THE YALE LAW JOURNAL FORUM

April 22, 2024

How to Get the Property Out of Privacy Law

Jane R. Bambauer

ABSTRACT. Privacy law emphasizes control over "your" data, but requiring consent for each data use is unprincipled, not to mention utterly impractical in the AI era. American lawmakers should reject the property model and use a framework that creates defined zones of privacy and clear safe harbors, irrespective of consent.

INTRODUCTION

In the United States, multiple attempts to pass an omnibus privacy law have faltered.[1] Explanations for these repeated failures usually home in on specific features of the proposals: a controversial preemption of state law, or a private right of action that was unacceptable to business-oriented legislators, for example.[2] These reasons are true in a sense; they accurately identify the portions of the bill that divide active stakeholders and break open political alliances. But there is also a deeper explanation—a latent tension between a property-based approach to privacy law and a torts-based one.

Property frameworks give people significant control over whether their data is collected and how it is used. Under this model, loss of control is a harm in itself (like loss of property), in addition to whatever downstream harms might also follow. By contrast, the torts framework manages risks related to activities. It assumes that nobody automatically has the right to exclude others from

[1] This is much to the chagrin of privacy scholars. *See, e.g.*, Priscilla M. Regan, *Fifty-Plus Years of Information Privacy Policy-Making: The More Things Change, the More They Remain the Same*, *in* Research Handbook on Information Policy 159, 159 (Alistair S. Duff ed., 2021).

[2] Qiuyang Zhao, *American Data Privacy and Protection Act: Latest, Closest, yet Still Fragile Attempt Toward Comprehensive Federal Privacy Legislation*, Jolt Dig. (Oct. 19, 2022), https://jolt.law.harvard.edu/digest/american-data-privacy-and-protection-act-latest-closest-yet-still-fragile-attempt-toward-comprehensive-federal-privacy-legislation [https://perma.cc/BAX2-97YG].

collecting, creating, or using information about them, but they *are* entitled to protection from unjustified risks and misuse of personal data that will foreseeably lead to physical, economic, and dignitary harms. The two frameworks have irreconcilable differences with respect to who decides how data will be collected or used, and how the decisions will be made. The conflict has smoldered and kept American lawmakers in paralysis.

Privacy advocates typically use the property/control framework,[3] and this has only increased over the last ten years under the influence of European and Californian privacy laws.[4] European law treats data as something that belongs to the people described in them. The data subjects have exclusive control over processing in most circumstances, just as individuals have a fundamental right to control access and use of their property.[5] Thus, under the European Union's General Data Protection Regulation (GDPR), any time a data controller wants

[3] Consider Alan Westin's formulation: "[P]ersonal information, thought of as the right of decision over one's private personality, should be defined as a property right, with all the restraints on interference by public or private authorities and due-process guarantees that our law of property has been so skillful in devising. Along with this concept should go the idea that circulation of personal information by someone other than the owner or his trusted agent is handling a dangerous commodity in interstate commerce, and creates special duties and liabilities on the information utility or government system handling it." ALAN F. WESTIN, PRIVACY AND FREEDOM 324-25 (1967). *See also* CHARLES FRIED, AN ANATOMY OF VALUES: PROBLEMS OF PERSONAL AND SOCIAL CHOICE 140 (1970) ("Privacy is not simply an absence of information about us in the minds of others; rather it is the control we have over information about ourselves."); ARTHUR R. MILLER, THE ASSAULT ON PRIVACY: COMPUTERS, DATA BANKS, AND DOSSIERS 25 (1971) ("[T]he basic attribute of an effective right of privacy is the individual's ability to control the circulation of information relating to him . . . .").

[4] For example, the White House Blueprint for an AI Bill of Rights says that "[d]esigners, developers, and deployers of automated systems should seek your permission and respect your decisions regarding collection, use, access, transfer, and deletion of your data," and admonishes companies against "burden[ing] users with defaults that are privacy invasive." Off. for Sci. & Tech. Pol'y, *Blueprint for an AI Bill of Rights*, WHITE HOUSE, https://www.whitehouse.gov/ostp/ai-bill-of-rights [https://perma.cc/Z9Y8-AADG].

[5] Regulation 2016/679 of Apr. 27, 2016, General Data Protection Regulation, 2016 O.J. (L 119) 37, art. 7 [hereinafter GDPR]; WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 29-30 (1987) (explaining that a property right confers to the owner "a substantial power to exclude others from the use and enjoyment" of the resource). While European regulators describe privacy as a fundamental right, this in itself does not determine whether those rights are protected through ownership or risk-management frames. For example, health and bodily integrity are also fundamental rights, but the health and safety of Europeans are protected from accidents using a risk-management frame rather than through exclusive control. *See Data Protection*, EUR. DATA PROT. SUPERVISOR, https://edps.europa.eu/data-protection/data-protection_en [https://perma.cc/7RFR-47EZ] ("In the EU, human dignity is recognised as an absolute fundamental right. In this notion of dignity, privacy or the right to a private life, to be autonomous, *in control of information about yourself*, to be let alone, plays a pivotal role.") (emphasis added).

to reuse data for a new purpose, it must seek the consent of the data subject.[6] Deviations from the property frame occur only to make it even *more* difficult for individuals to sell or give away control of personal data, lest they trade away their privacy too easily.[7] For example, the California Consumer Privacy Act (CCPA) and GDPR put limitations on what types of quid pro quo trades a data controller can offer in exchange for the consent of a data subject.[8] And even after permission is given, both legal regimes allow data subjects to renege and claw back their data under certain circumstances.[9] Thus, privacy law is attempting to make personal data an extra-sticky form of property.[10]

But the property frame, popular as it may be, is unworkable and unprincipled.[11] Consider how Europe's privacy law has already affected the region's approach to Artificial Intelligence (AI). Shortly after Open AI released ChatGPT to the general public, Italian privacy regulators forbade its access to the Italian

[6] Unless the new purpose falls within one of the narrow justifications under the GDPR that does not require data subject consent. GDPR, art. 6(1). The last category of allowable unconsented processing—category (f) for "legitimate interests" of the controller—is fairly narrow as interpreted by regulators and courts. The European Court of Justice has found, for example, that Meta's use of personal data to serve behavioral advertisements on its platforms does not quality as a legitimate interest. Case C-252/21, Meta Platforms v. Bundeskartellamt, ECLI:EU:C:2023:537, ¶ 117 (July 4, 2023).

[7] *See* Jessica Litman, *Information Privacy/Information Property*, 52 STAN. L. REV. 1283, 1283 (2000) (proposing a tort remedy for invasion of data privacy); ANITA ALLEN, UNPOPULAR PRIVACY: WHAT MUST WE HIDE? 10-11 (2011) (arguing that "it can be legitimate for liberal, egalitarian governments to mandate physical and informational privacy even when the privacy in question is unpopular—unwanted, resented, not preferred, or despised by intended beneficiaries or targets"); Ella Corren, *The Consent Burden in Consumer and Digital Markets*, 36 HARV. J.L. TECH. 551, 560, 599-600 (2023) (criticizing "empty consent" as a means of legitimizing harmful practices in digital markets, and praising regulations that substantively limit data processing irrespective of consent).

[8] Alysa Z. Hutnik, Aaron J. Burstein & Alexander I. Schneider, *The CCPA Non-Discrimination Right, Explained*, KELLEY DRYE (Apr. 29, 2020), https://www.kelleydrye.com/viewpoints/blogs/ad-law-access/the-ccpa-non-discrimination-right-explained [https://perma.cc/K56C-N3TJ]; Natasha Lomas, *Meta's EU Ad-Free Subscription Faces Early Challenge*, TECHCRUNCH (Nov. 28, 2023, 5:24 AM EST), https://techcrunch.com/2023/11/28/meta-ad-free-sub-noyb-complaint [https://perma.cc/S799-AF4N] (describing the challenge brought by one of Europe's privacy rights organizations arguing that a subscription charging thirteen Euros per month to avoid behavioral advertising was inappropriately overpriced).

[9] California Consumer Privacy Act, CAL. CIV. CODE § 1798.105 (2023).

[10] *Id.* §§ 1798.100, 1798.120.

[11] Ignacio Cofone offers one of the best critiques of the property frame. *See generally* Ignacio Cofone, *Beyond Data Ownership*, 43 CARDOZO L. REV. 501 (2021) (criticizing the property framing of privacy rights and arguing instead for tort-like liability rules). Like Cofone, I will argue in favor of liability rules that protect both the privacy of the data subjects and the liberties of the data processors. Ultimately, while we share a commitment to the torts frame, we emphasize different problems and needs within it.

market because the company collected and automatically analyzed users' queries.[12] Yet as advances in AI and machine learning place more power in the hands of end users to plot and commit a wide range of acts, both good and bad, AI safety, to say nothing of performance, will require AI companies to monitor the uses and misuses of their clients to avoid catastrophic risks.[13] Europe's AI Act will require all AI systems to guard against bias and other risks (which requires companies to take into account "characteristics or elements that are particular to the specific geographical, behavioural or functional setting"),[14] to maintain traceable logs of input data,[15] and to engage in post-market monitoring of users and information-sharing about threats.[16] Each of these decisions to collect, analyze, and occasionally disclose personal information undermines the data-subject control that was supposed to be so critical under GDPR. Europe's pretzel-shaped path for regulating the technology sector has a very uncertain future because the promise of fundamental rights to control personal information is simply not tenable.

Beyond its impracticality, treating personal information as property belonging to the data subject is unsound in principle, notwithstanding the widespread habit of referring to personal information as "*my* data." Privacy laws that attempt to create sticky privacy interests in personal data are not *merely* impractical. They are also incompatible with the philosophy of property rights. Treating personal data like sticky property—something that makes it difficult for the data subject to relinquish their control and easier to claw it back in most circumstances—lacks historical and logical foundation. Property rules rest on an assumption that the rights-holder has superior knowledge about the best uses of the property—when to exclude, when to share, and when to sell—and would do so without causing significant problems for others.[17] Outside of the special case of intellectual property, these conditions almost never hold when the object of the right is

[12] Karina Tsui, *Italy Bans ChatGPT over Privacy Concerns*, SEMAFOR (Mar. 31, 2023, 1:25 PM EDT), https://www.semafor.com/article/03/31/2023/chatgpt-banned-italy-privacy-concerns [https://perma.cc/9QTW-FNMA] (explaining that regulators found that the company has no legal basis under the GDPR to justify collecting query data and other personal data for the purpose of machine-learning training).

[13] Seyyed Ahmad Javadi, Richard Cloete, Jennifer Cobbe, Michelle Seng Ah Lee & Jatinder Singh, Poster Presentation, *Monitoring Misuse for Accountable 'Artificial Intelligence as a Service*,' AAAI/ACM CONF. ON AI, ETHICS & SOC'Y (Feb. 7-8, 2020), https://doi.org/10.1145/3375627.3375873 [https://perma.cc/YJ3N-UTAQ].

[14] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence and Amending Certain Union Legislative Acts, 2021/0106, art. 10(4)-(5) [hereinafter EU AI Act].

[15] EU AI Act, art. 12(4).

[16] *Id.* at art. 61 & 62.

[17] *See* discussion of the law and economics explanation for property rules *infra* Part I.

speech.[18] Just as "my" ideas, "my" opinions, and "my" observations are not really mine—not in any sense that allows me to exclude you from using them too—"my" data is also not my data.[19]

Privacy law should return to its roots in tort theory, where legal rules are intended to mediate conflicts between legitimate activities and interests without assigning veto power to anybody.[20] This is still the right frame. Good privacy policy will require lawmakers—courts, federal agencies, or what have you—to proactively protect people against risks that they may not have reason to know about. It will also require lawmakers to *permit* data processing that provides some benefit to the data processor, data subjects, or third parties without the necessity of getting the data subject's permission. This basic structure follows the American tradition of treating privacy as one of many objectives in a bustling zone of conflicting activities and interests.

This Essay argues that the American tradition of treating privacy as part of the management of social risks rather than as a sticky property bestowed to data subjects is a virtue of the American legal tradition that should not be cast aside in the rush to reign in technology companies.

