentally important right, instrumentally linked to personal and collective freedom. We question whether the material scope of these laws adequately captures and protects the underlying interests that appear to have motivated their adoption, such as privacy and the need to protect people from other data-driven harms. We also briefly note whether, and to what extent, these laws regulate or address the issue of law enforcement access to commercial databases. Finally, we examine to what extent the material scope of these laws—including how they protect privacy interests and limit corporate power—contributes to promoting neorepublican notions of liberty, non-domination, and antipower.

I. NEOREPUBLICANISM AND THEORIES OF PRIVACY: A BRIEF OVERVIEW

In this Part, we provide a brief overview of the major theoretical constructs that we use later on to reflect on the comparative findings

44 See our definition and discussion of this term, infra, Part VI.
of this research—namely, civic or neorepublican conceptions of liberty and domination, and the application of those ideas to the realm of privacy and data protection. At the outset, it should be noted that we have intentionally kept this analysis separate from our primary comparative analysis of these laws, so as to maintain the ability for the comparison to stand on its own, regardless of whether readers are sympathetic to the particular political philosophy employed or the normative conclusions we draw from it. Thus, even if one rejects the tenants of the neorepublican position we take on privacy and its application to data privacy law, we argue the comparative analysis will still be useful.

a. Data, Privacy, and Uncontrolled Interference

One of the primary faults of data privacy law (especially in the American context) is the entrenched dogma of “notice and choice” based on an intertwined overreliance on notions of individual

45 See, e.g., PHILIP PETTIT, JUST FREEDOM: A MORAL COMPASS FOR A COMPLEX WORLD (W.W. Norton & Co. 2014) [hereinafter PETTIT, JUST FREEDOM]; PHILIP PETTIT, REPUBLICANISM: A THEORY OF FREEDOM AND GOVERNMENT (David Miller & Alan Ryan eds., 1997); PETTIT, supra note 41; PETTIT, supra note 42; Samantha Besson & José Luis Martí, LAW AND REPUBLICANISM: MAPPING THE ISSUES, in LEGAL REPUBLICANISM: NATIONAL AND INTERNATIONAL PERSPECTIVES 3 (Samantha Besson & José Luis Martí eds., 2009).

control and individual consent, especially given its practical application, for decades, within a neoliberal framework favoring self-regulation by commercial actors.

The integration of these concepts into data privacy law originates, at least in part, from the development of Fair Information Practices ("FIPs"), a "set of aspirational principles developed over the past fifty years used to model rules for responsible data practices." That is not to say that notice, consent, and control are not important or even useful to protecting privacy, but their application has often proved insufficient to providing rigorous privacy protections. To that point,

a legal model grounded in ‘notice and choice’ cannot prevent data-based manipulation when notice is fictional, when choice can be manufactured by the tools of data and behavioral science, and when rules for individuals are used to regulate a problem with social dimensions.

As Hartzog argued, “while the FIPs have been remarkably useful, they have painted us into a corner.” This corner is

---

47 See Woodrow Hartzog, Privacy’s Blueprint: The Battle to Control the Design of New Technologies 95 (Harvard Univ. Press 2018) (“Privacy law’s first critical mistake was to overleverage the notion of control. By doing so, design is allowed to take functional control away from people by overwhelming them with structured choices.”).

48 See Neil Richards & Woodrow Hartzog, The Pathologies of Digital Consent, 96 Wash. U. L. Rev. 1461, 1463 (2019) ("[W]hile consent models permeate the digital consumer landscape, the practical conditions of these agreements fall far short of the gold standard.").

49 See Chander et al., supra note 5, at 1747-48 (recognizing “the much-criticized premise that disclosure and a right of refusal (so-called ‘notice and choice’) adequately empower consumers”); Hartzog, supra note 4747, at 62-67; Cohen, What Privacy Is For, supra note 4343, at 1930 (arguing that American privacy law suffers from “the deeply neoliberal philosophy that animates the new governance” and "that characterize[s] informational capitalism,” including “[i]ts emphasis on privatized regulation and control of information flows via notice and choice”); Cason Schmit et al., Data Privacy in the Time of Plague, 21 Yale J. Health Pol’y L. & Ethics 152, 176 (2022) (“While consumers often demand notice and choice rights, a growing body of literature suggests that the sense of control they provide may be illusory.”); Alicia Solow-Niederman, Information Privacy and the Inference Economy, 117 NW. U. L. Rev. 357, 364 (2022) (“[I]ndividual 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.”).


51 Richards & Hartzog, supra note 2, at 967.

52 Id. at 953.
characterized by the shift from “the first wave of personal computing”53 to “the age of Big Data” in which this exclusive, neoliberal, and legalistic focus on individual rights and interests (including individual control) is proving problematic “because Big Data does not revolve around individuals.”54 Additionally, as we argue below, the long-popular notion that privacy ought to be something like “the right to be let alone”55 is also inadequate to address today’s privacy concerns when we consider the networked nature of information flows and the aggregation of data from multiple sources that powers much of the commercial data market.

Against this backdrop, civic or neorepublican theories of privacy focus on privacy as instrumental to achieving the “central ideal of modern republican theory—freedom as non-domination.”56 According to Roberts, “the value of privacy for republicans lies in its capacity to shield individuals from the threat of domination. A consequence of loss of one’s privacy is that others may acquire dominating power—the capacity to interfere in one’s decisions on an arbitrary basis.”57

Broadly, neorepublican political philosophy takes aim at “the evil of subjection to another’s will,”58 and is particularly concerned with subjection within “important areas of personal choice,” framing the problem as one in which the state can play an important and instrumental role in preserving individual and collective liberty.59 The republican approach offers an alternative to current, failing approaches to protecting privacy based in neoliberal, market-favoring ideas like notice, choice, and (mostly non-informed) consent. Much of the recent work in this area is built on the work of Philip Pettit,60 but there is also a growing literature of aligned scholarship on privacy that does not directly rely on republican theory. For example, Julie Cohen’s concept of “semantic

53 Id.
54 Bart van der Sloot, The Individual in the Big Data Era: Moving Towards an Agent-Based Privacy Paradigm, in EXPLORING THE BOUNDARIES OF BIG DATA 177-79 (Bart van der Sloot et al. eds., 2016).
57 Id. at 321.
58 PETTIT, supra note 41, at 1.
59 Id.
60 See, for some central references, those cited supra at note 45.
discontinuity,61 which also takes aim at personal liberty and human flourishing by enabling what she calls “the play of everyday practice”62 is generally consistent with the republican theory we argue for here.63 As the first author has argued elsewhere,

Privacy promotes human flourishing and antipower by making individuals less visible (obscure), creating “breathing room to engage in the processes of boundary management that enable and constitute self-development,” as described by Cohen, and is an important and instrumental means of supporting the goal of maintaining individual liberty from . . . intrusion, interference, and/or domination.64

Semantic discontinuity refers to “gaps and inconsistencies within systems of meaning, and to a resulting interstitial complexity that leaves room for the play of everyday practice.”65 Balkin explains that, “although Cohen initially states this idea in terms of meanings, it soon becomes clear that she is talking about gaps in enforcement and in systems of surveillance and control.”66 Additionally, the notion of obscurity—and its explanatory emphasis on how “information can be obtained or kept out of reach, correctly interpreted or misunderstood”67—and its links to privacy and surveillance, as outlined by Hartzog, Stutzman, and Selinger is also very relevant and useful here.68

---

62 Cohen, supra note 43, at 1932; COHEN, supra note 43, at 266.
63 NEWELL, POLICE VISIBILITY, supra note 46, at 57 (“Cohen’s concept of semantic discontinuity resonates with much of the broader republican theory of privacy and access to information I argue for here.”).
64 Id. at 56 (quoting Cohen, supra note 43, at 1906).
65 Cohen, supra note 43, at 224.
Consider the classic, “paradigmatic neorepublican example” of the slave and (benevolent) master. If the master (slaveholder) never interferes with the slave’s life or choices, liberal theories of negative liberty cannot explain why the slave is worse off than if she had not been a slave in the first instance—that is, because the master has never actually interfered in her life. However, to the republican, this is confounding. The mere fact that the master could interfere with the slave’s life and choices suggests that the slave is subject to the arbitrary whims of the master, who thereby dominates the slave. This conception of domination as the antithesis of liberty has roots in ancient Rome, where concern about “the evil of being subject to a master, or dominus — suffering dominatio — and was contrasted with the good of libertas, or ‘liberty.’”

For Pettit, domination exists when one individual (or group of individuals), “A, will be dominated in a certain choice by another agent or agency, B, to the extent that B has a power of interfering in the choice that is not itself controlled by A.” Explaining further, he notes that

[when I say that B has a power of interference I mean that B has the uninvited and uninvaded capacity to interfere or not to interfere. And when I say that that power of interfering is not controlled by A, I mean that it is not exercised on terms imposed by A: it is not exercised in a direction or according to a pattern that A has the influence to determine.]

As such, neorepublican liberty has been framed as “freedom of agents, not of options.”

Some additional, important aspects of this position should be noted. First, if we understand privacy as instrumentally valuable to achieving the neorepublican idea of liberty (that is, non-domination), then we see that privacy must also be more than merely being left alone. Rather, privacy should take on characteristics

---

70 PETTIT, JUST FREEDOM, supra note 45, at 2.
71 Id. at 50; see also Capasso, supra note 69.
72 PETTIT, supra note 41, at 50.
73 Capasso, supra note 69 (citing Philip Pettit, Agency Freedom and Options Freedom, 15 J. THEORETICAL POL. 387 (2003)).
of what Pettit calls “antipower”—or the power to “command noninterference.”

Second, in addition to the core idea that domination can exist absent actual interference, not all interference is domination. According to Pettit, this is the case when interference is non-arbitrary, controlled, or non-alien. As such, neorepublican theory does not oppose regulation but actually promotes the application of regulatory control in ways that reduce domination and produce antipower. Domination, then, exists when interference is possible, subject only to the exercised will “or discretion of the interferer; interference that is uncontrolled by the person on the receiving end.” Interference, or the potential for interference, with your choices might be non-arbitrary in the sense that it “conforms to rules” (including the law), but it might still be “uncontrolled interference” when its application is not pursuant to nor aligned with your will. As noted by Pettit, “[t]he active, intentional restriction of your choice by any other agent or agency will be invasive only to the extent that it reflects a will that you do not control.”

Connecting these ideas to privacy and data protection, and specifically to the new wave of data protection laws in the United States, is relatively straightforward. Many of the concerns about unregulated commercial data markets and new computing technologies—such as artificial intelligence and machine learning algorithms—which facilitate automated decision-making, behavioral advertising, various other forms of consumer classification, and other uses of personal data—hinge on the fact that people’s personal data are being used in ways that affect them but that do not reflect their will. As such, commercial data practices involving consumer manipulation—such as targeted and behavioral

74 Pettit, supra note 42, at 589.
75 Capasso, supra note 69, at 182.
76 PETTIT, JUST FREEDOM, supra note 45, at 8-9.
77 PETTIT, supra note 41, at 58-59.
78 Philip Pettit, Republican Liberty: Three Axioms, Four Theorems, in REPUBLICANISM AND POLITICAL THEORY 102, 116 (Cécile Laborde & John Maynor eds., 2008).
79 PETTIT, supra note 41, at 58; Quentin Skinner, Freedom as the Absence of Arbitrary Power, in REPUBLICANISM AND POLITICAL THEORY 83 (Cécile Laborde & John Maynor eds., 2008).
80 PETTIT, supra note 41, at 58.
81 Id.
advertising, the use of dark patterns in designing user interfaces, or the use of algorithms in various contexts—such as in making employment, parole, criminal sentencing, and other decisions—are invasive because they are active, intentional restrictions on individuals’ choices that do not reflect the will of these individuals. They are, in short, forms of domination that must be resisted if privacy is to have any real, substantive existence within the new data protection regime gaining steam in the United States.

The role of law, however, is less clear. The concept of “notice and choice,” without much critical reflection, appears a valid mechanism for aligning peoples’ will with the use to which their data is put by companies who collect it. But the failure of that regime to substantiate peoples’ will into corporate data practices—even though people may have been presented with legal disclaimers and options to click or refuse to click buttons bearing the label, “I agree”—suggests that new approaches are necessary and that existing approaches need to be replaced. From a data abolitionist perspective, banning all commercial uses of personal data—at least those outside obvious and expected boundaries, such as might be captured through serious data minimization and purpose and retention limitation principles—would limit the amount of domination currently embedded in the system. But, as enticing as some aspects of that extreme position might be, they are probably not achievable nor a pragmatic solution.

In any case, if “privacy is valuable . . . because it shields the individual from domination and is a necessary condition for human flourishing,” attention to getting American data privacy law right is critical to preserving—or, perhaps more accurately, reclaiming—privacy as a fundamentally important aspect of our individual and collective liberty.

II. METHODOLOGY AND CASE SELECTION

States have adopted a variety of data privacy laws over the years. For purposes of this research, we sought to examine and compare only those laws considered “comprehensive.” We used the determination by the International Association of Privacy Professionals (“IAPP”), as represented on their U.S. State Privacy

82 NEWELL, POLICE VISIBILITY, supra note 46, at 56.
Legislation Tracker\textsuperscript{83} as a proxy for whether a state law should be included in our analysis. The text of these laws has also been subject to revision post-adoption.\textsuperscript{84} Notably, the versions used for this comparison were current as of August 1, 2022. Yet, even when these laws were current, not all their provisions were actually in force, as portions of the CCPA went into effect on January 1, 2023, the same effective date as the VaCDPA. Additionally, the effective dates of Colorado Privacy Act (“CoPA”) and Connecticut Data Privacy Act (“CtDPA”) was July 1, 2023, while the Utah Consumer Privacy Act (“UtCPA”) did not go into effect until December 31, 2023.

Our primary methodology is an interpretative approach to functional comparative legal analysis.\textsuperscript{85} As explained by De Coninck, explanatory functional comparative research is “aim\textsuperscript{[ed]} at explaining similarities and differences between legal systems and which, simultaneously, is theoretically informed and empirically supported.”\textsuperscript{86} This type of research addresses social issues in response to identified problems.\textsuperscript{87} We use this method for explanatory purposes both in “a spirit of inquiry,” as discussed by Samuel\textsuperscript{88} and under an interpretive philosophical approach like that advocated for by Michaels.\textsuperscript{89} As Michaels has framed and defended the interpretive approach:

> Explaining legal institutions functionally drives hypotheses that consider problems and structure not as realities (either empirical or philosophical), but rather make proposals about

\begin{flushleft}
\textsuperscript{83} See supra note 3.


\textsuperscript{87} Id. at 323.


\textsuperscript{89} Michaels, supra note 85, at 371-72.
\end{flushleft}
how societies can and should be understood, not just how they work. That a problem exists and that an institution is a response to it need not be proven; but the connection between events and institutions must be made plausible as a way of understanding. We may well say that a problem is constructed and still assign explanatory power to it; we may analyze from a particular non-universal viewpoint and offer this analysis as one of several possible interpretations. Functionalism thereby turns from a scientific to an interpretative approach to law, a way of ‘making sense’ that is distinct from the participants’ way of making sense of their legal systems. This would be problematic for a positive science. It is not so for comparative law understood as an argumentative and normative, purpose-oriented discipline.90

Despite its benefits, the functional comparative research approach has been criticized for, among other things, confusing “comparability with similarity.”91 We recognize issues of comparability between data protection legislation in the United States and that adopted in the EU under a different legal process and within a different legal and social context. Indeed, the functional method can be seen as “calling for the study of law-in-context, underlining the importance of legal as well as non-legal context to fully appreciate how certain problems are actually dealt with in the different legal systems under scrutiny.”92 This problem is less dramatic in the case of comparing the laws among U.S. states (although each also operates within a somewhat different legal context as well) than it is when comparing across international borders and with a supranational legal order. However, we posit that limiting our analysis to a comparison of the core elements relevant to the material scope of these laws mitigates these methodological challenges. In that light, we approach our analysis with the aim of assessing these laws for “likenesses and differences,” rather than representing them as being “similar.”93

Additionally, the functional method has been criticized in that “focusing on ‘problem’ and ‘function’ creates a mere illusion of objectivity” and that any pre-determined attention to particular

90 Id.
91 De Coninck, supra note 86, at 325.
92 Id. at 336.
93 Id. at 326.
values is culturally driven.\textsuperscript{94} We take this point and want to recognize that our personal standpoints throughout this research are limited to our individual experiences and backgrounds. This inherent bias is part of the “struggle” with which comparative researchers must engage while simultaneously recognizing the “the difficulty of, on the one hand, appreciating the function of a rule or institution within its own social context while, on the other hand, simultaneously attempting to reach comparability of rules or institutions by placing them in a neutral external comparative framework,”\textsuperscript{95} and addressing their own personal positionalities and perspectives as researchers and human beings.

With those caveats in mind, we argue that comparing the material scope of data privacy legislation of various U.S. states and the GDPR can be a fruitful exercise. This approach helps elucidate the likenesses and differences between these regulations, enhancing our understanding of their effectiveness in protecting individual rights—of privacy and related interests covered by data protection law—and in restricting corporate and state domination. Ultimately, a comparison allows for the “argumentative and normative, purpose-oriented”\textsuperscript{96} approach of legal scholarship.

III. COMPARING THE MATERIAL SCOPE OF U.S. AND EU DATA PRIVACY LAWS

“Personal data” and related terms, such as “personal information” or “personally identifiable information,” are foundational to modern data protection and information privacy laws.\textsuperscript{97} A defining characteristic of a true data protection law is that “legal protections follow personal data” rather than client-business relationships.\textsuperscript{98} This is in contrast to something like consumer...

\textsuperscript{94} Id. at 327.
\textsuperscript{95} Id. at 327-28.
\textsuperscript{96} Michaels, supra note 85, at 371.
\textsuperscript{97} Nadezhda Purtova, The Law of Everything, Broad Concept of Personal Data and Future of EU Data Protection Law, 10 L. INNOVATION & TECH. 40, 41-42 (2018) [hereinafter Purtova, The Law of Everything]; see also Lee A. Bygrave & Luca Tosoni, Article 4(1). Personal Data, in THE EU GENERAL DATA PROTECTION REGULATION (GDPR): A COMMENTARY 103, 105 (Christopher Kuner et al. eds., 2020) (“[P]ersonal data’ . . . is a threshold concept for the application of data protection law generally.”).
\textsuperscript{98} Chander et al., supra note 5, at 1749 (arguing that the CCPA has data protection characteristics, in part because “under the CCPA, legal protections
protection-based privacy law. Typically, the foundational core of what data protection laws regulate is the “processing” of personal data in particular contexts. For example, Article 2 of the GDPR defines the law’s material scope as “the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system.”