This Essay proceeds as follows. Part I distinguishes a tort-style, risk-based treatment of privacy law from a property-style, rights-based framework and traces the historical, meandering path through both. Part II explains why a tort rule is more fitting for personal information than a property rule. Part III describes in broad strokes how a risk-based privacy regime would work. Courts, legislators, or regulators would need to establish some clear zones of protection (per se violations) and zones of liberty (safe harbors or per se nonviolations) in order to serve the foreseeable and obvious needs of data subjects and data users. They would also need to establish some benchmarks for analyzing novel forms of data processing. I provide a more elaborate discussion of the zones of liberty (safe harbors) because the breadth of these allowances is what distinguishes a risk-based approach from a rights-based approach that has some exceptions.

A return to the torts frame will set the United States up for success as privacy law is forced to respond to new uses of personal data in AI, autonomous vehicles,

[18] *See* discussion of speech as property *infra* Part I.

[19] *See generally* Eugene Volokh, *Freedom of Speech, Information Privacy, and the Troubling Implications of a Right to Stop People from Speaking About You*, 52 STAN. L. REV. 1049 (2000) (describing potential free-speech concerns arising from broad informational privacy rules).

[20] On the origins of privacy in state tort law, see William Prosser, *Privacy*, 48 CALIF. L. REV. 383, 385-86 (1960). On the relationship tort law has to liberty of action, see OLIVER WENDELL HOLMES, THE COMMON LAW 77 (Mark DeWolfe Howe ed., 1963) (1881) ("Furthermore, the public generally profits by individual activity. As action cannot be avoided, and tends to the public good, there is obviously no policy in throwing the hazard of what is at once desirable and inevitable upon the actor.").

health innovations, and other areas where meaningful systems of privacy self-management will be impractical and undesirable.

I. TORT VERSUS PROPERTY AND THE BATTLE FOR PRIVACY

Across the many ways to define property, the common feature is the right to exclude.[21] As Guido Calabresi and Douglas Melamed put it, property laws give an individual or group enough control over something that they have a limited veto power over when or how it is used.[22] They differentiated property entitlements from liability rules by focusing on the nature of the legal remedies: an individual who contests the actions of another under a liability rule might be awarded compensation based on their damages, and even then only if they can also prove fault.[23] The property owner has a different remedy. Because the property owner alone has the right to decide whether the action should be taken, the property owner can have the court completely undo the other's actions – by, for example, enjoining the defendant to return the item or not take the action again – and can demand punitive damages or other strong deterrents to reinforce the right of exclusive control.[24]

For the purposes of this Essay, I want to focus particularly on who gets to manage behavior by recognizing when a wrong occurs.[25] With tort-liability rules, when two parties disagree over whether an action was wrong or not, that disagreement is resolved by a disinterested rule maker. This would be a judge under the common law, and the determination would be made ex post, after the putative harm has occurred. But I will also count as a "tort" framework other forms of lawmaking that identify wrongs without allocating a property interest,

[21] The two most common formulations are the "bundle of rights" conception (focusing on rights between individuals) and the *in rem* or dominion conception that focuses on the relationship between the owner and the object of dominion. *See generally* James Penner, *The "Bundle of Rights" Picture of Property*, 43 UCLA L. REV. 711 (1996) (describing and critiquing the notion that property is a set of relational rights between people); Thomas W. Merrill & Henry E. Smith, *What Happened to Property in Law and Economics?*, 111 YALE L.J. 357 (2001) (discussing the decline in the *in rem* conception and rise in the "bundle of rights" conception of property).

[22] Guido Calabresi & A. Douglas Melamed, Property Rules, *Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1092 (1972).

[23] LANDES & POSNER, *supra* note 5, at 30.

[24] *Id.* at 29-30.

[25] A. M. Honoré, *Ownership*, *in* OXFORD ESSAYS IN JURISPRUDENCE 107, 107 (A. G. Guest ed., 1961) (one of the "standard incidents" of property ownership is the right to manage – to be able to decide who is allowed to use the thing and how they may do so). The rights to exclude and to trade will usually do most of the work getting to a right (or opportunity) to manage. Shyamkrishna Balganesh, *Demystifying the Right to Exclude: Of Property, Inviolability, and Automatic Injunctions*, 31 HARV. J.L. & PUB. POL'Y 593, 626 (2008).

even if that work is done by legislators or administrative agencies. The important point is that a tort approach to managing behavior requires a disinterested public entity to decide what sort of conduct is wrongful based on their assessment of a myriad of societal benefits and risks. By contrast, a property framework does not require a disinterested assessment of wrongs. One of the interested parties – the one who holds the property right – has the final word on whether the other could use the property or not. If you need another person's consent to do something, a property-style interest is involved.[26]

It is natural to assume that tort rules attach to *activities* while property rules govern *things*. For example, "driving" is an activity that multiple people can pursue without asking for your permission, but they do have to ask permission before driving *your* car.[27] However, the distinction between activities and things becomes muddy with intangible or nonrivalrous things. Nowhere is this more obvious than in the context of privacy. If a company creates a log of your movements throughout a store, is this an activity ("creating" a log) or an invasion (creating a log of "your" movements)?

American privacy law has wrestled with this question for over a century. In *The Right to Privacy*, the famous article by Samuel D. Warren and Louis D. Brandeis that started it all, privacy was conceived as control: "The common law secures to each individual the right of determining, ordinarily, to what extent his thoughts, sentiments, and emotions shall be communicated to others."[28] Warren and Brandeis explicitly reject the idea of treating privacy like a property right, but this is only to distinguish it from the colloquial meaning of property that is typically commoditized and valued through its distribution and sale.[29] As for the formal definition of a property right, where the law offers exclusive control (including the hoarding of the property and the perpetual exclusion of others), Warren and Brandeis had precisely this in mind.

However, the earliest instantiation of privacy claims that could be brought against a private party emerged through common-law tort.[30] Privacy-related tort claims were recognized only when plaintiffs could show they suffered "outrageous" intrusions or disclosures that would "be offensive and objectionable to a reasonable man of ordinary sensibilities."[31] In other words, people were

[26] Barbara Prainsack, *Logged Out: Ownership, Exclusion and Public Value in the Digital Data and Information Commons*, 6 BIG DATA & SOC'Y, Jan.-June 2019, at 1, 3.

[27] Louis Kaplow & Steven Shavell, *Property Rules Versus Liability Rules: An Economic Analysis*, 109 HARV. L. REV. 713, 716 (1996).

[28] Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 198 (1890).

[29] *Id.* at 200.

[30] Prosser, *supra* note 20, at 397-98.

[31] *Id.* at 396.

generally at liberty to observe each other, share gossip, and otherwise invade what some could consider to be their personal bubble, just as they were at liberty to pursue other private activities like driving or playing frisbee in a park—even when those acts cause annoyance, delay, or accidental injury. But if the activities of collecting or sharing information are *unreasonable*, and if as a result of those activities the target of the privacy-intruding activity has suffered distress or concrete harm, then tort law recognizes a wrong.[32]

The emphasis on data-subject control reemerged in the 1960s, when computers became more common. The efficiency of computers changed the quality and quantity of personal data collection, especially by the government. In response to the anxiety around computers, an influential congressional report (the HEW Report) highlighted data-subject control more than American tort law traditionally had. According to the "Fair Information Practice Principles" (FIPPs) promulgated in the report, a data processor should not be able to share personal data with another entity without the informed consent of the data subject.[33] However, a closer read of the HEW Report reveals more nuance. The report explicitly rejected formulations of privacy that assume the data subject has exclusive control and instead favored the concept of "mutuality."[34] As it explained:

> [Some of the privacy formulations] speak[] of the data subject as having a unilateral role in deciding the nature and extent of his self-disclosure. None accommodates the observation that records of personal data usually reflect and mediate relationships in which both individuals and institutions have an interest, and are usually made for purposes that are shared by institutions and individuals. In fact, it would be inconsistent with this essential characteristic of mutuality to assign the individual record subject a unilateral role in making decisions about the nature and use of his record.[35]

The 1977 report *Personal Privacy in an Information Society*, produced by the Privacy Protection Study Commission, further refined the FIPPs to make clear that they do not place data in the data subject's absolute control. For example, where

[32] *Id.* at 391, 396-97; Gill v. Hearst Pub. Co., 253 P.2d 441, 444 (Cal. 1953); Samuel v. Curtis Pub. Co., 122 F. Supp. 327, 329 (N.D. Cal. 1954); Sidis v. F-R Pub. Corp., 113 F.2d 806, 808-09 (2d Cir. 1940).

[33] Dep't Health, Educ. & Welfare, No.(OS)73-94, Records, Computers and the Rights of Citizens: Report of the HEW Advisory Committee on Automated Personal Data Systems 53 (1973) [hereinafter HEW Report].

[34] *Id.* at 3, 40.

[35] *Id.* at 40.

the HEW Report says that "[t]here must be a way for an individual to prevent information about him obtained for one purpose from being used or made available for other purposes without his consent,"[36] the 1977 report uses the principle that "[t]here shall be limits on the external disclosures of information about an individual a record-keeping organization may make (the Disclosure Limitation Principle)."[37] Swapping data-subject consent for "limits" of some unspecified origin marks a shift away from the property frame.

Nevertheless, privacy statutes that were modeled after the FIPPs—starting with the Privacy Act[38] (which constrains how the federal government handles personal data) and including privacy statutes related to healthcare (HIPAA[39]), credit reporting (FCRA[40]), and electronic communications (ECPA[41])—prioritized a consent-based regime while leaving enough leeway and loopholes for the regulated industries to achieve some minimal level of innovation and operational efficiency.[42] Recent proposals for federal privacy legislation have pushed for more data-subject control, with fewer allowances, and the thrust of most legal scholarship runs in the same direction.[43]

Now, to be clear, the division between the property and tort frameworks is not so sharp. Nearly every privacy advocate and scholar understands that privacy ensures a personal sphere that is shielded, but not absolutely closed off, from the

[36] *Id.* at xx.

[37] PRIVACY PROTECTION STUDY COMMISSION, PERSONAL PRIVACY IN AN INFORMATION SOCIETY 502 (1977).

[38] Privacy Act of 1974, 5 U.S.C. § 552a (2018).

[39] Health Insurance Portability and Accountability Act Privacy Rule, 45 C.F.R. §§ 164.500 et seq. (2023).

[40] Fair Credit Reporting Act, 15 U.S.C. §§ 1681b et seq. (2018).

[41] Electronic Communications Privacy Act of 1986, 18 U.S.C. §§ 2510-23 (2018).

[42] The Privacy Act (which governs how the federal government manages its own collections of personal data) includes exceptions for criminal law enforcement, the investigation of tax and social security fraud, for purposes of "compelling circumstances affecting the health or safety of an individual," and for "routine uses" without getting the data subject's consent. 5 U.S.C. § 552a (2018). Many other U.S. privacy statutes carve out scenarios where data can be collected, used, or shared without consent, too. *See, e.g.*, Health Insurance Portability and Accountability Act (HIPAA), Pub. L. No. 104-191, 110 Stat. 1936 (1996) (codified as amended in scattered sections of 18, 26, 29 & 42 U.S.C.); Family Educational Rights and Privacy Act (FERPA), 20 U.S.C. § 1232g (2018); Electronic Communications Privacy Act, 18 U.S.C. §§ 2510-2523 (2018).

[43] American Data Privacy and Protection Act, H.R. 8152, 117th Cong. (2022); Daniel J. Solove, *Murky Consent: An Approach to the Fictions of Consent in Privacy Law*, 104 B.U. L. REV. (forthcoming 2024) (manuscript at 34-35), https://ssrn.com/abstract=4333743 [https://perma.cc/Y3AX-XXT7]; Luiza Jarovsky, *Improving Consent in Information Privacy Through Autonomy-Preserving Protective Measures (APPMs)*, 4 EUR. DATA PROT. L. REV. 447, 451-53 (2018); SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM 516-25 (2018).

uses and needs of others. Even when privacy law is built around a right of control, the right of a data subject to lock away information has been understood as a limited one that must be reconciled with, and sometimes superseded by, other compelling social needs. Leading privacy law scholars including Alan Westin,[44] Daniel J. Solove,[45] Helen Nissenbaum,[46] and Neil Richards[47] have recognized that information control is not always in the best interest of society or of the data subjects themselves, and privacy does not and should not require consent in every single conceivable case.[48]

So, the debate boils down to how wide or narrow the scope of freedom is for the data *user*—the potential privacy-violator, that is. It might be useful to separate the torts framework from the property one by thinking of defaults: Is it the case that individuals generally have control over the information that describes them, and exceptions are made to that general rule? (This would be the property frame.) Or is it instead more accurate to say that individuals generally have the freedom to observe, collect, and share information about others, and that this

[44] ALAN WESTIN, PRIVACY AND FREEDOM 374 (1967) ("[T]he needs of social order may require certain levels of exposure and confession even if these are involuntary. Consent is thus to be analyzed in the specific context of the purpose of [data collection] and the use to be made of the information so obtained.").

[45] Daniel J. Solove, *Introduction: Privacy Self-Management and the Consent Dilemma*, 126 HARV. L. REV. 1880, 1881 (2013) ("[I]n order to advance, privacy law and policy must face the problems with privacy self-management and start forging a new direction.").

[46] HELEN NISSENBAUM, PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE 85-88 (2010).

[47] NEIL RICHARDS, WHY PRIVACY MATTERS 71 (2022) ("[M]y own definition [of privacy] excludes other ways we could talk about privacy, such as its being a right to control our personal information or the ability to conceal disreputable information about ourselves.").

[48] Kirsten Martin, *Understanding Privacy Online: Development of a Social Contract Approach to Privacy*, 137 J. BUS. ETHICS 551, 556-57 (2016) ("Defining privacy as a state of inaccessibility is neither practical nor desirable and, ironically, renders privacy as a form of punishment . . . . These social contracts around what, to whom, and for what purpose information flows are the governing rules about privacy for a given community."); Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III & Kate Crawford, *Datasheets for Datasets*, MICROSOFT (July 9, 2018), https://www.microsoft.com/en-us/research/uploads/prod/2019/01/1803.09010.pdf [https://perma.cc/3GZ8-968K] (encouraging database creators to ask whether individuals were "told what the dataset would be used for and did they consent," but not requiring the answer to determine, on its own, the privacy assessment of a program). The concept of "forfeiture" in property law (particularly where that forfeiture is in the form of an *involuntary* relinquishment) can also produce some slippage between the property and tort frameworks. *See* Mark L. Hanin, *Privacy Rights Forfeiture*, 22 J. ETHICS & SOC. PHIL. 239, 259 (2022).

general rule of *permissiveness* is limited under circumstances that are harmful or risky? (This would be the torts frame.) [49]

Put this way, the torts framework is anathema to nearly every serious piece of scholarship or privacy proposal put out over the last several decades.[50] Even Daniel Solove and Ignacio Cofone, who have advocated for regulating privacy based on risk rather than property-style user self-management (and for reasons very similar to my own), have not veered very far from the property frame's center of gravity.[51] Cofone would treat unexpected repurposing of personal data as a form of privacy harm per se that can support liability.[52] Solove has embraced such a capacious definition of harm that his proposals would still require data-dependent firms to seek consent or stop what they are doing in order to avoid exposure to debilitating liability in a wide variety of real-world scenarios.[53] For example, Solove suggests that use of personal data to create predictions should be regarded as risky based on the chance of error.[54] He has also argued that

[49] Kaplow & Shavell, *supra* note 27, at 716 ("We are permitted to engage in such acts—from hunting to driving to construction—even though they create risks of harm and thus constitute probabilistic invasions of property interests, but we are often obligated to pay damages for any harm that we cause.").

[50] *See* LAWRENCE LESSIG, CODE VERSION 2.0, at 227-28 (2006); *see generally* Paul M. Schwartz, *Property, Privacy, and Personal Data*, 117 HARV. L. REV. 2056 (2004) (developing a framework of property protections for personal information); Richard S. Murphy, *Property Rights in Personal Information: An Economic Defense of Privacy*, 84 GEO. L. REV. 2381 (1996) (addressing economic arguments against privacy legislation with a property-based lens); Marc Rotenberg, *Fair Information Practices and the Architecture of Privacy (What Larry Doesn't Get)*, 2001 STAN. TECH. L. REV. 1 (reviewing LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE (1999)); Thomas D. Haley, *Illusory Privacy*, 98 IND. L.J. 75 (2022) (critiquing the prevailing "notice and consent" paradigm in privacy literature).

[51] Daniel Solove, *Data Is What Data Does: Regulating Based on Harm and Risk Instead of Sensitive Data*, 118 NW. U. L. REV. (forthcoming 2024) (manuscript at 47), https://ssrn.com/abstract=4322198 [https://perma.cc/84WN-G499] ("Treating all situations as equal often provides inadequate protections to high-risk situations. Another problem is treating low-risk situations with too many restrictions. Cumbersome and unnecessary restrictions trivialize privacy rules, making people perceive them as silly inconveniences and annoyances.").

[52] Indeed, he explains that the data subject retains significant property-like control over data that they had previously exchanged with others. Cofone, *supra* note 11, at 567, 569. Thus, the liability rule is really only operating for the function of making determinations, based on "reasonableness" about which data uses have been given away and which have not.

[53] For critiques of Daniel Solove's capacious definitions and taxonomies of privacy harms, see Ryan M. Calo, *The Boundaries of Privacy Harm*, 86 IND. L.J. 1131, 1139-42 (2011); Maria P. Angel & Ryan Calo, *Distinguishing Privacy Law: A Critique of Privacy as Social Taxonomy*, 123 COLUM. L. REV. (forthcoming 2024) (manuscript at 18-36), https://ssrn.com/abstract=4347191 [https://perma.cc/8AK4-XPA9].

[54] Solove, *supra* note 51 (manuscript at 46). Solove does not explain whether there should be some measure of acceptable versus unacceptable risk. He may have some idea in mind, but as written, the article currently seems to propose treating all predictive analytics as presumptively

anxiety about downstream consequences of a revelation should be recognized as a harm – even the fallout when a person is exposed as a liar.[55] If firms face credible threats of liability in these scenarios based on the risk of anxiety or error, the scope of freedom becomes severely limited to a range that looks, to me, not much different than that afforded by the GDPR.[56]

The privacy scholarship has created a misimpression for the general public that strong control-based privacy laws do not pose serious limitations on legitimate and useful activities – on freedoms to experiment and innovate, to perform research, to speak, to compete against dominant technology firms, or to offer content and services at a price that is heavily subsidized by behavioral advertising. A tort approach that emphasizes these liberties, and that creates legal liability only when a data practice foreseeably causes unjustified and concrete harm to others, offers much more promise for an enduring form of consumer protection.[57] A tort approach deters and provides recourse for activities that are harmful in a meaningful sense of causing real welfare reductions, and it also frees the data users to pursue activities that are *not* likely to cause harm.

II. PRIVACY LAW NEEDS TO MANAGE RISKS WHILE RECOGNIZING LEGITIMATE DATA ACTIVITIES

Why is it better to have a risk-management system overseen by a judge or regulator rather than a sticky property right managed by individual data subjects? After all, if the data economy is good for consumers, they can always

risky. Another source of confusion is that Daniel Solove was the reporter for the American Law Institute's Principles of Law, Data Privacy – a project that he seems to have personally endorsed. The Principles state that consent is "a core element of privacy law," and forbids data processing that would be "significantly unexpected" to the data subject even if the processing poses no foreseeable harm. Daniel J. Solove & Paul M. Schwartz, *ALI Data Privacy: Overview and Black Letter Text*, 68 UCLA L. REV. 1252, 1272, 1275-76 (2022). On endorsement, see Daniel Solove, *ALI Data Privacy Principles*, TEACH PRIV. (Dec. 17, 2020), https://teachprivacy.com/ali-data-privacy-principles [https://perma.cc/WUF7-U5SU].

[55] Daniel J. Solove & Danielle Keats Citron, *Risk and Anxiety: A Theory of Data-Breach Harms*, 96 TEX. L. REV. 737, 764-65 (2018) (using information about adultery revealed by the Ashley Madison website as an example).

[56] Ella Corren's recent work similarly advocates for a risk-regulation approach over a property/consent framework. But she, too, seems to have a capacious understanding of harm, assuming that collections or disclosures of personal data are presumptively risky. Corren, *supra* note 7, at 571 (describing "exploitations" of user data without articulating the harm); *id.* at 582 (assuming that firms that collect user data are in a zero-sum game with their consumers and therefore offer the least favorable terms to them); *id.* at 583 (implying that JetBlue's sale of customer data is harmful).

[57] HOLMES, *supra* note 20, at 77; Hanin, *supra* note 48, at 240, 244-47 (considering unfairness to privacy "duty-bearers").

choose to license or sell access to their personal information. Why should decision-making be taken out of their hands?

The theory and scholarship coming out of the law-and-economics movement continues to offer the richest and most sustained attention on questions about when human interactions should be regulated through private property rights or through liability rules. Generally speaking, law should recognize a property interest when individuals have special information about how to get the best value out of a resource and transaction costs are low enough to allow everybody to trade and rearrange their entitlements so that the resource is used for its most valuable purposes.[58] Conversely, the scale tips against recognizing a property interest if transaction costs are high, or if the decisions of the rightsholder are likely to cause negative externalities to third parties.

For reasons I explain, these factors cut against recognizing a property interest in personal data.

A. Information Gaps and Transaction Costs

Would data subjects know how to maximize the value of their personal data if they had full control over its uses? It is hard to believe they would. One problem is that people might not have stable values for their own privacy, as suggested by the so-called "privacy paradox." This is the frequently replicated phenomenon where individuals report high levels of concern about their privacy but

[58] Ronald H. Coase, *The Problem of Social Costs*, 3 J.L. & ECON. 1, 19 (1960). In a perfect marketplace, all mutually beneficial transactions would take place and no exploitative or harmful transactions would not. The ideal marketplace cannot exist because of the very real-world costs that have to be incurred to find and execute good transactions while avoiding bad ones. For example, the costs of transporting goods to a place where a buyer can receive them will of course have cost and cause some transactions to not occur. Thus, nearly all of the challenge of designing good law is in anticipating and understanding a wide range of transaction costs, some of which can be exploited and abused. *See also* LANDES & POSNER, *supra* note 5, at 31, 43 (explaining that in conditions of high transaction costs, either liability rules are preferable or defenses to a trespass action should be available). The arguments I raise here are similar (though not identical) to those raised in the context of intellectual property. Mark A. Lemley, *Property, Intellectual Property, and Free Riding*, 83 TEX. L. REV. 1031, 1032 (2005) ("[F]ull internalization of positive externalities is not a proper goal of [intellectual] property rights except in unusual circumstances, for several reasons: (1) there is no need to fully internalize benefits in intellectual property; (2) efforts to capture positive externalities may actually reduce them, leaving everyone worse off; and (3) the effort to capture such externalities invites rent-seeking."); Joseph E. Stiglitz, *The Revolution of Information Economics: The Past and the Future* 3 (Nat'l Bureau of Econ. Rsch., Working Paper No. 23780, 2017) ("But these property rights issues are different from and more complex than those concerning conventional property rights, where it is usually assumed the stronger the better. Here, the ambiguities in the assignment of property rights are apparent, and so called strong (intellectual) property rights may lead to poorer economic performance.").

are also willing to give it up for small payments or perks.[59] While there are several theories that could explain the paradox,[60] the most parsimonious explanation is that abstract questions about the value of privacy cannot account for what is actually a pretty utilitarian calculation. Concerns about privacy are adjusted up or down depending on the likely consequences of each data practice in context.[61] People will generally allow a data practice if they believe the benefits outweigh the risks.[62]

[59] Susan Athey, Christian Catalini & Catherine Tucker, *The Digital Privacy Paradox: Small Money, Small Costs, Small Talk* 17-18 (Nat'l Bureau of Econ. Rsch., Working Paper No. 23488, 2017); Kai-Lung Hui, Hock Hai Teo & Sang-Yong Tom Lee, *The Value of Privacy Assurance: An Exploratory Field Experiment*, 31 MIS Q. 19, 26-27 (Mar. 2007); Jan H. Schumann, Florian von Wangenheim & Nicole Groene, *Targeted Online Advertising: Using Reciprocity Appeals to Increase Acceptance Among Users of Free Web Services*, 78 J. MKTG. 59, 69-71 (Jan. 2014). Consumers rarely alter the defaults in privacy settings. *See* Omri Ben-Shahar & Adam Chilton, *Simplification of Privacy Disclosures: An Experimental Test*, 45 J. LEG. STUD. S41, S65-S66 (June 2016); Eric J. Johnson, Steven Bellman & Gerald H. Lohse, *Defaults, Framing and Privacy: Why Opting in-Opting out*, 13 MKTG. LETTERS 5, 13-14 (2002). And the drafting of privacy notices has almost no effect on behavior. *See* Lior Jacob Strahilevitz & Matthew B. Kugler, *Is Privacy Policy Language Irrelevant to Consumers?*, 45 J. LEGAL STUD. S69, S87-S93 (June 2016); Ben-Shahar & Chilton, *supra*, at S42. In my own work, I have found that privacy-related notices are often wasteful and do not change consumer choices. Dramatic just-in-time disclosures have the best potential to change behavior, but they also run the risk of exaggerating a sense of threat and distorting the consumer's evaluation of other criteria. Jane Bambauer, Jonathan Loe & Alex D. Winkelman, *A Bad Education*, 2017 ILL. L. REV. 109, 149-51. Even when consumers are paying attention to privacy options, they rarely forego a service or benefit that they would otherwise want in order to protect their privacy. Strahilevitz & Kugler, *supra*, at S79; Athey, Catalini & Tucker, *supra*, at 13-14; *see generally* Alessandro Acquisti, Laura Brandimarte & George Loewenstein, *Privacy and Human Behavior in the Age of Information*, 347 SCI. 509 (2015) (providing a summary of related scholarship).