Personal data and processing, along with their various elements, have also been called the “boundary concepts of data protection law.” The primary purpose of this Article is unpacking the definitions accorded to these terms, an essential first step towards understanding the scope of data protection.

Several of the definitions relevant to the material scope of these laws also reference the territorial scope of the law. As such, few short points on that issue are in order before we dive into the broader analysis. While the GDPR, by virtue of its application across the entire EU, maintains a broader territorial scope than the state-based laws in the United States, each of these laws is limited in application by references to the physical location of either data subjects or data processors (or controllers) or to their respective jurisdictional connections. For example, the GDPR applies to the conduct of controllers and processors located in the EU and to the conduct of non-EU controllers and processors when they process personal data about data subjects located in the EU, as long as

the processing activities are related to . . . the offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union; or . . . the monitoring of their behaviour as far as their behaviour takes place within the Union.

In the U.S. context, the jurisdictional reach of the law at issue is, as expected, tied in important ways to the geographical boundaries of the jurisdiction itself. Thus, an American federal law (such as the ADPPA, should it or something like it eventually be enacted), would encompass data processing operations relating to any “natural

---

99 GDPR, supra note 16, art. 2(1).
100 Nadezhda Purtova, From Knowing by Name to Targeting: The Meaning of Identification Under the GDPR, 12 INT’L DATA PRIVACY L. 163, 163 (2022) [hereinafter Purtova, From Knowing by Name to Targeting].
101 GDPR, supra note 16, art. 3(2).
person residing in the United States”\textsuperscript{102} and, since the ADPPA’s definition of “covered entity” includes those subject to the powers of the Federal Trade Commission Act,\textsuperscript{103} it would have (like the GDPR) some extraterritorial effect.\textsuperscript{104} Similarly, the laws of the various states are focused on commercial activities that occur within the geographical boundaries of each respective state, maintaining some extraterritorial effect as well. Each of the state laws examined here limits its definition of a data subject (or “consumer”) to natural persons who are state residents.\textsuperscript{105}

And, although it is not necessary to cover the full complexities of how each state law defines the entities subject to their regulations at this juncture, we do see that the laws at issue extend only to for-profit businesses that do business within the respective state or who target residents of the respective state with goods or services.\textsuperscript{106}

In the following subsections, we first examine and compare the subject matter to which these laws apply. Then, we examine the

\begin{itemize}
  \item \textsuperscript{102} American Data Privacy and Protection Act, H.R. 8152, 117th Cong. § 2(16) (2022).
  \item \textsuperscript{103} See id. § 2(9)(A)(i)(I) (“The term ‘covered entity’ . . . means any entity or any person, other than an individual acting in a non-commercial context, that alone or jointly with others determines the purposes and means of collecting, processing, or transferring covered data and—(I) is subject to the Federal Trade Commission Act . . . “).
  \item \textsuperscript{104} See, e.g., 15 U.S.C. 45(a)(4)(A) (extending the FTC’s enforcement authority over “unfair or deceptive acts or practices” to include “such acts or practices involving foreign commerce that—(i) cause or are likely to cause reasonably foreseeable injury within the United States; or (ii) involve material conduct occurring within the United States.”)
  \item \textsuperscript{105} See, e.g., CAL. CIV. CODE § 1798.140(i) (definition of “Consumer”); VA. CODE § 59.1-575 (definition of “Consumer”); COLO. REV. STATS. § 6-1-1303(6) (definition of “Consumer”); UTAH CODE § 13-61-101(10) (definition of “Consumer”); CONN. GEN. STAT. § 42-515(8) (2023) (definition of “Consumer”).
  \item \textsuperscript{106} See, e.g., CAL. CIV. CODE § 1798.140(d) (2022) (definition of “Business” encompasses certain entities that do business in the State of California”); VA. CODE § 59.1-576 (2023) (stating that the law “applies to persons that conduct business in the Commonwealth or produce products or services that are targeted to residents of the Commonwealth”); COLO. REV. STAT. § 6-1-1304(1)(a) (2023) (the law “applies to a controller that . . . Conducts business in Colorado or produces or delivers commercial products or services that are intentionally targeted to residents of Colorado”); UTAH CODE ANN. § 13-61-101(12) (2023) (defining a “controller” as “a person doing business in the state who determines the purposes for which and the means by which personal data are processed”); UTAH CODE ANN. § 13-61-102(1) (2023) (applying the law to any “controller or processor who . . . conducts business in the state; or . . . produces a product or service that is targeted to consumers who are residents of the state”); CONN. GEN. STAT. § 42-516 (2023) (noting that the law applies “to persons that conduct business in this state or persons that produce products or services that are targeted to residents of this state”).
\end{itemize}
legislation to determine what conduct regarding such subject matter is regulated—that is, what covered entities can or cannot do with personal data, typically referred to as forms of “processing” personal data. The laws of the various jurisdictions apply various definitional forms to “processing,” which can include the collection, use, storage, retention, analysis, sharing, disclosure, transfer, or sale of personal data. Finally, we examine which entities acquire obligations under the laws—that is, whose conduct is regulated. In the following subsections, we examine each of these elements in turn, beginning with questions of subject matter and then turning to processing and covered entities, comparing the law in each U.S. state to each other and to the GDPR.

a. Subject Matter

Within the GDPR, the definition of personal data has been rendered increasingly broad, causing some to argue that European data protection law may well become the “law of everything.” The GDPR defines personal data to mean “any information relating to an identified or identifiable natural person (‘data subject’).” The basic comparable definition in the CCPA is of “personal information,” which is defined as “information 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.” At the time of its adoption, and arguably still at present, the CCPA defined personal information much more broadly than almost all other US privacy laws. Notably, the basic definitions for “personal data” included in the VaCDPA, CoPA, UtCPA, and CtDPA are essentially identical. They each use the term to mean any “information that is linked or reasonably linkable to an identified [individual] or [an] identifiable natural

---

107 Purtova, The Law of Everything, supra note 97.
110 Chander et al., supra note 5, at 1750.
Because they are alike in many ways, we refer to these statutes, in the aggregate, as the “Virginia Model Statutes” at various points throughout our analysis. The draft text of the ADPPA took a slightly different approach, referring to “covered data,” which was defined as “information that identifies or is linked or reasonably linkable, alone or in combination with other information, to an individual or a device that identifies or is linked or reasonably linkable to an individual, and may include derived data and unique identifiers.”

Beyond the basic definitions, these laws also define and provide examples of specific types of data that fall within the general scope of what constitutes personal data or information, to varying degrees of specificity. For example, the CCPA’s definition of “personal information” also includes examples of several types of data that will constitute personal information as long the data “identifies, relates to, describes, is reasonably capable of being associated with, or could be reasonably linked, directly or indirectly, with a particular consumer or household.” These include, among others, “unique personal identifier,” “characteristics of protected classifications,” records related to consumers’ commercial activities, “biometric information,” information about a consumer’s interactions with websites or advertisements, “geolocation data,” “audio, electronic, visual, thermal, olfactory, or similar information,” records related to employment or education, sensitive information, and “inferences drawn from any of the information identified in this subdivision to create a

111 VA. CODE § 59.1-575 (2023) (definition of “Personal data”); see also COLO. REV. STAT. § 6-1-1303(17); UTAH CODE § 13-61-101(24) (2023); CONN. GEN. STAT. § 42-515(26) (2023).
112 We use “Virginia” because it was the first state to pass such legislation.
115 Id. § 1798.140(v)(1)(A) (discussed further infra at Part IV(A)(3)).
116 Id. § 1798.140(v)(1)(C).
117 Id. § 1798.140(v)(1)(D).
118 Id. § 1798.140(v)(1)(E).
119 Id. § 1798.140(v)(1)(F).
120 Id. § 1798.140(v)(1)(G).
121 Id. § 1798.140(v)(1)(H).
122 Id. § 1798.140(v)(1)(I).
123 Id. § 1798.140(v)(1)(J).
124 Id. § 1798.140(v)(1)(L).
profile about a consumer reflecting the consumer’s preferences, characteristics, psychological trends, predispositions, behavior, attitudes, intelligence, abilities, and aptitudes.”

The CCPA separately defines several of these categories. For example, the law defines biometric information quite broadly, especially in comparison with the Virginia Model Statutes, as

an individual’s physiological, biological, or behavioral characteristics, including information pertaining to an individual’s deoxyribonucleic acid (DNA), that is used or is intended to be used singly or in combination with each other or with other identifying data, to establish individual identity. Biometric information includes, but is not limited to, imagery of the iris, retina, fingerprint, face, hand, palm, vein patterns, and voice recordings, from which an identifier template, such as a faceprint, a minutiae template, or a voiceprint, can be extracted, and keystroke patterns or rhythms, gait patterns or rhythms, and sleep, health, or exercise data that contain identifying information.

In comparison, the VaCDPA defines “biometric data,” but does so in a much more limited way, excluding any “physical or digital photograph, a video or audio recording or data generated therefrom, or information collected, used, or stored for health care treatment, payment, or operations under HIPAA.” This definition clearly excepts biometric information that might be collected, for example, through the use of visual facial recognition technologies.

Utah’s law tracks the Virginia approach but provides a slightly broader exception, excluding from its definition data generated from “a physical or digital photograph” as opposed to only video and audio recordings. Connecticut’s basic definition tracks the law in Virginia and Utah. However, the CtDPA does not contain the exception for health care related information that is found in both

125 Id. § 1798.140(v)(1)(K).
126 See id. § 1798.140(c) (definition of “Biometric information”).
127 See Va. CODE § 59.1-575 (2023) (defining “biometric data” as “data generated by automatic measurements of an individual’s biological characteristics, such as a fingerprint, voiceprint, eye retinas, irises, or other unique biological patterns or characteristics that is used to identify a specific individual”).
129 See Utah CODE § 13-61-101(6)(a) (2023) (defining “Biometric data” as “data generated by automatic measurements of an individual’s unique biological characteristics”).
the VaCDPA and UtCPA, and it excepts data generated from photographs, video, or audio recordings, except when “such data is generated to identify a specific individual.”131 Colorado law omits any distinct definition for “biometric data,” only referring to the term as a form of “sensitive data” when processed “for the purpose of uniquely identifying an individual.”132

The CCPA also defines “sensitive personal information” quite broadly, to include:

information that reveals . . . [a] consumer’s social security, driver’s license, state identification card, or passport number[,] . . . account log-in, financial account, debit card, or credit card number in combination with any required security or access code, password, or credentials allowing access to an account[,] . . . precise geolocation133[,] . . . racial or ethnic origin, religious or philosophical beliefs, or union membership[,] . . . The contents of a consumer’s mail, email, and text messages unless the business is the intended recipient of the communication[,] . . . [or] A consumer’s genetic data.134

Sensitive personal information also encompasses the processing of biometric information to identify a consumer, and collection and analysis of information about a person’s health, sex life, or sexual orientation.135

---

133 The CCPA defines “precise geolocation” to mean “any data that is derived from a device and that is used or intended to be used to locate a consumer within a geographic area that is equal to or less than the area of a circle with a radius of 1,850 feet, except as prescribed by regulations.” Cal. Civ. Code § 1798.140(w) (2018). Relatedly, a “device” is defined as “any physical object that is capable of connecting to the Internet, directly or indirectly, or to another device.” Cal. Civ. Code § 1798.140(o) (2018). The VaCDPA also defines “precise location data” as a type of sensitive data, using different language than the CCPA that includes a 1,750-foot radius and excluding location data derived from “the content of communications or any data generated by or connected to advanced utility metering infrastructure systems or equipment for use by a utility.” Va. Code Ann. § 59.1-575 (2023) (definition of “Precise geolocation data”). The UtCPA and CtDPA are essentially identical to that in Virginia. Utah Code § 13-61-101(33) (definition of “Specific geolocation data”); Conn. Gen. Stat. § 42-515(27) (2023) (definition of “Precise geolocation data”). Colorado law does not define geolocation data, referring to it only as a means of identification. Colo. Rev. Stat. § 6-1-1305(16) (2022) (definition of “Identified or identifiable individual”).
The VaCDPA defines “sensitive data” as
a category of personal data that includes:

1. Personal data revealing racial or ethnic origin, religious beliefs, mental or physical health diagnosis, sexual orientation, or citizenship or immigration status;
2. The processing of genetic or biometric data for the purpose of uniquely identifying a natural person;
3. The personal data collected from a known child; or
4. Precise geolocation data.

This definition is substantially narrower than the CCPA in some respects, but it also includes data not encompassed by the CCPA’s language. For example, the VaCDPA language does not explicitly cover the contents of a consumer’s communications, their philosophical beliefs to the extent they are not religious beliefs, information about a person’s sex life, account numbers and passwords, and identity documents. However, it does include “citizenship or immigration status” and personal data collected from a person whom the business knows is a child.136 The Virginia law also includes something of a privacy-by-default position with an opt-in mechanism, effectively barring controllers from processing sensitive data unless and until they acquire “the consumer’s consent.”137 Comparably, the CCPA merely requires businesses who collect sensitive personal information to inform consumers about these practices138 and to provide them with the opportunity to limit—through an opt-out mechanism—some future processing of such information.139

Colorado’s definition for sensitive data is clearly based on Virginia’s language but does differ in some respects. For example, CoPA includes information about a person’s “sex life,” replaces “immigration status” with “citizenship status,” explicitly covers a physical or mental health condition rather than only diagnosis and does not include precise geolocation information.140 Like the

---

136 VA. CODE § 59.1-575.
137 Id. § 59.1-578(A)(5).
138 CAL. CIV. CODE §§ 1798.100(a)(2).
139 Id. §§ 1798.121(a).
140 COLO. REV. STAT. § 6-1-1303(24).
Virginia law, CoPA also prohibits processing sensitive data without the consent of the data subject or their parent.\textsuperscript{141}

Utah’s law largely mirrors Virginia’s text, but expands the range of health and medical information considered sensitive data.\textsuperscript{142} Oddly, Utah’s law excludes information about a person’s “racial or ethnic origin, if the personal data is processed by a video communication service.”\textsuperscript{143} Perhaps the biggest difference between Utah’s law and the laws in Virginia and Colorado is that Utah does not provide a privacy-by-default setting for processing sensitive data, merely requiring controllers and processors to provide the consumer with a “clear notice and an opportunity to opt out of the processing.”\textsuperscript{144} Finally, Connecticut law also draws on the Virginia language but, like Colorado, covers “sex life” and health “condition[s]”; unlike Colorado, Connecticut retains the language covering “immigration status” and “precise geolocation data.”\textsuperscript{145} The CtDPA also copies Virginia’s and Colorado’s opt-in position in regard to processing sensitive data.\textsuperscript{146}

The legal definitions for personal data or personal information contained in these statutes, including the GDPR, encapsulate several distinct elements, each of which potentially impacts the scope of what counts as personal data.\textsuperscript{147} As we see from the analysis, these elements include:

- **Inputs**: data \textit{and/or} information, although the laws rarely offer definitional guidance as to what these terms mean;
- **Links**: these terms link the inputs to the subjects, describing which links meet the laws’ standards, and include semantic or syntactic terms such as about, relating to, linked to, linkable to, describes, and capable of being associated with;
- **Identification**: the inputs must be linked to subjects who are \textit{identified} or \textit{identifiable};\textsuperscript{148} as such, this is a crucial definitional term;

\textsuperscript{141} Id. § 6-1-1308(7).
\textsuperscript{142} UTAH CODE § 13-61-101(32)(a) (although some exceptions also apply to the benefit of licensed health care facilities, as enumerated in § 13-61-101(32)(b)).
\textsuperscript{143} Id. § 13-61-101(32)(b)(i).
\textsuperscript{144} Id. § 13-61-302(3)(a).
\textsuperscript{145} CONN. GEN. STAT. § 42-515(38) (2023).
\textsuperscript{146} Id. § 42-520(a)(4).
\textsuperscript{147} See, e.g., Purtova, \textit{From Knowing by Name to Targeting}, supra note 100, at 245-46 (finding that the EU’s Working Party 29 breaks the definition of personal data into four elements that fit these categories).
\textsuperscript{148} However, in the ADPPA’s proposed language, the inputs need only be “linked or reasonably linkable to an individual,” regardless of whether the
Subjects: the laws variously refer to natural persons, consumers, households, individuals—and, in some cases, even consumer devices linkable to individuals—and these data subjects are the individuals who acquire rights under the law.\footnote{149}

Additionally, U.S. laws provide specific exceptions to the general subject matter definitions, including exceptions for “public information” and “de-identified” data, as just two initial examples. They also specifically refer to several additional subtypes of personal data, including “biometric data,” \footnote{150} “geolocation data,” \footnote{151} health information, \footnote{152} “sensitive data,” \footnote{153} and to the partly de-identified category of “pseudonymous data.” \footnote{154}

In the subsections that follow, we examine and compare how these laws define and scope the inputs, links, subjects, and identification that make up the primary subject matter addressed by each law.

i. Inputs

Even though all of these laws base their central definitions of personal data or personal information on data and information, respectively, they offer little or no definitional guidance as to what these terms are supposed to mean. The GDPR does not define individual is identified or identifiable, which seems to be a much broader notion than contemplated under any of the other legislation examined in this Article. See American Data Privacy and Protection Act, H.R. 8152, 117th Cong. § 2(8)(A).

\footnote{149} See, e.g., \textsc{Cal. Civ. Code} § 1798.140(i) (defining a “[c]onsumer” as a “natural person” resident in the state); \textsc{Colo. Rev. Stat.} § 6-1-1303(6)(a) (defining a “[c]onsumer” as an “individual” resident in the state); \textsc{Conn. Pub. Acts} § 42-515(8) (defining a “[c]onsumer” as an “individual” resident in the state); \textsc{Utah Code} § 13-61-101(10)(a) (defining a “[c]onsumer” as an “individual” resident in the state); \textsc{Va. Code Ann.} § 59.1-575 (defining a “[c]onsumer” as a “natural person” resident in the state).

\footnote{150} See \textsc{Va. Code} § 59.1-575 (defining “Biometric data”).

\footnote{151} See \textit{id.} (defining “Precise geolocation data”).

\footnote{152} See \textit{id.} (defining “Protected health information”); \textsc{Colo. Rev. Stat.} § 6-1-1303(13) (“Health-care information”); \textsc{Colo. Rev. Stat.} § 6-1-1303(21) (“Protected health information”).