[60] For a summary of the literature, see Christoph Lutz, Christian Pieter Hoffmann & Giulia Ranzini, *Data Capitalism and the User: An Explanation of Privacy Cynicism in Germany*, 22 NEW MEDIA & SOC'Y 1168, 1170-72 (2020); and Spyros Kokolakis, *Privacy Attitudes and Privacy Behavior: A Review of Current Research on the Privacy Paradox Phenomenon*, 64 COMPUTS. & SEC. 122 (2017).

[61] *See* Kirsten Martin & Helen Nissenbaum, *Measuring Privacy: An Empirical Test Using Context to Expose Confounding Variables*, 18 COLUM. SCI. & TECH. L.J. 176, 180-83 (2017). *See also* Kirsten Martin & Helen Nissenbaum, *Privacy Interests in Public Records: An Empirical Investigation*, 31 HARV. J.L. & TECH. 111, 130-37 (2017) (describing some of the contextual factors that appear to matter when public data is disclosed to different entities); Long Chen, Yadong Huang, Shumiao Ouyang & Wei Xiong, *The Data Privacy Paradox and Digital Demand* 27 (Nat'l Bureau of Econ. Rsch., Working Paper No. 28854, 2021) (finding that individuals who were the most privacy-sensitive were also the most interested in digital services, and therefore gave more permissions in exchange for those services).

[62] *See, e.g.*, Chen et al., *supra* note 61, at 3 (finding that individuals who have the greatest concern for privacy also get the greatest value from data-sharing); Mary J. Culnan & Robert J. Bies, *Consumer Privacy: Balancing Economic and Justice Considerations*, 59 J. SOC. ISSUES 323, 327 (2003); Tamara Dinev & Paul Hart, *An Extended Privacy Calculus Model for E-Commerce*

In other words, for most people, the value that they derive from personal data either being used or not used depends on what they get out of it, what they lose from it, and whether it helps other people. This would be a good match for a property and contract/market system if the type and value of data processing was obvious to data subjects, and if the consenting process was low cost. The trouble is, it is costly and time-consuming not only to manage consent processes, but even to analyze the risks and benefits of each data usage to figure out whether to consent in the first place.[63]

The main benefits that Big Data brings to consumers are the same ones that the data-using companies want, too: access to personal data helps drive down transaction costs.[64] Within the broad set of factors that can cause markets to be

Transactions, 17 INFO. SYS. RSCH. 61, 62 (2006); Hui et al., *supra* note 59, at 20; Heng Xu, Hock-Hai Teo, Bernard C.Y. Tan & Rita Agrawal, *The Role of Push-Pull Technology in Privacy Calculus: The Case of Location-Based Services*, 26 J. MGMT. INFO. SYS. 135, 137 (2009); Susanne Barth & Mennode Jong, *The Privacy Paradox—Investigating Discrepancies Between Expressed Privacy Concerns and Actual Online Behavior—A Systematic Literature Review*, 34 TELEMATICS & INFORMATICS 1038, 1044 (2017); Kirsten Martin, *Breaking the Privacy Paradox: The Value of Privacy and Associated Duty of Firms*, 30 BUS. ETHICS QUART. 65, 71-72 (2019) (summarizing this literature); Martin, *supra*, ("Uses of information deemed privacy violations in consumer surveys may be better judged after taking into consideration the benefits of sharing information online."); Miremad Soleymanian, Charles B. Weinberg & Ting Zhu, *Privacy Concerns, Economic Benefits, and Consumer Decisions: A Multi-Period Panel Study of Consumer Choices in the Automobile Insurance Industry* (Aug. 13, 2021), https://ssrn.com/abstract=3905034 [https://perma.cc/D73R-ZTN9]. My own experimental research has found that most Americans primarily value privacy for its instrumental purposes—to avoid concrete risks to themselves and others—and will therefore perceive less of a threat to privacy when the benefits of a data practice outweigh those risks. *See generally* Jane Bambauer, *Privacy Tradeoffs: Who Should Make Them, and How?*, (TPRC49: The 49th Research Conference on Communication, Information and Internet Policy Working Paper, 2021), https://ssrn.com/abstract=3905024 [https://perma.cc/M63X-EXHV] (tentatively finding that most Americans take a utilitarian cost-benefit approach to making judgments in the personal data context); Jane Bambauer et al., *supra* note 59 (considering how mandated disclosures interact with consumer cost-benefit analyses with regard to privacy).

[63] Corren, *supra* note 7, at 570; *see generally* Daniel Bjorkegren, *Nostalgic Demand* (June 25, 2018), https://ssrn.com/abstract=3220583 [https://perma.cc/MV65-2BZJ] (finding that when the quality of products are difficult to ascertain, consumers will rely on rough signals of quality like nostalgic brands, but that these preferences for nostalgia dissipate when consumers have better information about quality). The parallel here is that consumers may prefer privacy where they do not have good information about the ultimate uses and consequences of data processing. People are hypervigilant and overly wary of new information technologies given that technology (along with immigration) causes a persistent, exaggerated, and well-documented sense of threat to humans. BRYAN CAPLAN, THE MYTH OF THE RATIONAL VOTER: WHY DEMOCRACIES CHOOSE BAD POLICIES 23-93 (2007).

[64] Thus, I disagree with legal scholarship that insists that any surreptitious use of data to enhance efficiency presents a benefit for the firm and a cost for the customer. *See, e.g.*, Ryan Calo & Alex Rosenblat, *The Taking Economy: Uber, Information, and Power*, 117 COLUM. L. REV. 1623, 1654 (describing the use of data for broader logistics purposes as a taking). According to the

inefficient, unfair, or to even fail, many spring from the fact that the participants in the market lack relevant information to make the best choices. These include search and matching costs (the costs of making sure that prospective buyers know about what is offered for sale and that prospective sellers know where there is demand for their products and services); verification costs (making sure that goods and services offered for sale really are what sellers purport them to be); the costs of bargaining; and the costs of policing or enforcing performance of a contract.[65] When these costs are reduced, the surplus is usually shared by all participants in the market.

For example, one longstanding source of concern is the use of greater amounts of personal data for credit scoring and lending decisions. The unspoken presumption is that a scoring system that uses a greater amount of personal data creates a privacy imposition on the loan applicants and a financial benefit to the banks. But this is not so. A study of mortgages in the San Francisco Bay Area found that loan applicants living in counties with greater privacy protections (that set privacy as the default) paid higher interest rates *and also defaulted more often* than the applicants living in the counties that set data flow as the default, even after controlling for confounders.[66] Banks could not match applicants to loans as well, so there was more risk, and the costs of risk were, of course, passed along to the consumers.

Consider another example of reduced matching costs (though it is rarely discussed in quite this way): quarantine during a pandemic. The goal of a quarantine, whether voluntary or compulsory, is to restrict the movements of individuals who are most likely to be infected and contagious without interfering with the activities of those who are least likely to be so. In an environment with strong privacy defaults and information friction, this is very hard to do, with the result being that more people are quarantined, more people are infected, *or both.* South Korea's public health authority took the unusual step of publicly disclosing the time-stamped geolocation of individuals who later tested positive for COVID, allowing residents to self-assess whether they had been exposed to the virus and

authors, in the sharing economy, "You *are* the customer, quite literally, so you do not necessarily think of yourself as a product, too." *Id.* at 1652. To the contrary, reduced transaction costs are a win-win nonzero prospect.

[65] Subsets of this list appear in Carl J. Dahlman, *The Problem of Externality*, 22 J.L. & ECON. 141, 148 (1979); Joseph E. Stiglitz, *The Contributions of the Economics of Information to Twentieth Century Economics*, 115 Q.J. ECON. 1441, 1452 (2000); and Avi Goldfarb & Catherine Tucker, *Digital Economics*, 57 J. ECON. LIT. 3, 3 (2019).

[66] *See* Jin-Hyuk Kim & Liad Wagman, *Screening Incentives and Privacy Protection in Financial Markets: A Theoretical and Empirical Analysis*, 46 RAND J. ECON. 1, 8, 18 (2015). This is consistent with the more general phenomenon of risk-based lending markets. *See* Wendy Edelberg, *Risk-Based Pricing of Interest Rates for Consumer Loans*, 53 J. MONETARY ECON. 2283, 2283 (2006).

quarantine if necessary.[67] This marked a loss of control for the individuals whose location histories were shared automatically, and for every South Korean resident who may in the future become COVID-positive. But in exchange for this loss, Korean residents avoided an estimated 200,000 cases and 7,700 deaths during the subsequent four months.[68]

Transaction costs have become visible in the wake of the implementation of the GDPR in Europe. The GDPR caused investment in research and development for new technology startups in the EU to falter, and the productivity of EU firms fell as compared to U.S. counterparts.[69] Meanwhile, the internal budgets for privacy offices at companies of all sizes (including in the United States) increased twenty-nine percent.[70] Costs to American companies that comply with all aspects of GDPR-style laws are estimated to be approximately $480 per data subject.[71] Companies respond by raising prices for consumers—a cost that may well be worth it to some people and in some contexts, but probably not as a general rule. And this ignores the costs of inconvenience not only in the form of the time required to click through and manage consents, but also in terms of the degraded service that results from a less customized experience. For example, the introduction of GDPR seems to have resulted in consumers having to use twenty-one percent more search terms and to access sixteen percent more websites before making their online transactions.[72]

That said, when personal data winds up in the vault of a data company, there is no guarantee it will be used for net-beneficial purposes (like matching loans to applicants) rather than welfare-reducing purposes (like creating sucker

[67] Jung Won Sonn, *Coronavirus: South Korea's Success in Controlling Disease Is Due to Its Acceptance of Surveillance*, CONVERSATION (Mar. 19, 2020, 12:41 PM EDT), https://theconversation.com/coronavirus-south-koreas-success-in-controlling-disease-is-due-to-its-acceptance-of-surveillance-134068 [https://perma.cc/LS3N-DFTD].

[68] David O. Argente, Chang-Tai Hsieh & Munseob Lee, *The Cost of Privacy: Welfare Effects of the Disclosure of COVID-19 Cases* 16 (Nat'l Bureau of Econ. Rsch., No. 27220, 2020).

[69] Jian Jia, Ginger Zhe Jin & Liad Wagman, *The Short-Run Effects of GDPR on Technology Venture Investment* 4 (Nat'l Bureau of Econ. Rsch., Working Paper No. 25248, 2018); EUROPEAN COMMISSION, DIGITAL ECONOMY AND SOCIETY INDEX (DESI) 2021, at 6, 9 (2021).

[70] Müge Fazlioglu, *IAPP-EY Annual Privacy Governance Report 2021*, INT'L ASS'N OF PRIV. PROS. at xii (2021), https://ssrn.com/abstract=4227244 [https://perma.cc/P3BG-BQP7].

[71] Alan McQuinn & Daniel Castro, *The Costs of an Unnecessarily Stringent Federal Data Privacy Law*, INFO. TECH. & INNOVATION FOUND. 1 (Aug. 5, 2019), https://itif.org/publications/2019/08/05/costs-unnecessarily-stringent-federal-data-privacy-law [https://perma.cc/XBL7-L5NF].

[72] Yu Zhao, Pinar Yildirim & Pradeep Chintagunta, *Privacy Regulations and Online Search Friction: Evidence from GDPR*, NAT'L BUREAU ECON. RSCH. 1 (Aug. 2021), https://conference.nber.org/conf_papers/f160434.pdf [https://perma.cc/B9QC-M84A]. The authors also found that larger retailers benefited, relative to smaller retailers, from the search frictions caused by GDPR.

lists).[73] It's not just consumers but the data *collectors*, too, who do not know how information that is stored or shared can be used in the future—to the data subject's benefit or detriment.[74] The information vacuum leaves data subjects with no clear preference between sharing their personal information and trying to lock it away. Privacy discussions in the media and legal journals do not usually even attempt to net this out by comparing the chance of harm from lost control over personal data to the chance of benefits to the data subjects and to others. Our collective instincts about the Big Data economy may be overly pessimistic because so much of the benefit comes in an invisible form of lowering needless transaction costs. But the larger point is that the holders of the property right, if one were to exist, would not know which disclosures and uses inure to their benefit and which do not. If risk management were handed to a trusted and trustworthy authority—a judge or an agency that had the right incentives to identify and weed out harmful practices—the data subject would be freed from worry and from the relentless queue of consent requests. [75] Otherwise, data subjects are destined to be habitual consenters or nonconsenters. The habitual consenters will suffer the costs of oversharing (e.g., greater risk of identity theft), and the habitual nonconsenters will bear the costs of undersharing (e.g., higher interest rates).

B. Externalities and Collective Action Problems

If a property rule is likely to cause harmful externalities—that is, harm to individuals who are not represented by the parties bargaining over the use of the property—liability rules should apply instead.[76] A veto over the use of personal data will cause externalities, both bad and good. The good spillover effects include protecting others when a harmful practice like imputing information or predicting the behavior of others for a malicious purpose is stymied by others

[73] These are lists of consumers that are susceptible and vulnerable to fraud. Ginger Allen, *The Suckers List: Even if You Weren't Scammed, You Could Be on It*, CBS TEX. (Dec. 16, 2020, 8:00 PM), https://www.cbsnews.com/texas/news/suckers-list-scammed [https://perma.cc/8EPV-VLGF].

[74] Ginger Zhe Jin, *Artificial Intelligence and Consumer Privacy*, *in* THE ECONOMICS OF ARTIFICIAL INTELLIGENCE: AN AGENDA 439, 440 (Ajay Agrawal, Joshua Gans & Avi Goldfarb eds., 2019) ("Since data can be stored, traded, and used long after the transaction, future data use is likely to grow with data processing technology such as AI. More important, future data use is *obscure to both sides* . . . .").

[75] Kaplow & Shavell, *supra* note 27, at 720 ("[W]hat if bargaining is not always successful because parties sometimes misgauge what each other is willing to pay or accept? In this case, no unambiguous conclusion can be drawn . . . .").

[76] RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 70 (1973); Calabresi & Melamed, *supra* note 22, at 1108, 1119-20.

who refuse to supply their data during the training stage. These spillover effects of privacy are discussed in the privacy literature.[77] But there are also several types of negative externalities, and these are likely to be more common. The negative externalities include (1) evading detection of fraud, (2) frustrating attempts to improve accuracy and to test for and reduce bias, and (3) reducing competition in the technology industry.

Evading detection of fraud. Harmful externalities would arise from personal data ownership any time a data subject can exercise veto power during the course of an investigation for fraud, crime, or other misbehavior. If a fraudster exploits the privacy of their communications or bank accounts to obscure their bad intentions and gain the trust of a mark, the costs of privacy are borne not by the banks or the communications providers but by the third-party fraud victim. This problem was vividly captured in Richard A. Posner's *The Right of Privacy*, which cautioned proponents of privacy rights that it can become a right to commit a wide range of formal and informal frauds.[78]

Reducing accuracy and bias correction. Data practices in machine learning or basic social-science research depend on having a representative sample of data to perform well and avoid biased results.[79] These goals would be frustrated if some (nonrandom) set of data subjects refuse to allow access to their data.[80] Indeed, the timely and worthy goals of tackling unintentional bias in AI systems will require *more* personal data to avoid biased and unnecessary error in predictions.[81] In theory, because everybody benefits from a more accurate and fair AI or machine-learning system, everybody would be willing to pay their share, via money or reductions in privacy, to ensure that the AI system has access to an adequate amount of training and context data. But nobody has the incentive to pay extra to make up for data holdouts. Even the AI service providers, who would have *some* incentive to achieve a minimum level of accuracy to have a viable product, are likely to lack the incentive to pay for the optimal amount of personal data because a biased or error-prone system can still be marketable as long as it performs better than the available alternatives.

[77] Catherine Tucker, *Privacy, Algorithms, and Artificial Intelligence*, *in* THE ECONOMICS OF ARTIFICIAL INTELLIGENCE: AN AGENDA 423, 431-34 (Ajay Agrawal, Joshua Gans & Avi Goldfarb eds., 2019).

[78] Richard A. Posner, *The Right of Privacy*, 12 GA. L. REV. 393, 398 (1978).

[79] Michael Mannino, Yanjuan Yang & Young Ryu, *Classification Algorithm Sensitivity to Training Data with Non-Representative Attribute Noise*, 46 DECISION SUPPORT SYS. 743, 743-46 (2009).

[80] Jane Yakowitz Bambauer, *Tragedy of the Data Commons*, 25 HARV. J.L. & TECH. 1, 61, 64 (2011) (published as Jane Yakowitz).

[81] Sandra G. Mayson, *Bias in, Bias out*, 128 YALE L.J. 2218, 2224 (2019); Alice Xiang, *Being "Seen" Versus "Mis-Seen": Tensions Between Privacy and Fairness in Computer Vision*, 36 HARV. J.L. & TECH. 1, 45-49 (2022).

Reducing competition. Consent-based privacy controls also frustrate competition in the digital marketplace. The companies that are in the best position to collect and manage consents and to combine a variety of types of data are the largest companies that already dominate their markets—Google and Amazon, for example.[82] This can be understood as another collective-action problem:[83] Consumers as a whole know that they would benefit in the long run from more companies in more robust competition for their attention and money. But on the margin, no data subject would volunteer to spread their data around and subject themselves to the risks of misuse if the goal of greater competition is likely to fail due to the inaction of the other data subjects.

These problems of collective inaction are mitigated when data users are allowed to operate in a limited zone of freedom where their activities can proceed as long as they aren't likely to cause harm—a classic liability rule rather than a property rule.

C. *Speech as a Quintessential Liberty Zone*

Debates about data-privacy laws are so steeped in the language of consumer protection and digital markets that they obscure the fact that data privacy is a direct restriction on information, and on the means of production of knowledge.[84] In other words, privacy laws are speech restrictions. Modern free-speech law is rooted in the theory that speech and information create thorny collective-action problems: the benefits of free speech are often amorphous, hard

[82] James Campbell, Avi Goldfarb & Catherine Tucker, *Privacy Regulation and Market Structure*, 24 J. ECON. & MGMT. STRATEGY 47, 47 (2015); Garrett A. Johnson, Scott K. Shriver & Samuel G. Goldberg, *Privacy and Market Concentration: Intended & Unintended Consequences of the GDPR*, 69 MGMT. SCI. 5695, 5698 (2023); Christian Peukert, Stefan Bechtold, Michail Batikas & Tobias Kretschmer, *European Privacy Law and Global Markets for Data* 2 (Ctr. for L. & Econ., Working Paper No. 01/2020, 2020); Jia, Zhe Jin & Wagman, *supra* note 69, at 10; Ran Zhuo, Bradley Huffaker, KC Claffy & Shane Greenstein, *The Impact of the General Data Protection Regulation on Internet Interconnection*, 45 TELECOMMS. POL'Y (2020), https://ssrn.com/abstract=3761288 [https://perma.cc/FD37-SF34]; Zhao, *supra* note 2.

[83] *See generally* MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS 1-3 (1971) (discussing the mechanisms by which collective action work, where even "rational, self-interested individuals will not act to achieve their common or group interests").

[84] *See, e.g.*, Jane Bambauer, *Is Data Speech?*, 66 STAN. L. REV. 57, 87-91 (2014); Alan K. Chen, *Cheap Speech Creation*, 54 U.C. DAVIS L. REV. 2405, 2429-32 (2021); Komal S. Patel, Note, *Testing the Limits of the First Amendment: How Online Civil Rights Testing Is Protected Speech Activity*, 118 COLUM. L. REV. 1473, 1488 n.100 (2018); Sorrell v. IMS Health Inc., 564 U.S. 552, 570 (2011); Western Watersheds Project v. Michael, 869 F.3d 1189, 1195-96 (10th Cir. 2017).

to predict in advance, and spread across large numbers of people, and constitutional scrutiny helps counterbalance that tendency to undervalue it.[85]

The First Amendment effectively requires a risk-management approach to the regulation of speech. Outside of intellectual property, which has a distinct claim to matching the theory of property rights and its own constitutional basis for doing so,[86] speech has been treated as a special activity that should be constrained only when the harms are serious and nonspeculative.[87]

Consider the rules allowing plaintiffs to sue for defamation. Defamation law permits the subject of a communication to sue the speaker if the communication is false, impugns the character or reputation of the subject, and is made to a third party with the requisite level of fault.[88] The real-world harm that can be caused from spreading lies about a person are obvious. A defamatory lie harms both the subject of the lie and the listeners who may be duped by them. While harm can also occur from spreading true statements about a person, as a rough rule of thumb, we might expect that reputation damage is more unjust, and therefore more harmful, when the harsh judgments of character are based on fabrications. In other words, there is clear, concrete harm when actors engage in spreading falsehoods. And yet, under the pressure of constitutional law and the logic of tort law, a claim for defamation is highly constrained. A plaintiff cannot simply say, "This information pertains to me, ergo I can demand the removal and deletion of the publication and compensation for my lost reputation."

Defamation law looks nothing like a property claim. A plaintiff has to prove that the defendant's disclosure was false, was published with at least negligence with respect to its veracity, and caused concrete harm (in most cases).[89] Even then, defendants can raise several legal privileges that operate to ensure that frank conversations and normal operations are not impeded by fear of defamation liability. A person or company is privileged to share information, even if it's false, about another person if it is done in the course of an official proceeding,[90] if the information is offered in self-defense,[91] if they are warning others of

[85] *See* Jane Bambauer & Derek Bambauer, *Information Libertarianism*, 105 CALIF. L. REV. 335, 377-81 (2017); Seth F. Kreimer, *Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link*, 155 U. PA. L. REV. 11, 38 n.77 (2006).

[86] *See* Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 349-50 (1991).

[87] United States v. Alvarez, 567 U.S. 709, 717 (2012).

[88] RESTATEMENT (SECOND) OF TORTS § 558 (Am. L. Inst. 1977).

[89] *Id.*

[90] Walters v. Linhof, 559 F. Supp. 1231, 1234 (D. Colo. 1983).

[91] Mencher v. Shesley, 85 N.Y.S.2d 431, 434 (1948).

danger, if they have a common interest with the listener,[92] if the disclosure is in the public interest[93] or if the disclosure occurs entirely within a corporation.[94]

Thus, when it comes to falsehoods, the law is about as far from a property frame as one could get: not only do the subjects of falsehoods have no veto power over when they are being discussed, but the liability rule that applies to the information sharers is narrow, crafted with a good deal of concern about chilling information flows to the detriment of everyone. Defamation law has recognized several safe-harbor privileges that ensure speech activities are not chilled when the free flow of information is net beneficial, even if it isn't perfectly beneficial for each and every person.