\footnote{154} See \textsc{Va. Code} § 59.1-575 (defining “Pseudonymous data”).
The CCPA does not define "information." And the statutes based on the Virginia model—including VaCDPA, CoPA, UtCPA, and CtDPA—similarly define "personal data" as a type of "information" but do not define information itself. Presumably, the legislators (or drafters) assumed the terms would be interpreted broadly, as they are used in common parlance or as defined in traditional dictionaries. Outside the data protection context, at least one federal U.S. court has found that "information" is a "straightforward term [that] needs no further construction," and it seems the legislators behind these data protection laws may have a similar perspective.

In California, the state attorney general has held that personal information under the CCPA includes "internally generated inferences [a] business holds about [a] consumer," at least within the context of the right to access information granted by the CCPA. This origin-based classification of data as derived or inferred, alongside provided and observed data, was first made in the policy context during the Organization for Economic Cooperation and Development (“OECD”) privacy expert roundtable in 2014 and later adopted by the World Economic Forum “to raise awareness as to the scale of personal data processing, and the various types of personal data that are processed.” The distinction was later adopted in the EU in the context of the right to data portability which applies to actively provided and observed data but not derived and inferred data. The Article 29 Working Party (“WP29”), an influential group established under the Data Protection Directive before the GDPR’s enactment, issued

155 See Nadezhda Purtova & Ronald Leenes, Code as Personal Data: Implications for Data Protection Law and Regulation of Algorithms, 13 INT’L DATA PRIVACY L. 245, 250 (2023); see also Purtova, From Knowing by Name to Targeting, supra note 100, at 245, 250.
156 See Memorandum Opinion and Order at 47, Personalized Media Commc’n v. Motorola, Inc., 2:08-CV-70-CE (E.D. Tex. Sept. 30, 2011) (finding “both data and software code as types of information” and rejecting a definition that would limit information to simply “data.”).
159 WORLD ECONOMIC FORUM, RETHINKING PERSONAL DATA: A NEW LENS FOR STRENGTHENING TRUST 16(e).
160 Inge Graef et al., Data Portability and Data Control: Lessons for an Emerging Concept in EU Law, 19 GERMAN L.J. 1359 (2018), at 1373.
161 GDPR, supra note 16, art. 20(1).
guidelines with persuasive authority on interpreting the Directive. According to WP29, data types differ: *provided* data is “actively and knowingly provided by the data subject,” *observed* data is “provided by the data subject by virtue of the use of the service or the device,” and “inferred data and derived data are created by the data controller on the basis of the data ‘provided by the data subject’” and include outcomes like assessments, profiling (e.g., assigning a credit score), and other algorithmic results.\(^\text{162}\) In the *Nowak*\(^\text{163}\) case, the Court of Justice of the European Union (“CJEU”) ruled that personal data “potentially encompasses all kinds of information, not only objective but also subjective, in the form of opinions and assessments . . . .”\(^\text{164}\) Therefore, inferred and provided data, provided they relate to an identified or identifiable individual, are considered personal data under the GDPR.\(^\text{165}\)

However, scholars have rightly criticized data protection law for having undertheorized and underdeveloped definitions of these essential terms.\(^\text{166}\) According to Gellert, “it seems that information and its meaning are taken for granted, and in so doing justify the regulatory reach of data protection in the first place.”\(^\text{167}\) Similarly, Purtova has argued that the failure to define what information means “may have a truly explosive effect on the range of situations that would fall within the scope of data protection law.”\(^\text{168}\) Bygrave has also argued that contemporary reality of mass data collection and data processing requires better definitional clarity.\(^\text{169}\)


\(^{163}\) Case C-434/16, Peter Nowak v. Data Protection Commissioner, 2017 ECLI:EU:C:2017:994.

\(^{164}\) Id. at ¶ 34.

\(^{165}\) But see Sandra Wachter & Brent Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*, 2019 COLUM. BUS. L. REV. 494 (2019) (arguing that algorithmic inferences do not fall within the scope of personal data under the GDPR).


\(^{167}\) Gellert, *supra* note 166, at 157.


\(^{169}\) Bygrave, *supra* note 166, at 91-92.
One obvious source for definitional clarity is the academic field of information itself. However, within the realistic and formal information sciences (sometimes called informatics), information has a complicated and contested meaning. Much of the literature agrees that information has a semantic nature—that is, information relates to meaning or the communication of knowledge, as in informing or becoming informed—but other disciplinary perspectives have conceptualized and defined the term in other ways as well. And despite the fact that law and everyday parlance often treat information and data as synonyms, their relationship is also the subject of significant academic commentary, analysis, and debate.

On one influential view, data has been defined as the building blocks for information and knowledge. Relatedly, information is commonly understood as “data + meaning.” These propositions are also consistent with the widely-regarded idea that data, information, knowledge, and wisdom exist in a sort of pyramidal relationship in which each serves as the foundational element of the form that sits on top—that is, data + meaning = information; information + understanding = knowledge; and so forth. Likewise, this is consistent with arguments from multiple scholars that data “exist prior to argument or

170 Michael Buckland, Information and Society 1 (2017).
171 See, e.g., Bygrave, supra note 166, at 92 (using the term to “encompass a range of overlapping disciplines, most notably information science, computer science, information philosophy, and knowledge management.”).
172 See e.g., Rafael Capurro & Birger Hjørland, The Concept of Information, 37 ANN. REV. INFO. SCI. & TECH. 343 (2003); Gellert, supra note 166, at 157.
174 Kitchin, supra note 173, at 3.
176 Gellert, supra note 166, at 157; Kitchin, supra note 173; see R.L. Ackoff, From Data to Wisdom, 16 J. APPLIED SYS. ANALYSIS 3 (1989); see also Donald O. Case & Lisa M. Given, Looking for Information 74-76 (4th ed. 2016). But the basic concept of the “DIKW” pyramid has also been critiqued. See, e.g., Jennifer Rowley, The Wisdom Hierarchy: Representations of the DIKW Hierarchy, 33 J. INFO. SCI. 163, 177 (2007) (arguing that there is a lack of consensus in the literature about “the processes that transform elements lower in the hierarchy into those above them.”); Martin Frické, The Knowledge Pyramid: A Critique of the DIKW Hierarchy, 35 J. INFO. SCI. 131, 132 (2009) (“[T]he DIKW pyramid should be abandoned. It should no longer be part of the canon of information science, and such related disciplines as systems theory, information management, information systems, knowledge management, and library and documentation science.”).
interpretation”—the very processes that generate information, facts, or evidence. The semantic nature of information as a product of well-formed and meaningful data is represented in Floridi’s General Definition of Information (“GDI”), which states:

σ is an instance of information, understood as semantic content, if and only if: σ consists of n data, for n ≥ 1; the data are well formed; the well-formed data are meaningful.

Another contested idea within this literature is whether something is information if a human observer never observes it. Kind of like the old adage about whether sound exists in a forest if no one is there to hear the tree fall, some commentators have argued that information only exists if it is observed by, and possibly informative to, a human being. On this theory, information is “any difference that makes a difference to a conscious, human mind.” This argument complicates the application of informational concepts to the data protection realm when much of what concerns such law is the automated processing and analysis of (personal) data by machines. Alternatively, we might simply conclude that, for purposes of data protection law, a human observer is not required, and that data can be personal data even when only observed and processed by machines.

Lastly, in an influential and foundational piece within the discipline(s) of library and information science, Buckland attributes three primary meanings to the term information: information-as-process, information-as-knowledge, and information-as-thing. Information-as-process refers to the acts of communicating or receiving knowledge, while information-as-knowledge refers to that which has been perceived as part of such a process. Buckland’s third type refers to objects—or “things”—that “are regarded as being informative,” such as data, documents, or records. It seems plausible to conclude that data protection law generally refers to this form of

---


178 FLORIDI, supra note 175, at 21.

179 See, e.g., CASE & GIVEN, supra note 176, at 74 (noting the related, but “common notion that knowledge is information that has been sifted, organized, and understood by a human brain”).

180 Id. at 56, citing GREGORY Bateson, STEPS TO AN ECOLOGY OF MIND 453 (1972).

information within its definitions of personal data. That is, that data and information are things—signals, signs, or other instances of difference—that have the capacity to inform or bear meaning for an observer, whether human or machine.

ii. Links

The laws subject to comparative analysis in this paper contain several different, but related, semantic and syntactic linking terms. That is, depending on the specific law at issue, the inputs (data/information) must relate to, be linked or linkable to, describe, or be capable of being associated with an identified or identifiable data subject. Other provisions within these laws bring even more terms to bear, including whether data is “associated with” a data subject. Each of these linking terms, whether semantic or syntactic in nature, invites interpretation and may evolve differently in each jurisdiction as courts and data protection authorities—where they exist, such as in the EU and California—interpret and provide meaning to these terms when applied in actual cases and regulatory contexts. Indeed, the GDPR’s definition of personal data requires “a tailored, context-specific analysis for deciding whether or not personal data is present.”

We can draw some basic distinctions between the scope of subject matter by looking to the linking terms included in each state’s law, but these comparisons are limited as we don’t have much formal interpretation of some of these laws due to the fact that they have only recently been enacted.

As an important side note, Kitchin’s conceptualization of data encompasses multiple forms of links or relationships between data and subject: data can be representative, implied, or derived (what some also refer to as “inferred”); as measurements of some phenomena in the world, generated by implication, or derived from analysis of multiple pieces of data.

This account of data is also consistent with analyses of the relating to requirement in EU data protection law. As Purtova has shown, understanding what constitutes these semantic or syntactic link between data and data

---

183 KITCHIN, supra note 177, at 1.
184 For a good description of this requirement in EU data protection law, see Purtova, From Knowing by Name to Targeting, supra note 10097, at 166.
subject is crucial to understanding the material scope of the GDPR\textsuperscript{185} and other related data privacy laws. For example, Article 4(1) of the GDPR defines personal data as

any information relating to an identified or identifiable natural person (“data subject”); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.\textsuperscript{186}

The crucial linking term in the GDPR is “relating to,” linking “personal data” with “an identified or identifiable natural person” (or data subject). Neither the GDPR nor the Data Protection Directive that preceded it provide guidance as to how the term “relating to” should be interpreted.\textsuperscript{187} However, WP29—now replaced by the European Data Protection Board\textsuperscript{188}—provided some commentary on the concept of personal data and the meaning of “relating to” in EU data protection law.\textsuperscript{189} Specifically, WP29 explained that “data relates to an individual if it refers to the identity, characteristics or behavior[]r of an individual or if such information is used to determine or influence the way in which that person is treated or evaluated.”\textsuperscript{190} This commentary supports the interpretation offered by WP29 which, as summarized by Purtova, suggests that,

[i]nformation can “relate” to an individual in content, purpose, or result, meaning that information “relating to” a natural person includes but is broader than the information ‘about’ that person. The meaning of “relating to” grows even

\begin{itemize}
\item \textsuperscript{185} Purtova, The Law of Everything, supra note 9797, at 53-57, 66-72.
\item \textsuperscript{186} GDPR, supra note 16, art. 4(1).
\item \textsuperscript{187} Purtova, The Law of Everything, supra note 9797, at 44.
\item \textsuperscript{188} Michèle Finck & Frank Pallas, They Who Must Not Be Identified—Distinguishing Personal from Non-Personal Data Under the GDPR, 10 INT’L DATA PRIV. L. 11, 11 (2020).
\item \textsuperscript{189} Article 29 Working Party, Opinion 4/2007 on the Concept of Personal Data, June 20, 2007 [hereinafter WP 136].
\item \textsuperscript{190} See id. at 10; see also Sille Obelitz Søe et al., What Is the “Personal” in “Personal Information?” 23 ETHICS & INFO. TECH. 625, 627 (2021).
\end{itemize}
broader considering that these three conditions are meant as alternative and not as cumulative ones.\textsuperscript{191}

Thus, information \textit{about a person}—that is, when it relates to an individual in its semantic \textit{content}—is clearly personal data under the GDPR. But information that is not about a person may still relate to that person, and thus be considered personal data by virtue of the \textit{purpose} of its use, if the “the data are used or are likely to be used . . . with the purpose to evaluate, treat in a certain way or influence the status or behavior of an individual.”\textsuperscript{192} These links may be purely syntactic, and may not carry semantic meaning about a person at all. And, finally, information may also be considered personal data even if it is not \textit{about} a person and is not used for the purpose of evaluating or influencing a person, when its “use is likely to have an impact on a certain person’s rights and interests”\textsuperscript{193}. Obviously, then, the scope of personal data under the GDPR is tremendously broad.

Complicating matters of comparison, the laws in California, Virginia, Colorado, Utah, and Connecticut use a variety of additional linking terms. For example, the CCPA defines “personal information” as “information 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.”\textsuperscript{194} Setting aside analysis of the “or household” language, which potentially enlarges the scope of California law in several ways,\textsuperscript{195} we see that the CCPA applies to information that “relates to” a consumer or household, but also to information that “identifies” (which could be information \textit{about} a person or, in a syntactic sense, information that can function so as to identify a person, such as a cookie, IP address, or other identifier) or “describes” (which seems to be semantically \textit{about} a person in terms of \textit{content}) a consumer, and to information that is “reasonably capable of being associated with, or could reasonably be linked,

\begin{itemize}
\item \textsuperscript{191} Purtova, \textit{The Law of Everything}, supra note 97, at 54.
\item \textsuperscript{192} \textit{Id.} at 54 (quoting WP 136, supra note 189, at 10). See Soe et al., \textit{supra} note 190, at 627 (noting that the Data Protection Working Party specifies that personal information includes “information that is likely to have an impact on an individual”).
\item \textsuperscript{193} WP 136, supra note 189, at 11; Purtova, \textit{The Law of Everything}, supra note 97, at 54.
\item \textsuperscript{194} CAL. CIV. CODE § 1798.140(v)(1).
\item \textsuperscript{195} See Margot Kaminski et al., \textit{Symposium, The California Consumer Privacy Act}, 54 LOY. L.A. L. REV. 157, 162 (2020) (Felix Wu noting that the inclusion of “household” in the definition makes the definition “unusually broad”).
\end{itemize}
directly or indirectly” with a consumer. Whether this terminology will be interpreted in line with the GDPR’s *purpose* or *result* concepts—or in a more limited way—is still to be seen, but the phrasing does suggest a potentially broad scope for what constitutes personal information under the CCPA, potentially on par with the GDPR. Additionally, the existing regulations promulgated by the California Attorney General and the amended regulations proposed in 2022 by the California Privacy Protection Agency do not offer any guidance on the meaning or scope of these linking terms. At present, California-based courts have simply concluded that the law provides a “broad definition,” finding that it covers the sort of information collected by, for example, Google on users of its services. This includes data that identifies the individual being targeted through unique identifiers, device identifiers, geolocation, IP address, cookie match, and other information of account holders’ internet communications and highly detailed personal profile information about peoples’ interests, race, religion, sexual orientation, and health status.

Elsewhere, the VaCDPA provides what appears to (potentially) be, on its face, a narrower definition, wherein “personal data” means “any information that is linked or reasonably linkable to an identified or identifiable natural person.” As noted earlier, CoPA, UtCPA, and CtDPA all use nearly identical language. However, we have yet to see how the terms “linked” and “linkable” will be interpreted in practice, and whether this will be limited to data that is linkable merely because it is about a person—identifying or describing them, as a semantic link—or will be applied more

---

196 Felix Wu suggests that this framing means that “the definition goes beyond actually identifying an individual or a household,” which raises the “question as to what that would mean: What does it mean for information to relate to or describe an individual or household even if the information is not necessarily identifiable to that individual or household?” *Id.* at 162.


198 The proposed (at time of writing) regulations have been promulgated by the California Privacy Protection Agency and are available at https://cppa.ca.gov/regulations/ [https://perma.cc/6728-R7D2].


200 *In re Google*, *supra* note 199, at 939.

broadly such that it resembles information that relates to a person in
a broader syntactic sense, such as through purpose or result.

As noted already, data subjects must be identified or at least
identifiable in order for information to be considered personal data
under existing frameworks. In the next subsection, we examine how
these laws treat identification and identifiability as central
boundary-marking concepts in data protection law.

iii. Identification

Identification (and identifiability) is the “touchstone” of
personal data. It provides an essential distinction between
personal data and non-personal data upon which the rights and
obligations of existing data protection law rest. Data subjects
include persons who, for example, “can be readily identified,
directly or indirectly.”

Identification is, in some ways, a linking term, but it is also a
descriptive qualifier about the data subject in ways that the other
linking terms are not. For example, some definitions of personal data
refer to data that identifies a natural person—a semantic link—but
even when data is “personal” because it relates to a person only in a
syntactic sense, that person must still be identified or identifiable for
the law to regulate processing that data. For these reasons, and due
to its prominence in existing literature, we treat identification
separately here as a distinct aspect of the material scope of these
laws.

1. The “Identified” and/or “Identifiable” Data Subject

Under the GDPR, identification has received a wealth of
scholarly attention. Under Article 4(1) of the GDPR, information
is only personal data when it relates to an identified or identifiable natural person ("data subject"), where "identified" refers to the fact and "identifiable" to the possibility of identification. Identifiable” has traditionally been seen as a standard easier to meet and hence a de facto threshold for application of the GDPR. Yet, “identified” started gaining more practical significance as national caselaw emerged where courts found that a natural person was “sufficiently identified” (e.g., by the facial images captured by a CCTV camera) and that it was not necessary to address the question of identifiability.

In the context of the Data Protection Directive, WP29 wrote that “a natural person can be considered as ‘identified’ when, within a group of persons, he or she is ‘distinguished’ from all other members of the group.” Some national courts interpreted identification as individuation. Clearly, this concept of identification is broader than merely linking data to a person’s legal name. In addition to civil identification, Leenes distinguishes three other types of identification: recognition, classification, and session identification. Purtova adds targeting to this list. Marx also speaks of seven types of “identity knowledge.”

Article 4(1) of the GDPR defines an identifiable natural person as

one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic,

---

206 WP 136, supra note 189, at 12.
207 Id.
208 R (on the application of Edward Bridges) v. Chief Constable S. Wales Police, [2020] EWHC Civ 1058, at 124; see also Purtova, From Knowing by Name to Targeting, supra note 100, at 180–82.
209 WP 136, supra note 189189, at 12; Purtova, From Knowing by Name to Targeting, supra note 100, at *3.
210 Vidal-Hall v. Google, [2015] EWCA Civ 311, at 114 (“Identification for the purposes of data protection is about data that ‘individuates’ the individual, in the sense that they are singled out and distinguished from all others.”); Chief Constable S. Wales Police, EWHC Civ 1058, at 119.
212 Purtova, From Knowing by Name to Targeting, supra note 100, at 169.
mental, economic, cultural or social identity of that natural person.214

Thus, a data subject can be either identified or identifiable and such identification can be either direct or indirect and can refer to various types of identity. In reference to the legal texts examined here, personal information can also be de-identified, whether made anonymous or pseudonymous, and re-identified – for example, through the use of additional data and/or technological processes. Recital 26 of the GDPR provides interpretative guidance regarding when a natural person is identifiable or, in other words, when the possibility of identification is legally relevant.

To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments.215

Thus, Recital 26 established a test of means likely reasonably to be used for identification, which calls for a fact-dependent analysis that controllers and processors must engage in on a case-by-case basis.216 The test has been applied and further developed in the CJEU caselaw. In Breyer, the Court ruled that in order to satisfy the test of reasonable likelihood of identification, the means necessary for identification do not need to be in hands of one person.217 Furthermore, the Court ruled that identification would not be reasonably likely if it was “prohibited by law or practically impossible” due to “a disproportionate effort in terms of time, cost and man-power, so that the risk of identification appears in reality to be insignificant.”218

---

214 GDPR, supra note 16, art. 4(1)
215 Id. at Recital 26.
216 See Schwartz & Solove, supra note 182, at 886 (subjecting personal data to “a tailored, context-specific analysis for deciding whether or not personal data is present”).
218 Id. at 46.
In addition to “all objective factors” listed in Recital 26 (the cost and amount of time required for identification, and the available technology), WP29 suggests that the purpose of data processing is decisive for determining identifiability. In cases “where the purpose of the processing implies the identification of individuals, it can be assumed that the controller or any other person involved have or will have the means ‘likely reasonably to be used’ to identify the data subject.” This makes the boundaries of what constitutes identifiability broad and prevents an argument that the GDPR does not apply to the processing at hand while such processing “only makes sense if it allows identification of specific individuals and treatment of them in a certain way.” Yet, the resulting broad interpretation of identifiability where the purpose of processing is a decisive factor is not binding and has not yet been confirmed in the caselaw of the CJEU.

In the U.S. context, the Virginia Model Statutes all rely on identification. They all refer, in identical phrasing, to persons “who can be readily identified, directly or indirectly.” Additionally, CoPA specifies that this ready identification can be accomplished “by reference to an identifier such as a name, an identification number, specific geolocation data, or an online identifier.” Presumably, a name would provide direct identification while, for example, specific geolocation data would provide indirect identification—such that it is “linked or reasonably linkable” to a data subject. The CCPA’s definition of personal information does not depend as squarely on identification and, in any event, does not refer to ready identification. However, the CCPA does includes identification as a linking term in the definition of personal information alongside others and, in its definition of a “consumer” is limited to natural persons who have been identified.

The requirement in the Virginia Model Statutes that identification must be “readily” accomplishable clearly limits the scope of what identification means under these statutes, but what exactly this means has not yet been examined in any great detail.

219 WP 136, supra note 189, at 16.
220 Id.
221 VA. CODE ANN. § 59.1-575 (emphasis added); COLO. REV. STAT. § 6-1-1303(16) (emphasis added); UTAH CODE ANN. § 13-61-101(20) (emphasis added); Conn. Pub. Acts. § 1(15) (emphasis added).
222 COLO. REV. STAT. § 6-1-1303(16).
223 VA. CODE ANN. § 59.1-575.
224 CAL. CIV. CODE § 1798.140(i).
Must the covered entity (controller or processor) who wishes to process data be able to readily identify an individual? Or would it suffice that technological means exist to make such identification an easy and routine process, regardless of whether the covered entity itself has the capability to do so at the time of processing? It seems reasonable to assume that the word “readily” might be seen as a limit on the language of the GDPR that hinges identifiability on not only the ability of the controller, but also “by another person to identify the natural person directly or indirectly.”\(^{225}\) In any case, it seems reasonable to assume that this would exclude processes of identification that take time and extensive resources.

Within its definition of personal information, the CCPA also refers to various types of “identifiers,”\(^{226}\) which include “unique [personal] identifiers”\(^{227}\) and the undefined term, “online identifier.”\(^{228}\) These identifiers can also be either “probabilistic” or “persistent.”\(^{229}\) Identifiers include information such as “a real name, alias, postal address, unique personal identifier, online identifier, Internet Protocol address, email address, account name, social security number, driver’s license number, passport number, or other similar identifiers.”\(^{230}\) A “unique identifier” is defined as a persistent identifier that can be used to recognize a consumer, a family, or a device that is linked to a consumer or family, over time and across different services, including, but not limited to: a device identifier; an Internet Protocol address; cookies, beacons, pixel tags, mobile ad identifiers, or similar technology; customer number, unique pseudonym, or user alias; telephone numbers, or other forms of persistent or probabilistic identifiers that can be used to identify a particular consumer or device that is linked to a consumer or family.\(^{231}\)

Importantly, a “probabilistic identifier” is one that makes “identification of a consumer or a consumer’s device . . . more

\(^{225}\) GDPR, supra note 16, Recital 26 (emphasis added).
\(^{226}\) CAL. CIV. CODE § 1798.140(v)(1)(A).
\(^{227}\) See id.; see also Civ. § 1798.140(a)(j) (defining “unique identifier” or “unique personal identifier”).
\(^{228}\) Civ. § 1798.140(v)(1)(A).
\(^{229}\) See § 1798.140(x) (defining “[p]robabilistic identifier”); Civ. § 1798.140(a)(j) (including both persistent and probabilistic identifiers within the definition of “unique personal identifier” and “unique identifier”).
\(^{230}\) Civ. § 1798.140(v)(1)(A).
\(^{231}\) Civ. § 1798.140(a)(j).
"probable than not," based on the context. This bears resemblance to the GDPR’s identifiability standard, or of means reasonably likely to be used. Yet, it is not clear if the “probabilistic identifier” refers to one single piece of information that must be sufficient on its own for “more probable than not” identification, or whether it also implies that multiple pieces of information can be considered probabilistic identifiers when combined, thus enabling identification.

2. De-Identified Data

Examining the definitions provided for “de-identified data” is also helpful in constructing an image of what identification means. In this respect, the Virginia Model Statutes largely follow the language of the CCPA. All of the U.S. laws exempt de-identified data from their definitions of personal data or personal information. Under the CCPA, data is “deidentified” when it “cannot reasonably be used to infer information about, or otherwise be linked to, a particular consumer,” but only so long as the controller:

1. Takes reasonable measures to ensure that the information cannot be associated with a consumer or household.
2. Publicly commits to maintain and use the information in deidentified form and not to attempt to reidentify the information, except that the business may attempt to reidentify the information solely for the purpose of determining whether its deidentification processes satisfy the requirements of this subdivision.
3. Contractually obligates any recipients of the information to comply with all provisions of this subdivision.

232 CIV. § 1798.140(x) (emphasis added).
233 CIV. § 1798.140(v)(3); VA. CODE ANN. § 59.1-575 ("‘Personal data’ does not include de-identified data or publicly available information."); COLO. REV. STAT. § 6-1-1303(17)(b); UTAH CODE ANN. § 13-61-101(24)(b); Conn. Pub. Acts § 1(18).
234 CAL. CIV. CODE § 1798.140(m). The ADPPA draft language include a similar provision, although with some differences. See American Data Privacy and Protection Act, H.R. 8152, 117th Cong. § 2(10) (definition of “de-identified data”).
Thus, even though deidentified data is not considered personal data, the law does provide some obligations for processing deidentified data.\textsuperscript{235}

Virginia law also defines “de-identified data” as “data that cannot reasonably be linked to an identified or identifiable natural person, or a device linked to such person,”\textsuperscript{236} presumably by result of some technological process of obscuring or removing ready links between the data and a data subject. Colorado law defines “de-identified data” as “data that cannot reasonably be used to infer information about, or otherwise be linked to, an identified or identifiable individual, or a device linked to such an individual.”\textsuperscript{237} Utah and Connecticut law essentially match the Virginia and Colorado provisions.\textsuperscript{238}

Moving beyond “de-identified data,” and similarly to the GDPR, the Virginia Model Statutes also refer to “pseudonymous” data as a sub-type of personal data, although the CCPA’s treatment of this term is much more limited.\textsuperscript{239} Pseudonymous data is defined in the VaCDPA as:

personal data that cannot be attributed to a specific natural person without the use of additional information, provided that such additional information is kept separately and is subject to appropriate technical and organizational measures to ensure that the personal data is not attributed to an identified or identifiable natural person.\textsuperscript{240}

\textsuperscript{235} See also VA. CODE ANN. § 59.1-581(A) (“The controller in possession of de-identified data shall: 1. Take reasonable measures to ensure that the data cannot be associated with a natural person; 2. Publicly commit to maintaining and using de-identified data without attempting to re-identify the data; and 3. Contractually obligate any recipients of the de-identified data to comply with all provisions of this chapter.”); COLO. REV. STAT. § 6-1-1303(11) (nearly identical text); UTAH CODE ANN. e (nearly identical text); Conn. Pub. Acts § 1(13) (nearly identical text).

\textsuperscript{236} VA. CODE ANN. § 59.1-575 (definition for “de-identified data”).

\textsuperscript{237} COLO. REV. STAT. § 6-1-1303(11).

\textsuperscript{238} UTAH CODE ANN. § 13-61-101(14); Conn. Pub. Acts § 1(13).

\textsuperscript{239} While the CCPA also refers to the process of “pseudonymization,” this term is only applicable to data processing within the context of “research” as defined in the statute. CAL. CIV. CODE § 1798.140(aa)–(ab). The GDPR does not use the term “pseudonymous data” but has a definition of pseudonymization. GDPR, supra note 16, art. 4(5).

\textsuperscript{240} VA. CODE ANN. § 59.1-575 (definition of “pseudonymous data”); COLO. REV. STAT. § 6-1-1303(22) (nearly identical text); UTAH CODE ANN § 13-61-101 (nearly identical text); Conn. Pub. Acts § 1(24) (nearly identical text).
Similarly to the GDPR, as long as controllers keep pseudonymous data separate from other data that would enable re-identification and maintain appropriate technical and organizational measures in place to ensure re-identification does not take place, such data may not be subject to some of the data subjects’ rights—namely the right of access, the right of deletion, the right to correction (except under Utah law, which does not contain a right to correct inaccuracies), and the right to data portability.

In Europe, the GDPR’s definition of personal data excludes anonymous data—that which is “rendered anonymous in such a manner that the data subject is not or no longer identifiable” but includes pseudonymized data. “Pseudonymization” under the GDPR, means the processing of personal data in such a manner that the personal data 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 organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

Recital 26 declares that pseudonymous data, “which could be attributed to a natural person by the use of additional information

---

241 VA. CODE ANN. § 59.1-581(B); UTAH CODE ANN. § 13-61-303(2).
242 VA. CODE ANN. § 59.1-581(D) (exempting the right of access under § 59.1-577(A)(1)); UTAH CODE ANN. § 13-61-303(2) (exempting the right of access under § 13-61-201(1)); COLO. REV. STAT. § 6-1-1307(3) (exempting the right of access under § 6-1-1306(1)(b)); Conn. Pub. Acts No. 22-15 § 9(d) (exempting the right of access under § 4(a)(1)).
243 VA. CODE ANN. § 59.1-581(D) (exempting the right of deletion under § 59.1-577(A)(3)); UTAH CODE ANN. § 13-61-303(2) (exempting the right of deletion under § 13-61-201(2)); COLO. REV. STAT. § 6-1-1307(3) (exempting the right of deletion under § 6-1-1306(1)(d)); Conn. Pub. Acts No. 22-15 § 9(d) (exempting the right of deletion under § 4(a)(3)).
244 VA. CODE ANN. § 59.1-581(D) (exempting the right to correction under § 59.1-577(A)(2)); COLO. REV. STAT. § 6-1-1307(3) (exempting the right to correction under § 6-1-1306(1)(c)); Conn. Pub. Acts No. 22-15 § 9(d) (exempting the right of access under § 4(a)(2)).
245 VA. CODE ANN. § 59.1-577(A)(4), § 59.1-581(D); UTAH CODE ANN. § 13-61-303(2) (exempting the right of data portability under § 13-61-201(3)); COLO. REV. STAT. § 6-1-1307(3) (exempting the right to data portability under § 6-1-1306(1)(e)); Conn. Pub. Acts 22-15 § 9(d) (exempting the right of access under § 4(a)(4)).
246 GDPR, supra note 16, Recital 26.
247 Purtova, The Law of Everything, supra note 97, at 44.
248 GDPR, supra note 16, art. 4(5).
should be considered to be information on an identifiable natural person”\textsuperscript{249} to which the GDPR applies. However, if a controller can “demonstrate that it is not in a position to identify the data subject” —for example, by deleting additional information needed to re-identify a data subject using pseudonymous data\textsuperscript{250}— then the GDPR exempts the controller from responding to data subjects’ requests under the right to access, right to rectification (termed “correction” in U.S. laws), right to erasure (also known as the right to be forgotten), right to restrict future processing, and the right to data portability.\textsuperscript{251} However, using this exception by deleting additional information may not even be necessary, since a controller is not obliged “to maintain, acquire or process additional information in order to identify the data subject for the sole purpose of complying with [the GDPR],”\textsuperscript{252} and combining data sets which were previously kept separately constitutes such processing of additional information. The GDPR also relaxes several other controller obligations when pseudonymous data is concerned.\textsuperscript{253} Recital 28 declares that “the application of pseudonymization to personal data can reduce the risks to the data subjects concerned and help controllers and processors to meet their data-protection obligations,”\textsuperscript{254} e.g., the requirement of data protection by design and by default.\textsuperscript{255}

Thus, we see similar distinctions in U.S. and European laws between de-identified or anonymous data, on the one hand, and pseudonymized data, on the other hand. In their framing of anonymization, all the U.S. laws seem to anticipate the possibility of re-identification in a way that the GDPR’s definition of anonymous

\textsuperscript{249} Id. Recital 26.


\textsuperscript{251} GDPR, supra note 16, art. 11(2).

\textsuperscript{252} Id. art. 11(1).

\textsuperscript{253} See, e.g., id. art. 6(4)(e) (noting that pseudonymization may be an appropriate safeguard when authorizing future uses of data that are “compatible” with the purposes used for initial collection); id. art. 89(1) (finding that pseudonymization is an important safeguard for processing personal data for scientific, historical and statistical purposes); id. art. 25(1) (to help facilitate data protection by design aims of the GDPR); id. art. 32(1) (to meet data security requirements).

\textsuperscript{254} Id. Recital 28.

\textsuperscript{255} Id. art. 25.
data does not. This is explainable by the fact that if re-identification of “anonymous data” is reasonably likely, the data is no longer anonymous but personal under the GDPR. Regarding pseudonymous data, U.S. laws require controllers to keep additional identifying information separate and subject to technical and organizational controls to limit the risk of re-identification. This generally matches the GDPR’s requirements for data protection by design and by default, which include maintaining data in pseudonymous form—that is, keeping the identifying data separate and taking technical and organizational measures to prevent re-identification. It has been noted that, under the GDPR, “[t]he same piece of data can be anonymous at the time of collection, but turn into personal later, just sitting there, simply by virtue of technological progress.” And, under Recital 26 of the GDPR, designating a piece of information as personal data requires ascertaining “whether means are reasonably likely to be used to identify the natural person . . . .” In so doing, “account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments.”

The analysis above leads to the conclusion that across the U.S. data privacy laws and the GDPR, anonymous, pseudonymous, and de-identified data are not firm, unchanging designations, and such data may become personal (identifiable) data for a variety of reasons, from technological development to inadvertent data management practices that fail to operate as sufficient technical or organizational barriers to re-identification. Relatedly, numerous authors have argued that “anonymized information always carries some risk of re-identification,” suggesting that the common exception for so-called “de-identified” data may sit on a shaky conceptual foundation in the first place.

---

256 Purtova, The Law of Everything, supra note 97, at 44.
257 GDPR, supra note 16, Recital 26.
258 Id.
260 See infra, Part IV(A)(5).
Because each of these laws treat the concepts of identification and de-identification in relatively alike ways, substantial differences may only become apparent, if at all, through future interpretation of the statutory language in each jurisdiction. Much more has been written on these issues in relation to the GDPR, and it remains to be seen whether courts and data protection agencies in the United States (such as the California Privacy Protection Agency) tasked with interpreting the law will diverge in their positions about what identification means in practice.

iv. Subjects

The laws examined here use several terms to describe data subjects, referring to natural persons, consumers, households, and individuals. It is perhaps a basic point, but it is important to note that the subjects in all these laws are human beings, or “natural persons,” excluding other types of legal persons which may be recognized under the laws of each jurisdiction. Each law also requires that these data subjects be identified or identifiable, as discussed above.

The subject within the GDPR is an “identified or identifiable natural person,” who the law also refers to as a “data subject.”261 Showing their nature as consumer protection laws, all U.S. state laws enacted to date refer to the relevant data subject as a “consumer.”262 For example, the CCPA’s definition of personal data is linked to only that data as it relates to a “consumer,” which the law defines as “any natural person who is a California resident [under California law] however identified, including by any unique identifier.”263

The VaCDPA applies to an “identified or identifiable natural person,” which “means a person who can be readily identified, directly or indirectly.”264 However, a person who can exercise rights under the law is defined as a “consumer,” which means “a natural person who is a resident of the commonwealth acting only in an individual or household context. It does not include a natural person

---

261 GDPR, supra note 16, art. 4(1).
262 CAL. CIV. CODE § 1798.140(i); VA. CODE § 59.1-575; COLO. REV. STAT. § 6-1-1303(6); UTAH CODE ANN § 13-61-101(10); Conn. Pub. Acts No. 22-15 § 1(7).
263 CAL. CIV. CODE § 1798.140(i).
264 VA. CODE ANN. § 59.1-575.
acting in a commercial or employment context.” This is in contrast to the GDPR where personal data relating to the data subject does not have to concern private or family life, and could pertain to his or her life in professional and other capacities. The statute does not define what either “individual or household context” or “commercial or employment context” mean, although the exemptions do provide some possible explanations. The CCPA also exempts personal information collected by a business that is processed about a person acting in an employment context, including as a job applicant to the business, as well as personal information collected during a commercial transaction between the person and the business when the person is acting on behalf of another business or organization.

Like the VaCDPA, CoPA applies to the personal data of the “identified or identifiable individual,” which means, “an individual who can be readily identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, specific geolocation data, or an online identifier.” CoPA then defines a “consumer” as any “individual who is a Colorado resident acting only in an individual or household context,” excluding individuals “acting in a commercial or employment context, as a job applicant, or as a beneficiary of someone acting in an employment context.”