Defamation is not the best model for privacy law, but it is a useful guide because it showcases the theoretical, practical, and constitutional reasons to avoid assigning property rights in information.

D. Case Study: Facebook

Many of the problems I have described with treating personal information as property can be seen in the political crosswinds that are jostling the major U.S. technology companies. The demand for property-style privacy law spiked shortly after the revelation that Facebook had allowed third-party companies like Cambridge Analytica to collect a rich trove of Facebook user data (as well as some more basic information about the users' Facebook friends).[95] The immediate legal response, including the voter initiative that brought the California Consumer Privacy Act (CCPA) into being, was to create sticky property rights for data subjects so that data could not be disclosed or used for a new commercial purpose without a renewed and salient consent procedure.[96]

However, the intervening years have also brought a good deal of concern that the largest technology companies, including Facebook, were amassing such a rich trove of personal data about their users that startup companies would not be able to compete—hence efforts in Europe and the United States to affirmatively require digital platforms to make user data available to third-party

[92] *In re* Teleglobe Commc'ns Corp., 493 F.3d 345, 363-64 (3d Cir. 2007).

[93] Brown v. Hearst Corp., 862 F. Supp. 622, 627 (D. Mass. 1994).

[94] Simpson v. Mars, Inc., 929 P.2d 966, 969-70 (Nev. 1997).

[95] Jane Bambauer, *Cambridge Analytica and the Meaning of Privacy Harm*, PROGRAM ON ECON. & PRIV. 1 (2019), https://pep.gmu.edu/wp-content/uploads/sites/28/2019/01/Bambauer_PEP_White_Paper_Cambridge_Analytica.pdf [https://perma.cc/39DQ-6WTW].

[96] Katy Murphy & Steven Overly, *California Demands Facebook Records for Consumer Privacy Investigation*, POLITICO (Nov. 6, 2019), https://www.politico.com/states/california/story/2019/11/06/california-demands-facebook-records-for-consumer-privacy-investigation-1226526 [https://perma.cc/RM3Q-WSFP].

companies.[97] There has also been an increased understanding that small content producers are dependent, to varying degrees, on behavioral-advertising revenues, and that the costs of consent and high friction affect a large ecosystem of journalists and entertainment firms.[98] A risk-based legal rule would avoid these problems by forcing lawmakers and judges to be more honest and concrete about collective priorities when consumers' goals and interests are in tension.

[97] *See The Digital Markets Act: Ensuring Fair and Open Digital Markets*, EUR. COMM'N, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en [https://perma.cc/D566-C5XE]; American Innovation and Choice Online Act, S. 2992, 117th Cong. (2021).

[98] With the exception of one study, the empirical research shows that websites would lose between thirty-eight and sixty-six percent of their advertising revenues if behavioral advertising is banned. The majority of publishers would lose revenue. Avi Goldfarb & Catherine E. Tucker, *Privacy Regulation and Online Advertising*, 57 MGMT. SCI. 57, 68 (2011) (finding a sixty-five percent reduction in revenue, with the assumption that advertisers reduce their expenditure in line with the decrease in ad effectiveness); Howard Beales & Jeffrey A. Eisenach, *An Empirical Analysis of the Value of Information Sharing in the Market for Online Content* 8-9 (Navigant Econ., 2014), https://ssrn.com/abstract=2421405 [https://perma.cc/WHX3-8D2Q] (finding advertisers are willing to pay at least sixty percent more for advertisements informed by user behavior); Garrett A. Johnson, Scott K. Shriver & Shaoyin Du, *Consumer Privacy Choice in Online Advertising: Who Opts Out and at What Cost to Industry?*, 39 MKTG. SCI. 33, 33 (2020) (finding a fifty-two percent reduction in revenue); Deepak Ravichandran & Nitish Korula, *Effect of Disabling Third-Party Cookies on Publisher Revenue*, GOOGLE 1 (2019) (finding a sixty-four percent reduction in revenue); *The Value of Personalized Ads to a Thriving App Ecosystem*, META (June 18, 2020), https://developers.facebook.com/blog/post/2020/06/18/value-of-personalized-ads-thriving-app-ecosystem [https://perma.cc/48MY-7XKN] (finding a fifty percent reduction in revenue); Koen Pauwels, *What's a Cookie Worth Anyway?, Smarter Marketing Gets Better Results*, ANALYTIC DASHBOARDS (June 28, 2021), https://analyticdashboards.wordpress.com/2021/06/28/whats-a-cookie-worth-anyway [https://perma.cc/T48U-VQJX] (finding a 38.5% reduction in revenue). The only study that found a lower figure was based on a single high-value publisher and assumed that advertisers would still have access to a user's geolocation and device information. Veronica Marotta, Vibhanshu Abhishek & Alessandro Acquisti, *Online Tracking and Publishers' Revenues: An Empirical Analysis* 6 (Workshop on Econ. of Info. Sec., 2019), https://weis2019.econinfosec.org/wpcontent/uploads/sites/6/2019/05/WEIS_2019_paper_38.pdf [https://perma.cc/9F9H-KNWZ]. Moreover, ad-blocking and interference with tracking-based advertising diminishes the quantity and quality of ad-supported websites. Benjamin Shiller, Joel Waldfogel & Johnny Ryan, *The Effect of Ad Blocking on Website Traffic and Quality*, 49 RAND J. ECON. 43, 47 (2018) (showing that ad-blocking software, which decreases the effectiveness of advertising in ways that would have a similar revenue impact to a ban on targeted advertising, caused the quality of ad-supported websites to decrease). *But see* Vincent Lefrere, Logan Warberg, Cristobal Cheyre, Veronica Marotta & Alessandro Acquisti, *The Impact of GDPR on Content Providers: A Longitudinal Analysis* 40 (2022), https://ssrn.com/abstract=4239013 [https://perma.cc/7ZVV-GKUB] (finding no significant reduction in the amount of content produced by EU-based websites as compared to US-based websites).

We can see some of the logic of a torts framework refracted through the facts of *In re Facebook*.[99] Facebook was tracking the web-browsing behavior of Facebook users when they visited websites with an imbedded Facebook "like" button. The websites cooperated with the practice because they got an advertising boost if visitors clicked the "like" button,[100] and Facebook of course got access to more particularized user-behavior data that it could leverage in its ad-exchange business.[101] Perhaps the third-party websites provided notice in their privacy policies,[102] but the practices were not made salient. As a result, Facebook users were subjected to cross-site tracking without realizing it.[103]

[99] 956 F.3d 589 (9th Cir. 2020).

[100] "Clicking the LIKE button on the upper right-hand side of a Business Page serves two purposes. For a business, this is very important information to them, as it allows them to show the number of followers that they've gained utilizing social media, thereby tracking their Social Media ROI (Rate Of Influence). Secondly, although with Facebook's new algorithm it's not a guarantee, it most likely will boost the chance that you'll get updates, event notifications, and a deeper connection with that person or company." Dale Griffen, *The Importance of the Facebook "Like"*, GO! AGENCY, https://gosalesandmarketing.com/the-importance-of-the-facebook-like [https://perma.cc/Y8FK-9VBE].

[101] The browsing history data was subsequently used to help Facebook in its targeted advertising business, as Facebook was able to attract more advertisers and command higher ad-placement prices if it could promise the ad would reach a more relevant audience (that is, an audience more likely to click on the ad and make a purchase). *See* Brett R. Gordon, Florian Zettelmeyer, Neha Bhargava & Dan Chapsky, *A Comparison of Approaches to Advertising Measurement: Evidence from Big Field Experiments at Facebook*, 38 MKTG. SCI. 193, 195-96 (2019). The opinion says that Facebook creates behavioral profiles that are "sold to advertisers," *In re* Facebook, 956 F.3d at 596, but in fact, advertisers describe their target audience to Facebook without receiving any personal data, *see* META, AUDIENCE AD TARGETING, https://www.facebook.com/business/ads/ad-targeting [https://perma.cc/L49M-XY6F].

[102] Many websites have privacy policies that describe how data is collected, shared, and used. The California Online Privacy Protection Act of 2003 requires websites to disclose "whether other parties may collect personally identifiable information about an individual consumer's online activities over time and across different Web sites." CAL. BUS. & PROF. CODE § 22575(b)(6) (West 2023). Facebook's privacy policy affirmatively misled users, stating that users who were logged out would *not* be tracked, but tracking of those users did occur. Davis v. Facebook, Inc., 956 F.3d 589, 602-03 (9th Cir. 2020). It is not clear from the articulation of facts in the case whether the third-party websites provided accurate notice or not. Because of Facebook's misleading notice, the case against it should be stronger than the case against a third-party website that either accurately describes its practices in a privacy policy or makes no promises at all.

[103] The practice is commonplace because the predominant business model for websites is a modern, highly tailored variant of the magazine or broadcast model: internet companies offer elaborate and popular content and services for cheap or free, and they fund their operations through behavioral advertising. This is one of the most common business models for online content. Bernard Marr, *The 7 Most Successful Business Models of the Digital Era*, FORBES (Mar. 14, 2023, 3:30 AM EDT), https://www.forbes.com/sites/bernardmarr/2023/03/14/the-7-most-successful-business-models-of-the-digital-era [https://perma.cc/N8W3-PVTY].

To proponents of property-style privacy rights these facts constituted a straightforward violation of privacy rights. Facebook engaged in an unconsented observation and collection of personal data where its users weren't expecting it, and this interfered with Facebook users' exclusive control over that data. Case closed.

Tort principles, by contrast, required a different sort of analysis. Applying the intrusion upon seclusion tort,[104] the court's opinion made clear that the plaintiffs first had to prove that the personal information the defendant collected was private ("seclusion") to have any chance of recovery.[105] This requires plaintiffs to prove that the defendant's observations and data collection violated the norms and reasonable expectations of the observed.[106] It is not always going to be a straightforward inquiry because, unlike norms that have developed in physical space where real property and architectural features can double as markers of expectations, expectations are less visible in digital space.[107]

But even if the data *is* private, the defendant could still prevail if the collection was not unreasonable ("highly offensive").[108] This element essentially acknowledges that making observations and using information is an activity people are generally permitted to do. Thus, the consequences of observing or using somebody's personal information have to be significant in order for courts to curtail that freedom. The way the court put it:

> "[P]laintiffs must show more than an intrusion upon reasonable privacy expectations. Actionable invasions of privacy also must be 'highly offensive' to a reasonable person, and 'sufficiently serious' and unwarranted so as to constitute an 'egregious breach of the social norms.'" Determining whether a defendant's actions were "highly offensive to a reasonable person" requires a holistic consideration of factors such as the likelihood

[104] RESTATEMENT (SECOND) OF TORTS § 652B (AM. L. INST. 1977) ("One who intentionally intrudes, physically or otherwise, upon the solitude or seclusion of another or his private affairs or concerns, is subject to liability to the other for invasion of his privacy, if the intrusion would be highly offensive to a reasonable person.").

[105] *Davis*, 956 F.3d at 601.

[106] "We first consider whether a defendant gained 'unwanted access to data by electronic or other covert means, in violation of the law or social norms.' To make this determination, courts consider a variety of factors, including the customs, practices, and circumstances surrounding a defendant's particular activities. Thus, the relevant question here is whether a user would reasonably expect that Facebook would have access to the user's individual data after the user logged out of the application." *Id.* at 601-02 (quoting Hernandez v. Hillsides, Inc., 211 P.3d 1063, 1072 (Cal. 2009)).

[107] *See* Adam Pabarcus, *Are "Private" Spaces on Social Networking Websites Truly Private? The Extension of Intrusion Upon Seclusion*, 38 WM. MITCHELL L. REV. 397, 400 (2011).

[108] *Davis*, 956 F.3d at 601.

> of serious harm to the victim, the degree and setting of the intrusion, the intruder's motives and objectives, and whether countervailing interests or social norms render the intrusion inoffensive. While analysis of a reasonable expectation of privacy primarily focuses on the nature of the intrusion, the highly offensive analysis focuses on the degree to which the intrusion is unacceptable as a matter of public policy. [109]

The guiding light is risk management. When courts ask how the plaintiff might be harmed by the alleged practices, what the defendant intended to do with the information, and what society gets out of the whole affair, they are recognizing that the legitimate interests of defendants and others are *also* at stake, and any disagreements that arise must be managed without giving a veto or exclusive control to anyone.