The UtCPA also applies to personal data linkable to an “identified ... or an identifiable individual,” which means “an individual who [is or] can be readily identified, directly or indirectly.” A “consumer” likewise refers to “an individual who is a resident of the state acting in an individual or household context.

---

265 Id. (emphasis added).
266 WP 136, supra note 189, at 6.
267 See VA. CODE ANN. § 59.1-576(C)(14) (exempting personal data “processed or maintained (i) in the course of an individual applying to, employed by, or acting as an agent or independent contractor of a controller, processor, or third party, to the extent that the data is collected and used within the context of that role” for some definitional guidance as to what “employment context” or “commercial context” might mean); see also UTAH CODE ANN. § 13-61-102(2)(o)(i); Conn. Pub. Acts No. 22-15 § 3(b)(15)(A).
268 CAL. CIV. CODE § 1798.145(m)(1)(A).
269 CIV. § 1798.145(n)(1).
270 COLO. REV. STAT. § 6-1-1303(16).
271 COLO. REV. STAT § 6-1-1303(6).
273 Id. § 13-61-101(20).
[but] does not include an individual acting in an employment or commercial context.” The CtDPA shares a likeness, defining personal data in relation to an “identified or identifiable individual,” which means “an individual who can be readily identified, directly or indirectly.” Like the laws in Virginia, Colorado, and Utah, the CtDPA defines a “consumer” is an “individual who is a resident of [Connecticut]” except when “acting in a commercial or employment context.” However, the Connecticut law does provide a bit more guidance and specificity—as compared to the other Virginia Model Statutes—as to what this means:

“Consumer” does not include an individual acting in a commercial or employment context or as an employee, owner, director, officer or contractor of a company, partnership, sole proprietorship, nonprofit or government agency whose communications or transactions with the controller occur solely within the context of that individual’s role with the company, partnership, sole proprietorship, nonprofit or government agency.

Thus, we see that data subjects are human beings (natural persons), who are identified or identifiable in relation to data. The clear consumer protection grounding of the U.S. laws is an important limitation to the scope of those laws, especially when comparing the Virginia Model Statutes against the GDPR. If individuals are only rightsholders when acting as “consumers” in individual or household contexts, the rights granted by these laws clearly only follow the business-consumer relationship and not the data. As such, Chander, Kaminski, and McGeveran’s distinction between true data protection laws and lesser, more limited data privacy laws, is instructive. All these laws are clearly more limited in scope than the GDPR as relates to defining data subjects and the range of contexts in which data protection rights accrue under the law. However, the CCPA’s definition is also notably broader than the Virginia Model Statutes in that it includes data that relates to or could be linked to a particular household, rather than being limited to

274 Id. § 13-61-101(10).
276 Id. at § 1(7).
277 Id.
278 Chander et al., supra note 5, at 1749-50.
a specific individual. Under the CCPA, a “household,” as the subject which can be identified, includes “a group, however identified, of consumers who cohabitate with one another at the same residential address and share use of common devices or services.”

Other proposals have drawn on the language of the CCPA but would have gone even further. For example, Senate Bill 46 (2021-2022) in Massachusetts included “device” alongside identifiable individuals and households. And, notably, the ADPPA draft text referred to a “device” as being capable or identifying or linking to an “individual”—meaning “a natural person residing in the United States.”

v. Additional Exceptions

Several notable exceptions to what constitute the subject matter of these data privacy laws exist. As discussed previously, U.S. laws consider de-identified data to fall outside the scope of personal data, at least so long as the data controller meets certain conditions. U.S. laws also include major exceptions for “publicly available information,” while California includes a related exception for “lawfully obtained, truthful information that is a matter of public concern.” California and Utah also have exceptions for “aggregate consumer information.”

1. Publicly Available Information

All of the laws examined here, those in the United States as well as the GDPR, share the idea that some information may be placed in

---

279 CAL. CIV. CODE § 1798.140(v)(1), (q).
280 CIV. § 1798.140(q).
284 See CAL. CIV. CODE § 1798.140(v)(2); VA. CODE ANN. § 59.1-575 (“Personal data” is defined to “not include... publicly available information.”); COLO. REV. STAT. § 6-1-1303(17)(b); UTAH CODE ANN. § 13-61-101(24); Conn. Pub. Acts No. 22-15 § 1(18).
285 CAL. CIV. CODE § 1798.140(v)(2).
286 CIV. § 1798.140(b).
the public domain and therefore should not be subject to data processing restrictions. However, important differences between these jurisdictions exist. Of course, the balance of freedom of expression versus privacy (and data protection) is different in the United States and EU. In the United States, the First Amendment potentially complicates data protection aims, particularly within the context of information that has been disclosed widely or lawfully obtained by other means. Thus, we do see, as expected, a more robust exception for publicly available information in the U.S. laws. Where the GDPR’s exception is largely limited to allowing controllers to process sensitive personal data when the data are manifestly made public by the data subject herself.

The “publicly available information” exceptions present in U.S. state laws are arguably present as necessary conditions to keep these laws consistent with the First Amendment to the U.S. Constitution and its protections for free speech. The CCPA, for example, exempts “publicly available information or lawfully obtained, truthful information that is a matter of public concern”\textsuperscript{287} including “sensitive personal information” that meets the definition for being public.\textsuperscript{288} The CCPA defines “publicly available” as

information that is lawfully made available from federal, state, or local government records, or information that a business has a reasonable basis to believe is lawfully made available to the general public by the consumer or from widely distributed media; or information made available by a person to whom the consumer has disclosed the information if the consumer has not restricted the information to a specific audience. “Publicly available” does not mean biometric information collected by a business about a consumer without the consumer’s knowledge.\textsuperscript{289}

The Virginia and Utah laws include almost identical definitions, clearly informed by the CCPA, except that they do not include the final caveat about biometric information.\textsuperscript{290} Colorado’s law includes

\textsuperscript{287} Civ. § 1798.140(v)(2).
\textsuperscript{288} Civ. § 1798.140(ae)(3).
\textsuperscript{289} Civ. § 1798.140(v)(2).
\textsuperscript{290} See VA. CODE ANN. § 59.1-575 (defining “publicly available information” as “information that is lawfully made available through federal, state, or local government records, or information that a business has a reasonable basis to believe is lawfully made available to the general public through widely distributed media, by the consumer, or by a person to whom the consumer has disclosed the information, unless the consumer has restricted the information to a specific
very similar language, but in a much more concise formulation: “information that is lawfully made available from federal, state, or local government records and information that a controller has a reasonable basis to believe the consumer has lawfully made available to the general public.”

Connecticut’s law is almost identical to Colorado’s formulation, except that it also includes “information lawfully made available through . . . widely distributed media.”

No such broad exception for “publicly available information” applies in the GDPR context. Indeed, automated data processing has not been framed as a freedom of speech issue in Europe, except in the contexts of journalism and limitations on the “right to be forgotten,” or erasure. Moreover, the GDPR includes “the collection of personal data from documents in the public domain” within the scope of what constitutes “processing.” Instead, the GDPR refers to “personal data which are manifestly made public by the data subject” only as an exception to a general ban on processing “special categories of personal data”—the GDPR term for “sensitive” personal data, including data that might reveal a person’s racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning

---

292 See Conn. Pub. Acts No. 22-15 § 1(25) (defining “publicly available information” to mean “information that (A) is lawfully made available through federal, state or municipal government records or widely distributed media, and (B) a controller has a reasonable basis to believe a consumer has lawfully made available to the general public”).
293 GDPR, supra note 16, art 17(3).
295 GDPR, supra note 16, art. 9(2)(e) (emphasis added).
health or data concerning a natural person’s sex life or sexual orientation.\textsuperscript{296}

Importantly, if sensitive personal data is manifestly made public by the data subject, only the ban on processing such data is lifted, but it is not, on its own, sufficient to make the processing legitimate. Controllers would still need to rely on one of the general grounds for legitimate processing under Article 5.\textsuperscript{297}

The requirement that data be manifestly made public requires a data subject to have acted affirmatively in this regard, with the subjective understanding that the result would be the data becoming public.\textsuperscript{298} Making data public includes actions such as “publishing the data in the mass media, putting them on online social network platforms or similar actions.”\textsuperscript{299} But, according to the European Data Protection Board, merely presenting oneself in a public space—in view of surveillance cameras, for example—may not be sufficient.\textsuperscript{300}

Despite the fact that the EU similarly maintains that freedom of speech is a fundamental right, the privileged position of privacy and data protection as similarly situated fundamental rights in EU law may also be (part of) the reason why such a broad exception for publicly available information is not found in the GDPR. However, the GDPR also addresses this balancing in respect to access to official documents (freedom of information).\textsuperscript{301}

All the U.S. statutes exempt information contained in federal, state, or local government records and which is lawfully obtained under freedom of information laws. Each of the statutes also exempts information when a processor “has a reasonable basis to believe” that such information was “lawfully made available to the

\begin{footnotesize}
\begin{itemize}
  \item[296] GDPR, supra note 16, art. 9(1).
  \item[297] Id. Recital 51.
  \item[298] Ludmila Georgieva & Christopher Kuner, Article 9. Processing of Special Categories of Personal Data, in THE EU GENERAL DATA PROTECTION REGULATION (GDPR): A COMMENTARY 365, 378 (Christopher Kuner et al. eds., 2020).
  \item[299] Id.
  \item[300] See id.
  \item[301] See, e.g., GDPR, supra note 16, Recital 154 ("This Regulation allows the principle of public access to official documents to be taken into account when applying this Regulation."); see also id. ("[L]aws should reconcile public access to official documents and the reuse of public sector information with the right to the protection of personal data and may therefore provide for the necessary reconciliation with the right to the protection of personal data pursuant to this Regulation.").
\end{itemize}
\end{footnotesize}
general public by the consumer.” However, beyond those shared characteristics, the laws differ somewhat in four important respects.

Firstly, the CCPA, VaCDPA, UtCPA, and CtDPA exempt information that a processor reasonably believes has been lawfully published in “widely distributed media,” regardless of the original source of such information. The Colorado law, only exempting information made public by the consumer, is narrower in this respect.

Secondly, the laws in California, Virginia, and Utah include a third-party exception. Under this exception, any information a processor obtains from a third party to whom the data subject has shared the information is not considered personal data, as long as when disclosing the information to the third party, the data subject did not “restrict[] the information to a specific audience.”

Thirdly, the CCPA removes “biometric information collected by a business about a consumer without the consumer’s knowledge” from the exemption, meaning that such information is still personal information subject to regulation. This text appears to track the European Data Protection Board’s interpretation of biometric data collection under the GDPR.

Finally, the CCPA also includes the related exemption for “lawfully obtained, truthful information that is a matter of public concern.” This language is clearly linked to a concern for First Amendment values, but it does potentially extend the CCPA’s exception further than those contained in the other statutes.

Thus, it appears that the CCPA’s publicly available information exception is both broader in some respects and narrower in others as compared with the other U.S. statutes. Excluding the biometric information issue, Colorado’s law has the narrowest exception overall, as it is limited to information made public by the data subject and does not contain a third-party exception. The VaCDPA and UtCPA, which contain both of these, appear the broadest exceptions overall. The CCPA, with its extension to “lawfully obtained, truthful information” and restriction on biometric

---

302 CAL. CIV. CODE § 1798.140(v)(2).
304 COLO. REV. STAT. § 6-1-1303(17)(b).
306 See references cited supra note 332300.
307 CAL. CIV. CODE § 1798.140(v)(2).
information, and CtDPA, with its omission of the third-party exception, sit somewhere in the middle. However, even the narrowest of these exceptions, that of information made public by the data subject herself, provides a tremendously broad exception for commercial data practices. Indeed, “businesses scraping publicly available personal data [from online websites or social media profiles] remain unregulated even by the most expansive state data privacy laws.” As such, this exception provides a glaringly large gap in regulation as compared with the GDPR.

2. Aggregated Consumer Data

The CCPA and UtCPA also include an exception for what they refer to as “aggregate consumer information” or “aggregated data.” The CCPA defines this as information relating to a “group or category of consumers, from which individual consumer identities have been removed, that is not linked or reasonably linkable to any consumer or household,” including by a device, but “does not mean one or more individual consumer records that have been deidentified.” Under the UtCPA, “aggregated data’ means information that relates to a group or category of consumers: (a) from which individual consumer identities have been removed; and (b) that is not linked or reasonably linkable to any consumer.” The utility of these provisions is not readily apparent, as the data captured by these definitions would be excluded from the respective definitions of personal information and personal data in the first instance. As such, the presence of these exceptions does not appear to make a substantial difference to the overall material scope of these laws.

The GDPR does not define aggregate data, neither does it contain an explicit exception for aggregate data. However, so long as the aggregate data does not contain information relating to the identified or identifiable individuals, it will be not considered personal but anonymous data and the GDPR will not apply. The

---

non-binding Recital 162 GDPR is the only clause in the GDPR where aggregate data is mentioned. It provides that data processing for statistical purposes “implies that the result of processing . . . is not personal data, but aggregate data.”

3. Other Exemptions

These laws also exempt a variety of other information from their regulatory scope, although some of the particulars differ slightly. These exemptions include, among other things, certain medical information covered by state and federal medical information laws, including the Health Insurance Portability and Accountability Act of 1996 (“HIPAA”); personal information collected during clinical trials or other biomedical or human subjects research governed by federal rules; credit information subject to regulation under the Fair Credit Reporting Act (“FCRA”); certain financial information subject to regulation under the federal Gramm-Leach-Bliley Act (“GLBA”); personal information processed pursuant to the federal Driver’s Privacy Protection Act of 1994; personal data regulated by the federal Children’s Online Privacy Protection Act of 1998

---

313 CAL. CIV. CODE § 1798.145(c)(1)(A)-(B); VA. CODE ANN. §§ 59.1-576(B); 59.1-576(C)(1)-(3), (5)-(9); COLO. REV. STAT. §§ 6-1-1304(2)(a)-(c), (e)-(h); UTAH CODE ANN. §§ 13-61-102(2)(g)(i)-(ii), (2)(g)(vii)-(ix), (2)(h); Conn. Pub. Acts No. 22-15 §§ 3(b)(1)-(2), (6)-(10).

314 CAL. CIV. CODE § 1798.145(c)(1)(C); VA. CODE ANN. § 59.1-576(C)(4); COLO. REV. STAT. § 6-1-1304(2)(d); UTAH CODE ANN. §§ 13-61-102(2)(g)(iii)-(iv); Conn. Pub. Acts No. 22-15 §§ 3(b)(3)-(5).

315 CAL. CIV. CODE § 1798.145(d)(1); VA. CODE ANN. § 59.1-576(C)(10); COLO. REV. STAT. § 6-1-1304(2)(i)(II); UTAH CODE ANN. § 13-61-102(2)(i)(ii); Conn. Pub. Acts No. 22-15 § 3(b)(11).


317 CAL. CIV. CODE § 1798.145(e); VA. CODE ANN. §§ 59.1-576(B); COLO. REV. STAT. § 6-1-1304(2)(j)(II); UTAH CODE ANN. § 13-61-102(2)(k).

318 VA. CODE ANN. § 59.1-576(C)(12); COLO. REV. STAT. § 6-1-1304(2)(j)(V); UTAH CODE ANN. § 13-61-102(2)(m); Conn. Pub. Acts No. 22-15 § 3(b)(13). Cf. with CAL. CIV. CODE § 1798.140(v)(1)) (including education information subject to the federal Family Educational Rights and Privacy Act (“FERPA”) within its definition of personal information while limiting certain business obligations for such information) and CAL. CIV. CODE § 1798.145(q) (exempting businesses from complying with certain requests related to educational records).

(“COPPA”); information related to a vehicle or vessel held by dealers and manufacturers in anticipation of repairs, recalls, or warranties; certain information processed by air carriers (airlines); or information disclosed as part of a job application process (including emergency contact information). These numerous exemptions carve out substantial amounts of otherwise personal data, deferring in many cases to pre-existing (federal) law, and provide ample basis for concluding that these U.S. laws are much less comprehensive than the GDPR.

b. Processing: Regulated Use of Data

The purpose of data protection law is to regulate the “processing” of personal data. Indeed, the aim of the GDPR is to regulate “all or most stages of the data processing cycle, including registration, storage, retrieval and dissemination of personal data.” The GDPR provides a broad definition of processing, with the explicit intention that such a definition will help ensure that the definition is “technologically neutral in order to prevent risks of circumvention.” Under the definition, processing includes:

any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

---

320 COLO. REV. STAT. § 6-1-1304(2)(j)(IV).
321 CAL. CIV. CODE § 1798.145(g)(1)-(2).
322 COLO. REV. STAT. § 6-1-1304(2)(l); UTAH CODE ANN. § 13-61-102(2)(q); Conn. Pub. Acts No. 22-15 § 3(b)(16).
323 CAL. CIV. CODE § 1798.145(m)(1); VA. CODE ANN. § 59.1-576(C)(14); COLO. REV. STAT. § 6-1-1304(2)(k); UTAH CODE ANN. § 13-61-102(2)(o); Conn. Pub. Acts No. 22-15 § 3(b)(15).
324 Tosoni & Bygrave, supra note 294, at 117.
325 Id.; see also GDPR, supra note 16, Recital 15 (“In order to prevent creating a serious risk of circumvention, the protection of natural persons should be technologically neutral and should not depend on the techniques used.”).
326 GDPR, supra note 16, art. 4(2).
This list is not exhaustive but merely represents examples of the type of processing operations that fall within the scope of the GDPR.\textsuperscript{327} The CJEU has adhered to a broad interpretation of this definition such that “the definition essentially covers any data processing application”—including processing that does not result in any storage or recording of the data\textsuperscript{328}—and interpretations of the language included in the GDPR have suggested that processing covers everything “from the collection to the erasure of data, including organization, use, consultation, disclosure by transmission, dissemination or otherwise making available... comparison and suppression.”\textsuperscript{329} Processing includes all data processing operations performed “wholly or partly by automated means” as well as manual operations when the data is contained in (or intended to be contained in) a “filing system”\textsuperscript{330} that conforms to certain criteria enumerated in Article 4(6).\textsuperscript{331}

Turning to U.S. law, the CCPA offers a similarly broad definition, wherein “[p]rocessing” includes “any operation or set of operations that are performed on personal information or on sets of personal information, whether or not by automated means.”\textsuperscript{332} Likewise, Virginia law defines processing as “any operation or set of operations performed, whether by manual or automated means, on personal data or on sets of personal data, such as the collection, use, storage, disclosure, analysis, deletion, or modification of

\begin{thebibliography}{99}
\bibitem{327} See, e.g., Tosoni & Bygrave, \textit{supra} note 294, at 120; Joined Cases C-141/12 and C-372/12, YS and Others, 2013 (AG opinion), para. 64.
\bibitem{328} See Tosoni & Bygrave, \textit{supra} note 294, at 119; see also Article 29 Working Party, Opinion 1/2015 on Privacy and Data Protection Issues Relating to the Utilisation of Drones, June 16, 2015 [hereinafter WP 231], at 7, n. 13. However, it should be noted that at least one type of operation has been found to not constitute processing, at least in an advisory opinion: legal analysis that incorporates personal information when such personal data is relevant to the analysis. See Tosoni & Bygrave, \textit{supra} note 294, at 120-21 (citing Joined Cases C-141/12 and C-372/12, YS and Others (AG opinion), ¶¶. 62-65).
\bibitem{330} GDPR, \textit{supra} note 16, art. 2(1).
\bibitem{331} GDPR, \textit{supra} note 16, art. 4(6) (defining a “filing system” as “any structured set of personal data which are accessible according to specific criteria, whether centrally[ ]ed, decentrali[ ]ed or dispersed on a functional or geographical basis”).
\bibitem{332} CAL. CIV. CODE § 1798.140(y) (defining of “processing”).
\end{thebibliography}
personal data.” The laws in Utah and Connecticut essentially mirror the Virginia provision. Like the GDPR, each of these laws include an exemplary list of types of operations that qualify. Colorado’s provision, however, differs somewhat and provides a narrower, contained set of operations that qualify as processing, defining “processing” as the “collection, use, sale, storage, disclosure, analysis, deletion, or modification of personal data and includes the actions of a controller directing a processor to process personal data.” As such, Colorado’s law reaches a narrower range of operations than any of the others, limiting its practical application and material scope.