III. A PRACTICAL GUIDE TO RISK-BASED PRIVACY LAW

Privacy is part of a larger social contract.[110] American privacy law has been stuck in a perennial state of contestation because it is part of a complex set of trade-offs and coordinated actions that data subjects have with each other, with industry, and with the government.[111] While it may very well be that most people prefer, in general, and all else being equal, to have control over their data, these abstract preferences say little about where legal rights and obligations should be drawn in real-world contexts, where personal interests in privacy come into conflict with other pressing or pragmatic concerns. When control-based privacy rights come at a cost to threat detection, machine-learning applications, or even consumer convenience, "all else" is not equal, and consumers will be better off in many scenarios without weighing in on data processing.

[109] *Id.* at 606 (citations omitted).

[110] Helen Nissenbaum, *A Contextual Approach to Privacy Online*, DAEDALUS: J. AM. ACAD. ARTS & SCIS. 32, 35 (2011) (arguing that privacy should not be regarded as a rigid "procedural mechanism divorced from the particularities" of a data practice).

[111] The works of Helen Nissenbaum and Kirsten Martin have helped add definition to the idea of privacy as an intricate social contract where consent is just one possible route (neither necessary nor automatically sufficient) to achieving privacy. *See, e.g.*, Martin, *supra* note 48; Kirsten E. Martin, *Diminished or Just Different? A Factorial Vignette Study of Privacy as a Social Contract*, 111 J. BUS. ETHICS 519 (2012); Helen Nissenbaum, *Privacy as Contextual Integrity*, 79 WASH. L. REV. 119 (2004); HELEN NISSENBAUM, PRIVACY IN CONTEXT (2010). *See also* Mary J. Culnan & Robert J. Bies, *Consumer Privacy: Balancing Economic and Justice Considerations*, 59 J. SOC. ISSUES 323, 323 (2003) (discussing consumer-privacy concerns and their relationship to the "perceived fairness of corporate information practices"); J. (Hans) Van Oosterhout, Pursey P. M. A. R. Heugens & Muel Kaptein, *The Internal Morality of Contracting: Advancing the Contractualist Endeavor in Business Ethics*, 31 ACAD. MGMT. REV. 521, 522 (2006) (exploring integrative social contracts theory and the "internal morality of contracting").

The HEW Report stated that a respect for privacy requires data collectors to anticipate and behave responsibly when conflicts arise between their interests and those of the data subjects. Some data uses are in the mutual interests of the controllers and the data subjects, some are in their mutual interests but are "not perceived as such," and some are in direct conflict.[112] This observation was ahead of its time and explains why privacy rules are so difficult to articulate in advance. Add to this third parties' legitimate interests, and it is clear that the interests people have in a data practice sometimes run together and sometimes run against each other, and that data subjects will have a hard time knowing when their welfare is in jeopardy.

A risk- or harm-based approach to American privacy law is the right course. However, because the consent model has dominated privacy discourse for so long, lawmakers and legal scholars have not been focused on designing effective risk-based frameworks.

Risk-based privacy law needs to evolve three categories of practices: (1) per se privacy violations, which are harmful practices that should not be conducted unless the data processor has received clear and well-informed consent (and possibly not even then); (2) safe-harbor data practices, which can be thought of as per se nonviolations and are the data practices that are clearly warranted because of public needs or because of their benefit to the data subject, data processor, or others; and (3) the messy middle category consists of the practices that are neither obviously harmful nor obviously desirable. The legality of the practices in the messy middle should depend on the procedures that are in place to provide notice or transparency, the sensitivity of the data or inferences, the ability of data subjects to avoid the practice if they wish, and the costs and benefits of the practice. Liability for the middle category will depend on a common-law-like process that sorts new use cases into the first two categories (per se violations or safe harbors).[113]

A. *Per Se Violations*

Per se privacy violations involve observations, disclosures, or uses of data that are nearly universally unwanted and disturbing or are unnecessary for the welfare of the data subject and the community. Lawmakers have a head start in identifying per se violations based on existing privacy torts, rules, and statutory laws that have stood the test of time. Examples of established per se privacy

[112] HEW Report, *supra* note 33, at 46.

[113] Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 638-43 (2014) (describing a common-law-like process the FTC has used to establish certain minimum requirements to avoid engaging in unfair practices).

violations include intrusive observations or recordings of private spaces and conversations, including upskirt photography,[114] wiretapping phone lines and other private communication channels,[115] and the needless infliction of embarrassment.[116]

A few themes can be discerned from this collection of longstanding privacy violations. First, the collection of information from uninvited observers will cause people to have less candor with one another, so it is imperative to define some spaces and contexts that are private—where those within the seclusion zone can be assured they are not monitored by outsiders, and where outsiders have effective notice that the usual freedoms to observe and record are not available unless they are invited in. Second, while the zones of seclusion will not be easy to define in all cases, there are some areas and contexts (e.g., bathroom stalls) for which there is near-universal agreement. And third, the law can and should recognize when personal information is disclosed to an audience that will foreseeably take advantage of vulnerable data subjects or will foreseeably overreact and retaliate against them.

I suspect other categories of per se violation could be added to a new statute without much controversy based on these principles. For example, uses of data that are purely extractive and designed to facilitate fraud or addiction could be recognized as per se violations.[117] The indiscriminate publication of large amounts of private information (for no apparent public purpose) may be another. A firm's knowing or reckless noncompliance with its own privacy policies could also be considered a per se violation, with statutory damages or actual damages awarded depending on the firm's mental state.[118] Legislators or regulators could also prohibit the use of data that has the purpose or the unjustified effect of discriminating against protected classes of individuals. For example, using personal data to infer an employment applicant's race, pregnancy status, or

[114] The Video Voyeurism Prevention Act, 18 U.S.C. § 1801 (2018).

[115] The Wiretap Act, 18 U.S.C. §§ 2510-20 (2018).

[116] RESTATEMENT (SECOND) OF TORTS § 652D (AM. L. INST. 1977) ("Publicity Given to Private Life"). This tort may be too narrow, though, by failing to capture embarrassing revelations that do not rise to the level of "publicity" in terms of its dissemination but that nevertheless reach an audience with the result of pure humiliation, without counteracting social benefits. *See* Hanin, *supra* note 48, at 248 (suggesting a duty not to inflict "gratuitous embarrassment").

[117] *See, e.g.*, Maddy Varner & Aaron Sankin, *Suckers List: How Allstate's Secret Auto Insurance Algorithm Squeezes Big Spenders*, MARKUP (Feb. 25, 2020, 5:00 ET), https://themarkup.org/allstates-algorithm/2020/02/25/car-insurance-suckers-list [https://perma.cc/347N-A893]; Gema de las Heras, *Scammers Target Disaster Victims. Spot Their Traps*, FTC (Aug. 2, 2022), https://consumer.ftc.gov/consumer-alerts/2022/08/scammers-target-disaster-victims-spot-their-traps [https://perma.cc/9JWW-LFY4].

[118] *Cf.* Solove & Hartzog, *supra* note 113, at 638-43 (discussing the development of various existing theories for what constitutes an unfair trade practice).

religious affiliation in order to affect the chance of hiring would be a violation no matter what type of data was used, how it was collected, or how it was processed.[119]

B. Safe Harbors (or Privileges, or Per Se Nonviolations)

Existing statutes and precedents can also supply a partial list of data practices that the law has historically *exempted* from privacy-related restrictions. These should be treated categorically as nonviolations in a risk-based system. Data practices that have become commonplace and that could not be forbidden without a significant shock to popular digital-media services may also be good candidates for legal privileges. As a general rule, safe harbors should be created if there would be broad agreement, even if not universal agreement, among well-informed observers that a data practice is good for society on balance.[120] Together, a set of safe harbors can establish a zone of liberty for data processors and innovators.

What follows is a starter set of data collection and processing safe harbors based on common privacy-law exceptions, free-speech-related privileges, and common industry practices. They are listed from least to most controversial.

1. With Consent

A data practice done with the knowledge and voluntary consent of the data subject should be considered a per se nonviolation, as long as the practice is not on the list of per se violations. Few would find this controversial since this exception replicates the key feature of popular privacy-law proposals—data-subject control. Operationalizing consent is another matter, though. There will be ambiguities over whether a data subject has sufficient knowledge about the bargain and whether the consent of a data subject is voluntary or performed with an unacceptable level of duress. Data-use disclosures that are buried in an end-user agreement may not constitute evidence of "knowledge" unless the practices

[119] *See* Solove, *supra* note 51 (manuscript at 46). *See generally* Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671 (2016) (advocating for a disparate-impact doctrine for victims of discriminatory data mining).

[120] What it means to be "good for society" could be thought of as a contractualist position on privacy rules. *See generally* Richard M. Re, *Fourth Amendment Fairness*, 116 MICH. L. REV. 1409 (2018) (proposing a contractualist approach to Fourth Amendment privacy). Alternatively, what is "good for society" could be understood from behind the Rawlsian veil of ignorance. *See generally* Akira Inoue, Masahiro Zenkyo & Haruya Sakamoto, *Making the Veil of Ignorance Work: Evidence from Survey Experiments*, *in* OXFORD STUDIES IN EXPERIMENTAL PHILOSOPHY 53 (Tania Lombrozo, Joshua Knobe & Shaun Nichols eds., 2021) (describing the Rawlsian theory and experimental studies).

are also well publicized or generally known to data subjects. And consent that is obtained in circumstances where the data subject has little choice—such as at a hospital or with an employer—could also fall short of the proper definition of consent. But these issues are not intractable, and they are not unique to the privacy context. Consent doctrines in battery,[121] intellectual property,[122] and police searches[123] can serve as models.

The government could simplify and encourage effective consent by developing voluntary labeling schemes that help firms quickly signal the sorts of data practices the firm has committed to using (or not using). This would allow firms to compete on privacy more efficiently as a salient feature of their services.[124] But the important thing, for this Essay, is that consent offers just one of many routes for companies to avoid legal complications.

2. *For the Direct Benefit of the Data Subject*

When a wallet is returned to its owner, the owner will not resent the Good Samaritan who looked inside to find an ID. The same will be true when personal data is used to locate individuals suffering from a mental-health crisis, track down displaced children, find individuals for the purpose of relaying payments to them (such as class-action settlements, child-support payments, or refunds), provide warnings about known or credible threats to health or safety, or complete forms or bypass red tape for a transaction that the data subject initiated. To generalize, services performed under the reasonable and good-faith belief that the services will assist and benefit the data subject should be per se nonviolations of privacy. They can be analogized to the tort doctrine of "presumed consent."[125]

[121] *See* Kenneth W. Simons, *A Restatement (Third) of Intentional Torts?*, 48 ARIZ. L. REV. 1061, 1066-78 (2006).

[122] *See* Mark R. Patterson, *Must Licenses Be Contracts? Consent and Notice in Intellectual Property*, 40 FLA. ST. U. L. REV. 105, 115-16 (2012).

[123] *See* Alafair S. Burke, *Consent Searches and Fourth Amendment Reasonableness*, 67 FLA. L. REV. 509, 516-25 (2015).

[124] *See generally* Sarah Holland, Ahmed Hosny, Sarah Newman, Joshua Joseph & Kasia Chmielinski, *The Dataset Nutrition Label: A Framework to Drive Higher Data Quality Standards*, *in* DATA PROTECTION AND PRIVACY: DATA PROTECTION AND DEMOCRACY 1 (Dara Hallinan, Ronald Leenes, Serge Gutwirth & Paul De Hert eds., 2020) (proposing a "Dataset Nutrition Label" to assess and publicize data-analysis practices).

[125] *See* Restatement (Third) of Torts: Intentional Torts to Persons § 16 (Am. L. Inst., Tentative Draft No. 4, 2019).

3. *For Self-Protection or the Protection of Others*

Occasionally, the data subject is an aggressor who is attempting to use deceit to harm others.[126] In these situations, the past or future victims of their deceit have more compelling interests in discovering the misrepresentation than the data subject has in hiding it.[127] Data should be able to be used without consent for detecting or warning about fraud, criminal activity, or other misbehavior, and for complying with federal or local "Know Your Customer" laws.[128] To be sure, detecting fraud and crime requires the analysis of the data of many innocent individuals. But when access and analysis of personal data is done for the purpose of exposing misconduct, the data processor does not need consent.