We also see definitions for various identified types of processing within these laws, although most of the exemplary types of processing remain undefined. For example, the CCPA defines “collection,” “sharing,” and “selling” personal information, but the Virginia Model Statutes only define “sale,” omitting any definitions for collection, sharing, or any other form identified within their definitions for processing. The CCPA defines “collection” as “buying, renting, gathering, obtaining, receiving, or accessing any personal information pertaining to a consumer by any means. This includes receiving information from the consumer, either actively or passively, or by observing the consumer’s behavior.”

Additionally, the CCPA defines “sharing” as

sharing, renting, releasing, disclosing, disseminating, making available, transferring, or otherwise communicating orally, in writing, or by electronic or other means, a consumer’s personal information by the business to a third party for cross-context behavioral advertising, whether or not for monetary or other valuable consideration, including

---

333 VA. CODE ANN. § 59.1-575 (defining “process” and “processing”).
335 Conn. Pub. Acts No. 22-15 § 1(20) (defining “process” and “processing”).
336 COLO. REV. STAT. § 6-1-1303(18) (defining of “process” and “processing”).
338 See CAL. CIV. CODE § 1798.140(f) (2022) (“‘Collects,’ ‘collected,’ or ‘collection’ means buying, renting, gathering, obtaining, receiving, or accessing any personal information pertaining to a consumer by any means. This includes receiving information from the consumer, either actively or passively, or by observing the consumer’s behavior.”).
transactions between a business and a third party for cross-context behavioral advertising for the benefit of a business in which no money is exchanged.\textsuperscript{339}

Notably, this definition of “sharing” is limited to instances of data processing within the context of “cross-context behavioral advertising” and does not apply to similar operations in other contexts or settings.

The one type of processing we see defined in all five of these laws is the “sale” of personal data. The definitions provided are complex, subject to several exceptions. Under the CCPA, “sell,” “selling,” “sale,” and “sold” are defined as “selling, renting, releasing, disclosing, disseminating, making available, transferring, or otherwise communicating orally, in writing, or by electronic or other means, a consumer’s personal information by the business to a third party for monetary or other valuable consideration.”\textsuperscript{340} This does not apply when a consumer intentionally directs the controller or processor (a “business” under the CCPA) to “disclose personal information or intentionally interact with one or more third parties,”\textsuperscript{341} when the business shares an identifier to communicate to a third party that the data subject has opted out of the sale or processing of their (sensitive) personal information,\textsuperscript{342} or when the information is transferred as part of a “merger, acquisition, bankruptcy, or other transaction in which the third party assumes control of all or part of the business.”\textsuperscript{343}

The VaCDPA defines “sale of personal data” more succinctly as “the exchange of personal data for monetary consideration by the controller to a third party.”\textsuperscript{344} Virginia law also includes several

\begin{footnotesize}
\textsuperscript{339} See Civ. § 1798.140(ah)(1). Some exceptions to this general definition are also addressed in subsection 2, including when “[a] consumer uses or directs the business to intentionally disclose personal information or intentionally interact with one or more third parties.” See CAL. CIV. CODE § 1798.140(ah)(2)(A) (2022)). The business uses an identifier to alert “persons that the consumer has opted out of the sharing of the consumer’s personal information or limited the use of the consumer’s sensitive personal information.” See CAL. CIV. CODE § 1798.140(ah)(2)(B) (2022)). When the “business transfers to a third party the personal information of a consumer as an asset that is part of a merger, acquisition, bankruptcy, or other transaction in which the third party assumes control of all or part of the business . . . .” See CAL. CIV. CODE § 1798.140(ah)(2)(C) (2022).

\textsuperscript{340} CAL. CIV. CODE § 1798.140(ad)(1).

\textsuperscript{341} Civ. § 1798.140(ad)(2)(A).

\textsuperscript{342} See Civ. § 1798.140(ad)(2)(B).

\textsuperscript{343} Civ. § 1798.140(ad)(2)(C).

\textsuperscript{344} VA. CODE ANN. § 59.1-575 (2022).\end{footnotesize}
exceptions, including when the controller discloses personal data “to a processor that processes the personal data on behalf of the controller; . . . to a third party for purposes of providing a product or service requested by the consumer; . . . to an affiliate of the controller; . . . “ when the data is publicly available information because the data subject made the personal data “available to the general public via a channel of mass media and . . . did not restrict to a specific audience,” and when the disclosure or transfer is made as “part of a merger, acquisition, bankruptcy, or other transaction in which the third party assumes control of all or part of the controller’s assets.”

The Colorado and Connecticut laws essentially track the Virginia language, except that they also specifically exclude the disclosure of personal data that a consumer has intentionally directed the controller to disclose or when the consumer directs the controller to “interact with a third party.” Utah law tracks the Colorado and Connecticut language but adds an additional exception for circumstances in which, “considering the context in which the consumer provided the personal data to the controller, a controller’s disclosure of personal data to a third party . . . is consistent with a consumer’s reasonable expectations.”

Overall, although we see much consistency in how these laws define processing, we do see a general trend among the enacted statutes toward adding exceptions and diluting the scope of what is included within the definition of “sale” — the only type of processing specifically defined by all of the U.S. statutes — with Utah law clearly the most limited in this respect, especially because of its additional exception wherein a controller can circumvent the “sale” of personal data based whenever an operation is adjudged to be “consistent

345 Id.
346 See COLO. REV. STAT. § 6-1-1303(23)(b)(V)(A) (2023) (“Sale,” ‘sell,’ or ‘sold’ does not include the following: . . . [t]he disclosure of personal data: That a consumer directs the controller to disclose or intentionally discloses by using the controller to interact with a third party.”); 2022 Conn. Pub. Acts 22-15 (sSB Sess.) § 1(26)(D) (“Sale of personal data does not include . . . the disclosure of personal data where the consumer directs the controller to disclose the personal data or intentionally uses the controller to interact with a third party.”). 
with a consumer’s reasonable expectations.” Additionally, since Colorado’s law provides only a finite list of operations within the scope of processing, its scope is also significantly limited in comparison to the others. As such, we see a substantial difference between the GDPR and (to a somewhat lesser extent) the CCPA, on the one hand, and at least several of the Virginia Model Statutes, on the other hand.

c. Covered Entities: Controllers and Processors

In this part, we examine which entities are obligated to comply with the law. Following a discussion of the law in each state, we provide a comparative analysis of the similarities and differences between these laws and examine how the examined provisions compare to relevant aspects of the GDPR.

In California, the CCPA applies to for-profit businesses (and their subsidiaries or parent companies) that collect “consumers’ personal information” when the “business” participates in determining the “purposes and means of processing” of such information; does business in California; and either (A) had annual gross revenues of more than twenty-five million dollars in the previous calendar year, (B) “annually buys, sells, or shares the personal information of 100,000 or more consumers or households [or] (C) [d]erives 50[%] or more of its annual revenues from selling

---

350 Id. It is also not clear who is tasked with making this determination, and what deference a decisionmaker should give to a controller’s own judgment on this issue.

351 See Vanessa Perumal, The Future of U.S. Data Privacy: Lessons from the GDPR and State Legislation, 12 NOTRE DAME J. INT’L & COMP. L. 99, 116 (2022) (attempting to place some state data privacy laws on a spectrum alongside the GDPR, from most to least expansive in terms of regulation, in large part in reliance on which entities are obligated as controllers to respect data subjects’ rights). The analysis situates the GDPR as the most expansive, with various U.S. state laws situated along a spectrum towards least expansive, with the CCPA placed as more expansive than the VaCDPA.

352 See CAL. CIV. CODE § 1798.140(d)(1) (2022) (defining “business” 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” that also meets the other requirements set out in the text).

353 See CIV. § 1798.140(d)(2) (“Any entity that controls or is controlled by a [covered] business . . . and that shares common branding with the business and with whom the business shares consumers’ personal information.”).
or sharing consumers’ personal information.”\textsuperscript{354} The CCPA also applies to certain joint ventures and partnerships,\textsuperscript{355} as well as other persons who do business in California.\textsuperscript{356} Businesses “collect” personal information whenever they buy, rent, gather, obtain, receive, or access “any personal information pertaining to a consumer by any means.”\textsuperscript{357}

Notably, beyond the exemptions for certain types of data processing noted earlier,\textsuperscript{358} the CCPA expressly exempts a variety of entities from its regulations, including non-profit organizations, financial institutions regulated under the GLBA, consumer reporting agencies regulated under the FCRA, medical information and health care providers dealing with information regulated by various state and federal medical privacy laws (including HIPAA).\textsuperscript{359} Additionally, CCPA-covered businesses are exempt from liability when they share or provide personal information in compliance with federal, state, or local laws, while under court order, subpoena,\textsuperscript{360} or other inquiry by federal, state, or local authorities.\textsuperscript{361} The CCPA also allows for preservation requests from law enforcement agencies,\textsuperscript{362} absolves businesses from liability for cooperating with law enforcement in certain circumstances,\textsuperscript{363} and includes an exemption for providing government access to personal information in exigent circumstances.\textsuperscript{364}

\textsuperscript{354} CIV. § 1798.140(d).
\textsuperscript{355} See CIV. § 1798.140(d)(3). ("‘Business’ means: . . . [a] joint venture or partnership composed of businesses in which each business has at least a 40[ %] interest. For purposes of this title, the joint venture or partnership and each business that composes the joint venture or partnership shall separately be considered a single business, except that personal information in the possession of each business and disclosed to the joint venture or partnership shall not be shared with the other business.”).
\textsuperscript{356} See CIV. § 1798.140(d)(4) (2022). ("‘Business’ means: . . . [a] person that does business in California, that is not covered by paragraph (1), (2), or (3) and that voluntarily certifies to the California Privacy Protection Agency that it is in compliance with, and agrees to be bound by, this title.”).
\textsuperscript{357} CIV. § 1798.140(t).
\textsuperscript{358} See infra Part IV(A)(5)(c).
\textsuperscript{359} CIV. § 1798.145(c)-(e) (2021).
\textsuperscript{360} See CIV. § 1798.145(a)(1) ("The obligations imposed on businesses by this title shall not restrict a business’ ability to: (1) Comply with federal, state, or local laws or comply with a court order or subpoena to provide information.”).
\textsuperscript{361} CIV. § 1798.145(a)(2).
\textsuperscript{362} Id.
\textsuperscript{363} CIV. § 1798.145(a)(3).
\textsuperscript{364} CIV. § 1798.145(a)(4).
The VaCDPA applies to “persons” who conduct business within Virginia or who “produce products or services that are targeted to residents” of Virginia as long as those “persons” either “control or process personal data of at least 100,000 consumers or (ii) control or process personal data of at least 25,000 consumers and derive over 50% of gross revenue from the sale of personal data.”365 The term “person” is not defined in the statute, which in various places refers to entities as either a “natural person” or “natural or legal person.”366 Presumably, “person[s]” in this context refers to either legal or natural persons, as a “controller” is defined to include both.367 Similarly, the statute does not define what it means to “control” personal data—although a reasonable reading of the statute could be that to control personal data is to act as a “controller.” A controller “means the natural or legal person that, alone or jointly with others, determines the purpose and means of processing personal data.”368 Virginia law exempts state and local (public) entities within the Commonwealth, financial institutions subject to the GLBA, entities subject to HIPAA, nonprofit organizations, and post-secondary institutions of higher education from obligations under the law.369

Colorado law resembles the VaCDPA in many respects. It creates legal obligations for any “controller” that “[c]onducts business in Colorado or produces or delivers commercial products or services that are intentionally targeted to residents of Colorado”370 as long as the controller:

(I) Controls or processes the personal data of one hundred thousand consumers or more during a calendar year; or

366 See § 59.1-575.
367 Id.
368 Id.
369 See § 59.1-576(B) (“This chapter shall not apply to any (i) body, authority, board, bureau, commission, district, or agency of the Commonwealth or of any political subdivision of the Commonwealth; (ii) financial institution or data subject to Title V of the federal Gramm-Leach-Bliley Act (15 U.S.C. § 6801 et seq.); (iii) covered entity or business associate governed by the privacy, security, and breach notification rules issued by the U.S. Department of Health and Human Services, 45 C.F.R. Parts 160 and 164 established pursuant to HIPAA, and the Health Information Technology for Economic and Clinical Health Act (P.L. 111-5); (iv) nonprofit organization; or (v) institution of higher education.”).
370 COLO. REV. STAT. § 6-1-1304(1)(a) (2023).
(II) Derives revenue or receives a discount on the price of goods or services from the sale of personal data and processes or controls the personal data of twenty-five thousand consumers or more.\textsuperscript{371}

In language almost identical to that in the VaCDPA, a “controller” is any “person that, alone or jointly with others, determines the purposes for and means of processing personal data.”\textsuperscript{372} As with the VaCDPA, CoPA does not define “person” but presumably this term means any natural or legal person. Similarly, “controls” presumably means to act as a “controller,” but this is not explicit in the text. Likewise, what it means to conduct business in the state or intentionally target state residents is not defined. Like the CCPA and VaCDPA, CoPA contains several exemptions for specific institutions, including for financial institutions subject to GLBA;\textsuperscript{373} national securities associations registered under the federal Securities Exchange Act;\textsuperscript{374} and covered entities, business associates, and qualified service organizations subject to federal health-related legislation (including HIPAA).\textsuperscript{375} CoPA also exempts public entities\textsuperscript{376} and public utilities,\textsuperscript{377} except as otherwise allowed by state or federal law. However, CoPA’s list of expressly exempted entities includes air carriers\textsuperscript{378} (as defined by federal law), a provision not seen in the CCPA or VaCDPA. CoPA also enumerates

\textsuperscript{371} Id. § 6-1-1304(1)(b)(I)-(II).
\textsuperscript{372} Id. § 6-1-1303(7).
\textsuperscript{373} See id. § 6-1-1304(2)(q) (“This part 13 does not apply to: . . . a financial institution or an affiliate of a financial institution as defined by and that is subject to the federal “Gramm-Leach-Bliley Act”, 15 U.S.C. sec. 6801 et seq., as amended, and implementing regulations, including Regulation P, 12 CFR 1016.”).
\textsuperscript{374} See id. § 6-1-1304(2)(m) (“This part 13 does not apply to: . . . a national securities association registered pursuant to the federal “Securities Exchange Act of 1934”, 15 U.S.C. sec. 78o-3, as amended, or implementing regulations.”).
\textsuperscript{375} See id. § 6-1-1304(2)(h) (“This part 13 does not apply to: . . . information maintained in the same manner as information under subsections (2)(a) to (2)(g) of this section by: (I) A covered entity or business associate; (II) A health-care facility or health-care provider; or (III) A program of a qualified service organization as defined in 42 CFR 2.11.”).
\textsuperscript{376} See id. §§ 6-1-1304(2)(n)-(o).
\textsuperscript{377} See id. § 6-1-1304(2)(n) (“This part 13 does not apply to: . . . customer data maintained by a public utility as defined in section 40-1-103 (1)(a)(I) or an authority as defined in section 43-4-503 (1), if the data are not collected, maintained, disclosed, sold, communicated, or used except as authorized by state and federal law.”).
\textsuperscript{378} See id. § 6-1-1304(2)(l) (“This part 13 does not apply to: . . . an air carrier as defined in and regulated under 49 U.S.C. sec. 40101 et seq., as amended, and 49 U.S.C. sec. 41713, as amended.”).
a long list of exemptions to controllers’ obligations, for example, when processing is required to comply with legal orders, cooperate with law enforcement, pursue internal research and development, or mitigate security breaches.\textsuperscript{379}

The Utah law applies to “any controller or processor” who (1) conducts business in Utah or “produces a product or service that is targeted to consumers who are residents of the state”\textsuperscript{380}; (2) has an annual revenue of at least $25 million; and (3) meets either of the following thresholds:

(i) during a calendar year, controls or processes personal data of 100,000 or more consumers; or (ii) derives over 50% of the entity’s gross revenue from the sale of personal data and controls or processes personal data of 25,000 or more consumers.\textsuperscript{381}

The UtCPA defines a “controller” as any “person doing business in the state who determines the purposes for which and the means by which personal data is processed, regardless of whether the person makes the determination alone or with others.”\textsuperscript{382} A “’[p]rocessor’ means a person who processes personal data on behalf of a controller.”\textsuperscript{383} Utah’s law contains express exemptions for governmental entities or third parties under contract and acting on behalf of a government entity,\textsuperscript{384} tribes,\textsuperscript{385} higher education institutions,\textsuperscript{386} nonprofit organizations,\textsuperscript{387} covered entities and business associates under HIPAA,\textsuperscript{388} consumer reporting agencies

\textsuperscript{379} See COLO. REV. STAT. § 6-1-1304(3) (2023).
\textsuperscript{380} UTAH CODE ANN. § 13-61-102(1)(a) (2023).
\textsuperscript{381} UTAH CODE § 13-61-102(1)(c)(i)-(ii).
\textsuperscript{382} Id. § 13-61-101(12).
\textsuperscript{383} Id. § 13-61-101(26).
\textsuperscript{384} See id. § 13-61-102(2)(a) (“This chapter does not apply to: (a) a governmental entity or a third party under contract with a governmental entity when the third party is acting on behalf of the governmental entity.”).
\textsuperscript{385} See id. § 13-61-102(2)(b).
\textsuperscript{386} See id. § 13-61-102(2)(c).
\textsuperscript{387} See id. § 13-61-102(2)(d).
\textsuperscript{388} See id. §§ 13-61-102(2)(e)-(f).
subject to the FCRA, financial institutions subject to GLBA, and air carriers.

In Connecticut, the CtDPA includes provisions much like those in the other states, with the law applying to “persons” who conduct business in—or who produce products or services that are targeted to residents of—Connecticut, and which also meet either of the following provisions during the preceding calendar year:

1. Controlled or processed the personal data of not less than one hundred thousand consumers, excluding personal data controlled or processed solely for the purpose of completing a payment transaction; or
2. Controlled or processed the personal data of not less than twenty-five thousand consumers and derived more than twenty-five per cent of their gross revenue from the sale of personal data.

Likewise, CtDPA defines a “controller” as “an individual who, or legal entity that, alone or jointly with others determines the purpose and means of processing personal data,” and a “processor” as “an individual who, or legal entity that, processes personal data on behalf of a controller.” Likewise, the Connecticut law also exempts state government agencies, nonprofit organizations, higher education institutions, national securities associations registered under the federal Securities Exchange Act, financial institutions subject to GLBA, and covered entities and business associates under HIPAA.

---

389 See id. § 13-61-102(2)(j).
390 See id. § 13-61-102(2)(k) (“This chapter does not apply to . . . (k) a financial institution or an affiliate of a financial institution governed by, or personal data collected, processed, sold, or disclosed in accordance with, Title V of the Gramm-Leach-Bliley Act, 15 U.S.C. Sec. 6801 et seq., and related regulations.”).
391 See id. § 13-61-102(2)(q).
393 Id. § 1(8).
394 Id. § 1(21).
395 Id. § 3(a)(1).
396 Id. § 3(a)(2).
397 Id. § 3(a)(3).
398 Id. § 3(a)(4).
399 Id. § 3(a)(5).
400 Id. § 3(a)(6).
In contrast to the U.S. provisions that limit their application to certain businesses and provide several categorical exemptions, the GDPR applies more broadly, defining controllers and processors to include a much wider range of agents. Indeed, “the GDPR covers anyone that processes personal data, including not only companies but also individuals, non-profit organizations, and governments.”

As such, the GDPR is not restricted to the “complex set of overlapping requirements” established under the CCPA and retained in the CCPA and Virginia Model Statutes. As Chander et al. have noted, this difference “reflects the dominant approach on each side of the Atlantic[,] wherein a data protection model inherently aims to be comprehensive [while U.S. law] still limits its aim to protecting consumers from certain data handling practices within a specific context defined by commerciality, geography, and scale.”

The GDPR does contain some important exceptions, however. For example, the GDPR’s rules do not extend to processing activities that “fall outside the scope of [EU] law,” such as national security-related activities, or that relate to “common foreign and security policy,” or to data processed for law enforcement purposes by “competent authorities,” by EU institutions, or that would interfere with several specific EU directives that also encompass data privacy concerns.

One of the most cited exceptions to the GDPR’s reach is the “personal or household activity” exception explained in Recital 18, which reads, in part, “[t]his Regulation does not apply to the processing of personal data by a natural person in the course of a purely personal or household activity and thus with no connection

---

401 Chander et al., supra note 5, at 1758 (citing GDPR, supra note 16, art. 2(1)).
402 Id.
403 Id. at 1758-59.
404 Herke Kranenborg, Article 2. Material Scope, in THE EU GENERAL DATA PROTECTION REGULATION (GDPR): A COMMENTARY 60, 60 (Christopher Kuner et al. eds., 2020) (noting that this exclusion may be limited in some circumstances, namely those in which the personal data was initially collected by private companies for commercial purposes) (citing GDPR, supra note 16, art. 2(2)(a)).
405 Id.; GDPR, supra note 16, Recital 16.
406 Kranenborg, supra note 404, at 70 (citing GDPR, supra note 16, art. 2(2)(d)).
407 Id. (citing GDPR, supra note 16, art. 2(2)(b)).
408 Id. at 71.
409 Id. at 71-72.
410 GDPR, supra note 16, art. 2(2)(c).
to a professional or commercial activity.” 411 This provision provides a limit on who can be defined as a controller or processor, protecting certain basic, personal data practices from potential liability. The recital states that the exemption covers, at the very least, “correspondence and the holding of addresses, or social networking and online activity undertaken within the context of such activities.” 412 Following decisions by the CJEU, the exclusion “must be interpreted as covering only activities that are carried out in the context of the private or family life of individuals.” 413 This means that even some personal use of social networking websites involving the processing of personal data could fall outside this exclusion. Thus, the exclusion may not apply when, for example, the personal data is posted publicly for “an unrestricted number or people,” when posted in collaboration with an association or company, or when the data processing extends beyond purely personal or household space—for example, the CJEU has held that using a video surveillance camera to record private property surrounding a home for security purposes constituted a regulated form of processing when the view of the camera extended beyond the private property. 414 As such, the exclusion is quite limited in practice.

IV. COMPARING THE LAW IN VARIOUS CONTEXTS

In this section, we briefly compare and analyze the material scope of these laws in reference to four factual scenarios illustrating some of the most common contexts in which individuals are affected by data in the modern digital society. We use these scenarios to provide additional clarity in comparing the material scope of these laws, placing the application of the cited provisions into a more practical context, using practical examples to illustrate differences

411 GDPR, supra note 16, Recital 18. Notably, several of the U.S. statutes contain similar exceptions, but they are unlikely to do much work due to the way in which the statutes define their regulated entities. See, e.g., VA. CODE § 59.1-582(E) (“Nothing in this chapter shall be construed as an obligation imposed on controllers and processors that . . . applies to the processing of personal data by a person in the course of a purely personal or household activity.”); COLO. REV. STATS. § 6-1-1304(3)(e); UTAH CODE § 13-61-102(2)(p); CONN. GEN. STAT. § 10(e)(2).

412 GDPR, supra note 16, Recital 18.

413 Kranenborg, supra note 409, at 68.

414 Id. (citing Case C-212/13, Ryneš v. Úřad pro ochranu osobních údajů, ECLI:EU:C:2014:2428, ¶ 33 (Dec. 11, 2014)).
between the scope of the legal protection provided by the GDPR and U.S. laws. In the paragraphs that follow, we briefly describe the facts of each scenario and then compare and contrast the application of the enacted U.S. laws and the GDPR. We use these scenarios to exemplify the scope of what constitutes personal data and the processing of personal data, and what differences in material scope exist between the various laws subject to our analysis.

a. Uber and Its Drivers: Automated Employment Decisions

Since at least 2018, Uber has been accused of firing (or rescinding access to its service for) its ride-share drivers based on automated analysis of data available to Uber about the drivers’ activities on the platform. The analysis supporting the suspensions reportedly revealed “security concerns related to account sharing” or detection of multiple log-in attempts at different locations, or suspicion of “fraudulent activities.” Based on an analysis of Article 15 and 22 of the GDPR, the Amsterdam Court of Appeal handed down a judgment in favor of the drivers because it determined the actions taken by Uber constituted decisions based solely on automated processing, including profiling and the drivers had a right to be informed about the fact of such decision-making and receive all information useful for challenging such a decision. Given the broad scope of what constitutes personal data under the GDPR, the data about drivers’ accounts and activity on the Uber app clearly constitute personal data, and Uber clearly is obligated to comply with the GDPR as it relates to drivers and users based within the EU.

415 See Mary-Ann Russon, Uber Sued by Drivers over “Automated Robo-Firing”, BBC News (Oct. 26, 2020), https://www.bbc.com/news/business-54698858 (The App Drivers & Couriers Union (ADCU), which is bringing the legal challenge, says that since 2018, it has seen well over 1,000 individual cases where drivers have allegedly been wrongly accused of fraudulent activity and immediately had their accounts terminated without a right of appeal.); see also Sarah Butler, Court Tells Uber To Reinstate Five UK Drivers Sacked by Automated Process, GUARDIAN, (Apr. 14, 2021), https://www.theguardian.com/technology/2021/apr/14/court-tells-uber-to-reinstate-five-uk-drivers-sacked-by-automated-process [https://perma.cc/D3CE-FWAA].

416 Butler, supra note 420.

417 Russon, supra note 420.

However, whether U.S. data privacy laws examined above apply to this situation is dictated by the material scope and exceptions to each law. Perhaps surprisingly, U.S. state laws examined here would generally exempt Uber’s processing in this case. This is so, primarily, because the information about Uber’s drivers can be framed as data about employees, or as being processed within a commercial or employment context. As we saw earlier, these laws all exclude from their scope data collected or otherwise processed about “a natural person acting in a commercial or employment context.” Thus, regardless of whether the law restricts decision-making based purely on automated data processing, data about a person acting in an employment-related context falls outside the scope of regulation. This is true even when the data at issue are linkable to an identified natural person and the controller meets the threshold requirements for acquiring obligations under the law.

b. Paint Chips as Personal Data: The Scope of Personal Data

On May 31, 1970, Eugene Edwards was arrested and charged with attempting to break into a post office in Lebanon, Ohio. While he was being placed in a cell at the local jail, investigators discovered that the attempted entry to the post office had been accomplished through the use of a pry bar on a wooden window, which apparently had caused paint chips to be scattered along the windowsill and screen. The next morning, the police purchased clothing for Edwards and then booked his old clothing, which contained chips of paint, into evidence. Examination and analysis confirmed that the paint chips on Edwards’s clothing matched those from the post office window, and this finding was instrumental in his prosecution for attempted breaking and entering. The Supreme Court acknowledged that “the articles of clothing

---

419 VA. CODE § 59.1-575; see also CAL. CIV. CODE § 1798.145(m)(1)(A); COLO. REV. STATS. § 6-1-1303(6); UTAH CODE § 13-61-101(10); CONN. GEN. STAT. § 1(7).
421 Id. at 801-02.
422 Id. at 802.
423 Id.
[Edwards] wore were themselves material evidence of the crime for which he had been arrested.”

Of course, all the laws examined here exclude, or do not cover, data processed by law enforcement for law enforcement purposes. Thus, we need to adjust the factual scenario a bit to make the comparison more relevant and interesting for present purposes. As such, let us assume that the entity processing the paint chips in this case is a business covered by all the laws at issue rather than a police agency. If we make that adjustment, the question posed is one that gets to the heart of the subject matter of these laws—that is, whether something like the paint chips themselves can be characterized as personal data, thus making their processing subject to regulation.

Under the GDPR, as discussed earlier, any data “relating to” an identified natural person constitutes personal data. Scholars have argued that the material scope of the GDPR is so broad that it may well encompass the processing of any and all data. More concretely, Purtova has argued that the current definition and technological context suggest that personal data encompasses an exceedingly wide range of data, including computer code and weather data. Specifically, Purtova considers a case of weather in the context of a living lab run in the Dutch city of Eindhoven where weather data, among others, is used for predictive policing. Depending on one’s theoretical inclinations regarding what information is, either weather itself or information about the weather, is information as meant in the GDPR’s definition of personal data (“any information regardless of its nature, content or format”). This information relates to the inhabitants of the city by reason of purpose, because its processing is meant to aid in predictive policing and public order and hence its purpose is “to

424 Id. at 805.
425 GDPR, supra note 16, art. 4(1).
426 Purtova, The Law of Everything, supra note 97, at 45 (“The concept ‘personal data’ [has the] potential to become an all-encompassing notion.”); Purtova & Leenes, supra note 155, at 250.
427 Purtova & Leenes, supra note 155, at 250.
428 Purtova, The Law of Everything, supra note 97, at 57-59 (“Weather, while neutral and of no consequence as a subject of small talk, may be considered information relating to an identified or identifiable natural person in the context of a living lab, for instance . . . .”).
evaluate, treat in a certain way or influence the status or behavior of the city inhabitants (e.g., to prevent them from acting in a criminal way). Purtova also notes that, in the context of a rich database to which weather information is added, advanced data analytics techniques and the complex web of public and private partners involved in the living lab project, including commercial parties and law enforcement, not all the factual impact of the data processing may be intended. Hence, this information may relate to people in impact or result. Finally, it might be argued that the city’s inhabitants within the ambit of the living lab are likely to be identifiable at some point if not by the weather information alone, certainly in combination with the rich data available via the living lab. Thus, in at least certain imaginable contexts, Purtova argues, even weather data may become personal data.

Moving to the scenario from United States v. Edwards, as in the case of weather data, either the chips themselves or the information inferred from their analysis is “any information” in the sense encompassed by the GDPR. As we saw earlier, under the GDPR, “information can ‘relate’ to an individual in content, purpose, or result.” The paint chips removed from Edwards’s clothing and the results of their analysis may not have been semantically about Edwards, but they are clearly either linked to him in purpose or result—that is, given that they are processed in the context of an investigation and then used in a trial aimed at establishing Edwards’s guilt, the purpose of their use is “to evaluate, treat in a certain way or influence the status or behavior of Edwards” and/or their “use is likely to have an impact on Edwards’ rights and interests.” Finally, Edwards—the data subject to whom information relates—is identified (his civil identity is known). Thus, the paint chips would be personal data under the GDPR.

Turning to the CCPA, the law defines “personal information” as “information 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.” Based on the plain meaning of the words used, it seems reasonable to conclude that the paint chips either relate to Edwards, are reasonably capable of being associated with Edwards, or could reasonably be linked to Edwards in the scenario provided. Indeed, the various courts involved in Edwards’s lawsuit concluded that the paint chips were evidence of his presence at the windowsill and thus of his probable guilt—they had, in fact, linked the paint chips to Edwards and associated them with him and his actions. After all, “associated” means “related, connected, or combined together,” where “linkable” means “capable of being linked,” and “relates” means “to show or establish logical or causal connection between.” Any of these seem sufficient to categorize the paint chips as personal information under the CCPA within this context.

The Virginia Model Statutes all require data to be “linked or reasonably linkable” to an identifiable natural person. Whether this will be interpreted to include a context like that imagined here is yet to be seen, but an argument could be made, just like as with the CCPA, that the paint chips as data might be “linkable” to Edwards. Additionally, all this analysis assumes “data” and “information” will be interpreted by courts and other bodies within each jurisdiction to include things like paint chips.

c. Posting on a Website: Individuals Posting Personal Data on the Internet

Does referencing someone by name on the internet constitute the processing of personal data? In 1998, Bodil Lindqvist created a webpage while completing a course on data processing. Part of the site, which was focused on providing information related to her

---

436 CAL. CIV. CODE § 1798.140(v)(1).
440 VA. CODE § 59.1-575.
church in Sweden, included information that would “allow parishioners preparing for their confirmation to obtain information they might need.” However, she also published pages that detailed, “in a mildly humorous manner, the jobs held by her colleagues and their hobbies. In many cases family circumstances and telephone numbers and other matters were mentioned. She also stated that one colleague had injured her foot and was on half-time on medical grounds.”

After some of her parish colleagues objected to this information being posted online, Lindqvist removed the pages. Subsequently, she was prosecuted for criminal violation of the Swedish data protection law for unlawfully processing personal data. The CJEU held that she had, in fact, processed personal data without legitimate grounds and that no exceptions to the law absolved Mrs. Lindqvist of liability, including the exclusion for data processed purely in a personal or household context. The court also held that such a finding was not inherently at odds with the right to freedom of expression.

We see in this example an example of processing by an individual that moves beyond a “purely personal and household activity.” Even though the case was decided under the 1995 Data Protection Directive which is no longer in force following the enactment of the GDPR, the law retained this general exception in Article 2(2)(c) and Recital 18. However, unsurprisingly, Bodil Lindqvist’s conduct would not be subject to regulation under any of the U.S. data privacy laws, primarily since an individual is generally not subject to the laws unless they meet the threshold requirements as a covered business. Thus, Lindqvist is simply not the sort of entity to which the laws are designed to capture, highlighting the differences between the consumer protection approach of U.S. laws and the fundamental rights approach to protecting data privacy under EU law.

442 Id. ¶ 13.
443 Id. ¶ 14.
444 Id. ¶ 15.
445 Id. ¶ 27.
446 Id. ¶ 100.
447 See supra note 416 and accompanying texts.
After the Supreme Court decided Dobbs, the New York Times declared that we had entered “the post-Roe era of digital privacy” — a “moment” the authors argued, “that underscores how the use of technology has made it practically impossible for Americans to evade ubiquitous tracking.” After Dobbs, privacy experts argued that “call histories, text messages and emails [as well as] ... location data, online payment records, Google searches and fertility tracking apps” would all be used to investigate and prosecute women seeking abortions or information about abortion procedures in states that had instituted bans and criminalized seeking or providing abortions. These concerns focus on three primary scenarios in which law enforcement may get access to such data: (1) through their own surveillance; (2) by using legal processes to acquire the data collected by private companies; or (3) by purchasing the information directly from businesses (including data aggregators).

Collection and other forms of processing personal data by law enforcement are not directly covered by any of the data protection or data privacy laws examined in this Article. However, the Law Enforcement Directive, which complements the GDPR in the EU does cover data processing in the context of law enforcement’s own surveillance (in reference to scenario 1, above), as long as the processing is done by the competent authorities “for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties.” Both the GDPR and the LED will apply to acquiring the data collected by private companies through using legal processes (scenario 2) and through purchasing the information directly from businesses (scenario 3), and the GDPR applies fully to private parties sharing the data.
Without diving into the specifics of the interoperability between the GDPR and the LED in cases of public-private legally mandated or voluntary cooperation, the question remains whether, and to what extent, U.S. data privacy laws may, even indirectly, limit or regulate law enforcement access to personal data. Of course, as noted earlier, certain communication-related records are subject to procedural protections under federal laws like the Stored Communications Act, which provides several legal processes for obtaining such records.

However, our analysis here is directed squarely at the so-called comprehensive consumer data privacy laws adopted in California, Virginia, Colorado, Utah, and Connecticut. In short, we see that these laws do little, if anything to directly limit law enforcement access to personal data. Some indirect effects may occur, of course, if regulated businesses do stop collecting or maintaining as much personal data as they have in the past, as there will then be more limited information to share with law enforcement. However, these laws do not apply directly to law enforcement agencies—that is, law enforcement agencies are not regulated as controllers or processors. This means that these laws, in any case, do not restrain law enforcement authorities from collecting data on their own (scenario 1). Further, the laws generally either require regulated businesses to share data with law enforcement agencies or are agnostic towards such data sharing, thereby facilitating scenarios 2 and 3.

For example, U.S. laws exempt regulated businesses from liability for complying with legal requests for data (facilitating data sharing and acquisition by law enforcement under scenario 2), empower law enforcement agencies to force businesses to retain information that may be relevant to an on-going investigation (also facilitating scenario 2), and require businesses to cooperate with law enforcement when they “reasonably and in good faith” believe that such cooperation will assist law enforcement in

452 See discussion supra Introduction.
454 See, e.g., CAL. CIV. CODE §§ 1798.145(a)(1)-(2); VA. CODE §§ 59.1-582(A)(1)-(2); COLO. REV. STATS. §§ 6-1-1304(3)(a)(I)-(II); UTAH CODE §§ 13-61-304(1)(a)-(b); CONN. GEN. STAT. §§ 10(a)(1)-(2).
455 CAL. CIV. CODE §§ 1798.145(a)(2). Businesses subject to the laws in Virginia, Colorado, Utah, and Connecticut may also be subject to data retention requests under general provisions requiring them to comply with “federal, state, or local law, rule[s], or regulation[s].” VA. CODE § 59.1-582(A)(1); see also COLO. REV. STATS. § 6-1-1304(3)(a)(I); UTAH CODE § 13-61-304(1)(a); CONN. GEN. STAT. § 10(a)(1).
investigating violations of the law\cite{CAL_CIV_CODE §§ 1798.145(a)(3); see also VA. CODE § 59.1-582(A)(3); COLO. REV. STATS. § 6-1-1304(3)(a)(III); UTAH CODE § 13-61-304(1)(c); CONN. GEN. STAT. § 10(a)(3).} or in emergency situations when a “person is at risk of danger of death or serious physical injury”\cite{CAL_CIV_CODE §§ 1798.145(a)(4); see also VA. CODE §§ 59.1-582(A)(6); CONN. GEN. STAT. § 10(a)(8).} (potentially facilitating, or at least not prohibiting, voluntary data sharing that would go beyond legal requests under scenario 2). The CCPA also allows businesses to sell deidentified or aggregate personal data to law enforcement (and others)\cite{CAL_CIV_CODE §§ 1798.145(a)(6).}, while the Virginia Model Statutes also allow such sales as long as they comply with certain requirements\cite{CAL_CIV_CODE §§ 1798.145(a)(4); see also VA. CODE §§ 59.1-582(A)(6); CONN. GEN. STAT. § 10(a)(8).} (related to scenario 3).

This analysis suggests that these new state laws do little or nothing to directly regulate law enforcement access to personal data collected and held by private companies. Absent broader regulatory schemes dedicated to this issue of law enforcement access to and use of personal data, whether collected on their own or acquired from private businesses, the U.S. legal landscape here remains relatively like what data privacy has historically been as compared to the EU: a largely unregulated space with only sectoral-like protections for things like information acquired through activities that meet the standard for Fourth Amendment searches or seizures\cite{Regulated by the Fourth Amendment to the US Constitution. U.S. CONST. amend. IV.}, accessing stored communications records\cite{18 U.S.C. § 2701, et seq.}, and intercepting live communications\cite{18 U.S.C. § 2510, et seq.}.

V. CONCLUSION

As Richards and Hartzog noted in 2021, “[i]t wasn’t supposed to be like this.”\cite{Richards & Hartzog, supra note 2, at 963.} While “the internet promised human empowerment,” what we have now is a massive combination of technology and industry that “manufacture consent to boilerplate terms that no one reads” and a society flooded with “corporate
promises of empowerment [that] have instead delivered manipulation, disempowerment, and distrust.” Importantly, state-based data privacy laws examined in this Article have moved the needle forward with regard to protecting information privacy in the United States. Prior to enactment of the CCPA in 2018, commercial data processing was virtually unregulated in the U.S.—a wild west of data brokers and commercial data market that trafficked in billions of dollars of personal data. These laws have also spurred real legislative movement in the U.S. Congress, as federal legislators and corporate lobbyists seek a unified regulatory framework applicable across the entire country.

However, the material scope of these laws is also quite limited in various ways, as discussed earlier. These differences stem from, among other reasons, the fact that “[d]ata protection laws like the GDPR proceed from the principle that data protection is a fundamental human right safeguarded through constitutional protections . . . . This places data protection rights on the same plane as free speech or due process.” Thus, in comparison to the unregulated, wild-west approach in the United States where data can be bought and sold for any reason absent some legal restriction, the default situation in the EU “is that personal information cannot be collected or processed unless there is a specific legal justification for doing so.” Indeed, according to Chander et al., the most meaningful difference between the GDPR and the CCPA is that the CCPA does not require that processing be lawful – that is, based on grounds for lawful processing and in accordance with general principles. Rather, it “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.”

464 Id. at 964.
466 Most importantly, through the bipartisan support garnered for the ADPPA in 2022.
467 Chander et al., supra note 5, at 1747.
468 Id.
469 Id. at 1756.
470 Id.
Based on the distinction made by Chander et al. between data protection laws—where protection follows the data—and consumer-protection-based data privacy laws, where protection follows a commercial relationship between businesses and consumers, we see that U.S. laws differ quite substantially in their scope as compared to the “gold standard,” or GDPR. While California law retains its consumer protection framing, it does resemble a broader protection that tracks the data itself to a larger extent than do the Virginia Model Statutes. However, even the CCPA is much more limited in its material scope and application than the GDPR, applying to a smaller range of regulated entities and containing a broader list of exceptions, thus contributing to why others have argued that it is “lacking major structural elements of the GDPR.”

Indeed, the CCPA and the Virginia Model Statutes contain lengthy lists of exemptions and defer to other, potentially weaker, state and federal laws in important contexts, thus limiting whether they can even be called “comprehensive” data privacy laws in the first place. If many types of health and medical data, financial data, and credit data, among other categories of otherwise personal data, fall outside the scope of these laws, they are really only “comprehensive” data privacy laws in the sense that a block of Swiss cheese is a solid block of cheese without holes—that is, they are not comprehensive data privacy laws at all. The Virginia Model Statutes are also clearly consumer privacy laws and not true data protection laws, given that both their protection and application track the consumer-business relationship. The deficiencies in these laws are also apparent in that the federal draft ADPPA would encompass a broader material scope, suggesting that political will exists to do more than what these laws encompass. As argued by Chander et al., “[i]f the GDPR is a doctoral thesis, the CCPA is a term paper written the night before the deadline.” We suggest that if such a

471 Id. at 1749.
473 See supra Part IV(C); see also Chander et al., supra note 5, at 1755 (“[W]hile the CCPA is broader than past American sectoral laws, it still regulates a much narrower set of entities than does the GDPR.”).
474 See supra Part IV(A)(5).
475 Chander et al., supra note 5, at 1746.
476 Id.
comparison is accurate, then the Virginia Model Statutes appear to be something more like an out-sourced term paper, bought and paid for by the legislature but written by an outside, financially incentivized author without an understanding or concern for the needs and requirements of a real data protection statute or of the rights embodied in the text.

In the end, however, whether these laws adequately capture and protect the underlying interests that appear to have motivated their adoption, such as privacy and the need to protect people from other data-driven harms, may become more clear after the laws go fully into effect and the scope of their terms—including, most importantly, the linking terms used to define their subject matter—is interpreted by courts and data protection agencies (where they do now exist or may exist in the future). The First Amendment clearly plays a role in limiting the potential scope of certain aspects of these laws compared to the GDPR, requiring, at a minimum, some exceptions for publicly accessible information held by government institutions. But the consumer protection framing and some (potential) limits to the material scope of what constitutes personal data in these laws are likely more limiting than required by the First Amendment. Besides their provisions requiring corporate compliance and cooperation with law enforcement requests that potentially go beyond what other laws require, the mere fact that these states have not enacted any similar data privacy legislation applicable to law enforcement agencies and the criminal investigations context highlights a shortcoming in relation to the GDPR/LED pairing in Europe.

Finally, we examine to what extent the material scope of these laws—including how they protect privacy interests and limit corporate power—contributes to promoting neorepublican notions of liberty, non-domination, and antipower. Domination, as we framed it earlier in reliance on Pettit’s work, exists when “uncontrolled interference” is possible—that is, when it “exercised at the will or discretion of the interferer; interference that is uncontrolled by the person on the receiving end.”477 Interference, or the potential for interference, with an individual’s choices might be non-arbitrary in the sense that it “conforms to rules” (including the law), but it might still be “uncontrolled interference” when its application is not pursuant to nor aligned with the individual’s

477 Pettit, supra note 41, at 58.
As noted by Pettit, and quoted earlier, “[t]he active, intentional restriction of your choice by any other agent or agency will be invasive only to the extent that it reflects a will that you do not control.”

In our view, this perspective is largely consistent with Julie Cohen’s concept of surveillance as modulation, which she describes as, “a set of processes in which the quality and content of surveillant attention is continually modified according to the subject’s own behavior, sometimes in response to inputs from the subject but according to logics that ultimately are outside the subject’s control.”

Modern information technologies enable this mode of surveillance, pursuing logics that operate outside the will and control of individuals. Modulation is, in Cohen’s terms, “a mode of privacy invasion, but it is also a mode of knowledge production designed to produce a particular way of knowing and a mode of governance designed to produce a particular kind of subject.”

Indeed, a primary purpose of commercial data processing in the American data economy “is to produce tractable, predictable citizen-consumers whose preferred modes of self-determination play out along predictable and profit-generating trajectories... tailored to play to existing inclinations, nudging them in directions that align with profit-maximizing goals.” These logics and forms of consumer privacy invasion and manipulation permitted by existing U.S. law are examples of how networked information technologies have been deployed by private companies to produce modulation under a faulty neoliberal approach to regulating (or failing to meaningfully regulate) commercial data practices.

According to Cohen, this existing mode of “privacy governance is particularly ill-equipped to respond effectively to emerging practices of modulation. Its emphasis on privatized regulation and

\[478\] Id.
\[479\] Id.
\[480\] Cohen, supra note 43, at 1915.
\[481\] See id. (“Networked information technologies enable surveillance to become modulation. Surveillance may be defined generically as attention that is ‘purposeful, routine, systematic and focused.’ Networked information technologies enable surveillant attention to become continuous, pervasively distributed, and persistent.”).
\[482\] Id. at 1917.
\[483\] Id.
\[484\] Id. at 1930.
control of information flows via notice and choice reinforces precisely those aspects of modulation that are most troubling and most intractable.” Furthermore, Cohen posits that “[i]f privacy regulation is to provide effective shelter, . . . we will need to acknowledge that privacy is the opposite of modulation and can exist only to the extent that processes of modulation are gap-ridden, transparent, and incomplete.”

Building from the ideas of Pettit and Cohen, we construct the notion of *informatic domination*, which we define as a state of affairs made possible by the ability of an agent to use or process information relating to a natural person in an arbitrary or uncontrolled way and that could meaningfully affect or influence the behavioral choices or range of available choices available to an individual or group of individual persons. This also responds to what Richards and Hartzog have described as, “the risks of opportunism that arise when people trust others with their personal information and online experiences.” Pettit’s concept of antipower would be promoted when means of commanding non-interference are made available to these persons. Viewed in that light, data privacy laws, whether based in consumer protection or fundamental rights, can be analyzed and critiqued based on how much and how well they promote antipower and reduce the potential for informatic domination.

One way to do this, in the words of Cohen, is to institute “gaps and inconsistencies within systems of meaning” and surveillance—which can be read as regulating data processing practices. Similarly, Selinger and Hartzog argue that “obscurity” could mean keeping information “out of reach . . . or

485 See id.
486 Id.
487 Others have also used this term in various ways. See, e.g., Majid Tehranian, *Taming Capital, Holding Peace, Democratizing Global Governance* 28, 29 (Eşref Aksu & Joseph A. Camilleri eds., 2002) (“The hierarchical structure of the world is a continuing characteristic of *de facto* global governance. The global system has gone through three major interlocking phases of hegemonic development, evolving from agrarian to industrial and informatic domination.”); see also Christopher M. Kelty, *Too Much Democracy in All the Wrong Places: Toward a Grammar of Participation*, 58 *Current Anthropology* 77, 81 (2017) (“Participation is at once a liberating experience for those in poverty and a tool of informatic domination through monitoring.”).
488 Richards & Hartzog, *supra* note 2, at 966.
A complementary pursuit could be to make data privacy regulation work in support of aligning possible data practices with the will of the individual data subjects (or groups of individuals, where a common will and interests are present). This must be done in ways that go beyond the pre-existing and neoliberal “notice and choice” model of industry self-regulation dominant in the United States. These responses should include a variety of approaches and elements, including defaulting processing (including collection) to an opt-in status, reforming what real informed consent procedures looks like, enforcing substantive data minimization and purpose limitation rules, adopting duties of loyalty with corresponding obligations and remedies, and

490 Selinger & Hartzog, supra note 67, at 119 (stating “misunderstood” in the context of breaking syntactic links and maintaining gaps in information held and analyzed rather than reducing the effect of rights to rectify or correct bad information).

491 Notably, all the Virginia Model Statutes (except Utah) and the ADDPA include language that prohibits controllers from processing “sensitive” personal data in some circumstances. See American Data Privacy and Protection Act, H.R. 8152, 117th Cong. § 102(2) (2022) (“Notwithstanding section 101 and unless an exception applies, with respect to covered data, a covered entity or service provider may not . . . collect or process sensitive covered data, except where such collection or processing is strictly necessary to provide or maintain a specific product or service requested by the individual to whom the covered data pertains, or is strictly necessary to effect a purpose enumerated in paragraphs (1) through (12) and (14) through (15) of section 101(b).”); VA. CODE § 59.1-578(A)(5) (2022) (“A controller shall . . . not process sensitive data concerning a consumer without obtaining the consumer’s consent, or, in the case of the processing of sensitive data concerning a known child, without obtaining consent from the child’s parent or lawful guardian.”); 2022 Conn. Pub. Acts 22-15 (sSB Sess.) § 6(a)(4) (“A controller shall . . . not process sensitive data concerning a consumer without obtaining the consumer’s consent, or, in the case of the processing of sensitive data concerning a known child, without processing such data in accordance with COPPA.”).

492 See CAL. CIV. CODE § 1798.140(h) (2018) (identifying a positive example of this wherein any “agreement obtained through use of dark patterns does not constitute consent”). The CCPA’s definition of “consent” also states 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.” Id. However, the CCPA could be read as not doing any more than requiring the excision of “unrelated information” from another broad and difficult to understand legal document about commercial data practices.

493 For a good argument and explanation of this possible path to help redeem privacy law, see generally Richards & Hartzog, supra note 2 (presenting a theory of
allowing for meaningful private rights of action instead of restricting enforcement to states attorneys general or data protection agencies.\footnote{See Lauren Henry Scholz, Private Rights of Action in Privacy Law, 63 WM. & MARY L. REV. 1639, 1676 (2022) ("All of these proposals share an understanding that merely forcing companies to announce data breaches and subjecting them to public actions from state attorneys general is insufficient to prevent data insecurity.").}

Avoiding informatic domination and promoting antipower could serve as compelling “normative reasons”\footnote{See Richards & Hartzog, supra note 2, at 966 ("Any account of a duty of loyalty must offer normative reasons for having the duty in the first place, specifying the values served by imposing such a duty of loyalty on companies in the context of what we have elsewhere called ’information relationships.’")} for adopting, for example, a robust duty of loyalty that ties controllers’ decisions to the will and interests of the data subject, with ample opportunities...
for enforcement, including under a private right of action.\textsuperscript{496} Instructive here is Richards' and Hartzog's version (and vision) of a duty of loyalty, in which

loyalty would manifest itself primarily as a prohibition on designing digital tools and processing data in a way that conflicts with a trusting party's best interests. Data collectors bound by such a duty of loyalty would be obligated to act in the best interests of the people exposing their data and engaging in online experiences, but only to the extent of their exposure.\textsuperscript{497}

Importantly, this vision is for a "best interests" form of loyalty that would help rid "trusting consumers of the burdens of privacy self-management and other "privacy work.""\textsuperscript{498} If tied to substantive and accessible enforcement mechanisms, including private rights of action, this version of data loyalty obligations—one that works to ensure the data subject's will and interests are respected and that respects robust data minimization and purpose limitation principles (relative to "exposure")—could be seen as complementary to achieving "semantic discontinuity" and preserving the obscurity of additional personal information that need not be processed or otherwise brought to bear in enacting the best interests of the data subject. Adopting such principles more fully into data privacy law would reduce the potential for uncontrolled and arbitrary interference arising from the indiscriminate processing that so characterizes the modern American data economy.

In our final remarks, we focus on the big picture. The interlocking issues of data privacy, (commercial) data practices, liberty, and human flourishing combine to make these questions compelling and vitally important. Richards and Hartzog argue that the "digital transformation" of the past few decades was based on great promises about "human well-being and flourishing," but that these promises remain "unfulfilled" under the current legal regime.\textsuperscript{499} Relatedly, Cohen makes a compelling argument that

\textsuperscript{496} For an argument in favor of private rights of action in privacy law generally, see Scholz, \textit{supra} note 499.

\textsuperscript{497} Richards & Hartzog, \textit{supra} note 2, at 966. Notably, this sort of test would place greater power in courts to determine what is in an individual's best interest. Specifying specific criterion or general factors to guide judicial interpretation might also be warranted.

\textsuperscript{498} \textit{Id.} at 968.

\textsuperscript{499} \textit{Id.} at 967.
A society that permits the unchecked ascendancy of surveillance infrastructures cannot hope to remain a liberal democracy. Under such conditions, liberal democracy as a form of government is replaced, gradually but surely, by a different form of government that I will call modulated democracy. Citizens within modulated democracies—citizens who are subject to pervasively distributed surveillance and modulation by powerful commercial and political interests—increasingly will lack the ability to form and pursue meaningful agendas for human flourishing.\textsuperscript{500}

On that note, and considering that when even the strongest, most comprehensive U.S. data privacy law can be criticized for not being focused on “reinforcing good data practices or creating substantive protections for consumers,”\textsuperscript{501} we know there is still plenty of work to be done.

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{500} COHEN, supra note 43, at 1912.
\item \textsuperscript{501} Chander et al., supra note 5, at 1758.
\end{itemize}
\end{footnotesize}