For example, Apple once had plans to check all images uploaded to iCloud against the images of known child pornography.[129] If (and only if) an accountholder's photos produce ten matches, the software would have

[126] Crime takes a large toll on U.S. residents, imposing as much as a trillion dollars in losses per year, according to one estimate. *See* Aaron Chalfin, *The Economic Costs of Crime*, *in* THE ENCYCLOPEDIA OF CRIME AND PUNISHMENT 543, 551 (Wesley G. Jennings ed., 2016). Crime flourishes in low-information environments where the perpetrator's identity or activity cannot be observed. Such crimes are, in other words, crimes of opportunity that exploit low information. *See generally* Eric L. Piza, Brandon C. Welsh, David P. Farrington & Amanda L. Thomas, *CCTV Surveillance for Crime Prevention: A 40-Year Systematic Review with Meta-Analysis*, 18 CRIMINOLOGY & PUB. POL'Y 135 (2019) (demonstrating that surveillance cameras are associated with crime reduction).

[127] *See, e.g.*, NISSENBAUM, *supra* note 46, at 178 (noting that in the employment context, "[i]t is clear why presentation of self would be important for applicants, but just as clear why companies might resist applicants' claims trumping their own"); Posner, *supra* note 78, at 401 ("To the extent that people conceal personal information in order to mislead, the economic case for according legal protection to such information is no better than that for permitting fraud in the sale of goods."); Posner, *supra* note 78, at 403 ("[T]here is a prima facie case for assigning the property right [to personal data] *away* from the individual where secrecy would reduce the social product by misleading the people with whom he deals." (emphasis added)).

[128] For what it is worth, a survey of privacy professionals conducted by the Ponemon Institute at Accenture found that most privacy professionals agree it is acceptable to use personal information to identify and authenticate customers, to share information with law enforcement, for fraud prevention, and for government and national-security purposes. *How Global Organizations Approach the Challenges of Protecting Personal Data*, PONEMON INST. 23 (2009), https://www.ponemon.org/local/upload/file/ATC_DPP%20report_FINAL.pdf [https://perma.cc/Q8ZG-EXZZ].

[129] Apple had planned to use a function that would convert an image into one particular string of numbers, referred to as hashes. This would have allowed Apple to check the hash of every image against a library of hashes that came from known Child Sexual Abuse Material (CSAM). If the hashes matched, or nearly matched, Apple could have confidence that they have detected child pornography. *See generally CSAM Detection: Technical Summary*, APPLE (2021), https://www.apple.com/child-safety/pdf/CSAM_Detection_Technical_Summary.pdf [https://perma.cc/BYX4-8D4W] (describing how the system works).

automatically alerted Apple employees so they could share the information with authorities. Apple has since abandoned its plans in response to criticisms and concerns related to privacy.[130] The safe harbor I suggest here would allow Apple to proceed with the program without the threat of legal penalties because the purpose of the program is to detect threats to third parties. Of course, this would not *require* Apple or any other firm to scan for threats to third parties. But if they choose to do so, that act would be immunized from privacy-related suits.

4. *For the Purposes of Statistical Research or Internal Research and Development*

Most privacy laws allow data processors to use data for internal research, product improvement, and new-product development, as well as for general statistical-research purposes.[131] They also allow data processors to prepare deidentified versions of the data and share them to other researchers. These research uses of personal data often fall outside the statutory definitions of "personal data" because the data is expected to be used in a manner that does not directly link back to the individual data subjects. The reasons this exception to privacy-related liability is controversial at all is that there is significant concern over the potential for deidentified data to be reidentified, or to be used in a manner that is highly stigmatizing for a particular identity group.[132] Thus, the crafting of this safe harbor will depend on a thoughtful approach to make sure data processors are taking reasonable efforts to prevent the reidentification of data subjects.

5. *Cooperating with Civil, Criminal, or Regulatory Investigations*

Every privacy statute has an exemption for firms that respond to subpoenas, summonses, or warrants as long as the judicial or law-enforcement requests are

[130] *See* Lily Hay Newman, *Apple Kills Its Plan to Scan Your Photos for CSAM. Here's What's Next*, WIRED (Dec. 7, 2022, 1:11 PM), https://www.wired.com/story/apple-photo-scanning-csam-communication-safety-messages [https://perma.cc/VR5D-6VSG].

[131] *See, e.g.*, 45 C.F.R. § 164.512 (2023) (enacting a HIPAA regulation on the uses or disclosure of personal health information for which authorization is not required).

[132] Jules Polonetsky, Omer Tene & Kelsey Finch, *Shades of Gray: Seeing the Full Spectrum of Practical Data De-identification*, 56 SANTA CLARA L. REV. 593, 609-21 (2016). I have argued that the malicious reidentification of the subject in a deidentified database that was prepared for purely statistical purposes should be prohibited by law. Yakowitz Bambauer, *supra* note 80, at 48. This conduct would be another example of a possible per se violation.

consistent with Fourth Amendment law and other applicable statutes.[133] Legal scholars have been critical of these exemptions,[134] but they are part of an enduring trade-off between privacy and public safety. At its best, data can be used by police not only to detect or deter a perpetrator of a serious crime, but also to clear a suspect or exonerate a criminal defendant.[135] Thus, it probably makes sense to provide a law-enforcement safe harbor in a generally applicable privacy law, and to encourage lawmakers to place appropriate restrictions on law-enforcement access through more targeted legislation.

6. *For Matching, with the Direct Participation of the Data Subject*

Finding reliable information about potential clients, customers, or business partners is a market transaction cost that can frustrate matching between two sides of a market or search process. [136] Just as consumers often need information about businesses to have confidence that they know enough about the quality of goods and services, businesses, too, need information to find their customers and clients.

Sometimes, this matching is performed with the proactive participation of the data subject, as when the data subject applies for a loan, seeks admission to a school, or enters search terms into a flight aggregator website or a search engine. When the data subject initiates a matching process, the data processor responding to the request should be permitted to use independently sourced data—that is, information that goes beyond what the applicant has supplied—in order to find the best match between the applicant and the supplied product, service, or content.

Let us use lending as an example, since the matching process in this market is more familiar than other matching practices. When a lender uses a credit report to make lending decisions, these independent sources of information can

[133] Communications Assistance for Law Enforcement Act, 47 U.S.C. §§ 1001-1010 (2018) (requiring telecommunications providers to design equipment to facilitate targeted surveillance for law enforcement); Erin Murphy, *The Politics of Privacy in the Criminal Justice System: Information Disclosure, the Fourth Amendment, and Statutory Law Enforcement Exemptions*, 111 MICH. L. REV. 485, 485 (2013).

[134] Murphy, *supra* note 133, at 503-07.

[135] *See* Joshua A.T. Fairfield & Erik Luna, *Digital Innocence*, 99 CORNELL L. REV. 981, 981 (2014) (arguing that, although "the fruits of mass surveillance have been used almost exclusively to convict," such data could be used "to prevent wrongful convictions and to provide hard proof of actual innocence").

[136] *See* Joseph E. Stiglitz, *Information*, CONCISE ENCYCLOPEDIA ECON. (1st ed.), https://www.econlib.org/library/Enc1/Information.html [https://perma.cc/E6CL-CXRT] (describing reputation, advertising, and third-party intermediaries as common methods to overcome search costs).

reassure lenders about the creditworthiness of a loan applicant.[137] Credit reports help the *applicants* as much as they help the lender. While there may be some threshold beyond which additional data is not useful for improving matching and performance, society has not reached that threshold.

The value of matching extends well beyond credit markets. In the health sector, for example, advances in machine learning are changing the practice of medicine because new programs can digest and learn from vast amounts of data about both the patient and others to make customized, time-sensitive recommendations.[138] Data-driven medical-adherence scoring systems, which use information about patients to predict whether they are likely to stick with a prescribed treatment, can improve health by better matching patients to services, such as treatments and pharmacy interventions.[139] Some of the benefits of personal-data use can be harnessed with the data subject's consent and active participation, as when a health or wellness app asks users for permission to access data from sensors or other digital services. But this is not always the case. First, managing consent and permissions adds a layer of costs that will often be impractical or financially detrimental. For example, in the United States, healthcare-privacy regulation caused hospitals to slow or stop the adoption of electronic medical records due to the costs of consent and compliance burdens, preventing implementation of this cost-saving technology.[140] The transactions would be even more difficult and costly if every company had to negotiate with each individual data subject over payment and other contract terms. The practical effect of such a rule would mean that all industries would simply have to

[137] John M. Barron & Michael Staten, *The Value of Comprehensive Credit Reports: Lessons from the U.S. Experience*, *in* CREDIT REPORTING SYSTEMS AND THE INTERNATIONAL ECONOMY 273, 273 (Margaret J. Miller ed., 2003).

[138] *See, e.g.*, Basel Kayyali, David Knott & Steve Van Kuiken, *The Big-Data Revolution in U.S. Health Care: Accelerating Value and Innovation*, MCKINSEY & CO. (Apr. 1, 2013), https://www.mckinsey.com/industries/healthcare/our-insights/the-big-data-revolution-in-us-health-care [https://perma.cc/A7EC-HUY3] (discussing Ginger.io, a mobile application that tracks user data to assist with the provision of behavioral-health therapies). Some of the most promising applications would have to merge data from disparate sources both within and outside the health sector, and the incentives to do this are likely to depend on being able to collect, purchase, and reuse personal data. *See* Sonja Marjanovic, Ioana Ghiga, Miaoqing Yang & Anna Knack, *Understanding Value in Health Data Ecosystems*, RAND EUR. 21-22 (2017), https://www.rand.org/content/dam/rand/pubs/research_reports/RR1900/RR1972/RAND_RR1972.pdf [https://perma.cc/FR8Q-7BW6]. For a definition of "machine learning," see M.I. Jordan & T.M. Mitchell, *Machine Learning: Trends, Perspectives, and Prospects*, 349 SCI. 255, 255 (2015).

[139] *See* Inmaculada Hernandez & Yuting Zhang, *Using Predictive Analytics and Big Data to Optimize Pharmaceutical Outcomes*, 74 AM. J. HEALTH-SYS. PHARMACY 1494, 1495 (2017).

[140] Amalia R. Miller & Catherine Tucker, *Privacy Protection and Technology Diffusion: The Case of Electronic Medical Records*, 55 MGMT. SCI. 1077, 1081 (2009).

operate with much less information and would be much more wasteful and higher priced as a result.[141]

To be sure, the effects on individual data subjects will be mixed when personal data is used to match people to scarce resources. Not every data subject will be made better off. Some will receive less favorable treatment (e.g., worse home-mortgage terms) than they would in the absence of secondary sources of personal data. But as long as the purpose of analyzing the data subjects is legitimate (e.g., to match credit-card offers, dating-site users, or search results), access to more complete data will result in more "winners" than "losers," and the average user of the matching service (and society at large) will be well served.[142]

This recommendation does not currently exist in any privacy statute, so far as I am aware. And the concept of matching, which is often performed through algorithmic predictions and scoring, has tended to provoke the suspicion and fear of consumers.[143] Thus, I count this as one of the more controversial recommendations for a safe harbor. Nevertheless, the demonstrated benefits of using additional data to perform better matching (in terms of both reduced error and reduced bias), as compared to alternative methods, are sizable enough to justify it. And that's to say nothing of the free-speech interests involved.[144]

[141] *See* Goldfarb & Tucker, *supra* note 65, at 23.

[142] This is similar to how consumers match themselves to businesses. *Cf.* Michael Luca, *Reviews, Reputation, and Revenue: The Case of Yelp.com* 2, 19 (Harvard Bus. Sch., Working Paper No. 12-016, 2016), https://www.hbs.edu/ris/Publication%20Files/12-016_a7e4a5a2-03f9-490d-b093-8f951238dba2.pdf [https://perma.cc/RHW7-5ANF] (finding that consumers respond more strongly to Yelp reviews of restaurants when those reviews contain more information); Brett Hollenbeck, *Online Reputation Mechanisms and the Decreasing Value of Chain Affiliation*, 55 J. MKTG. RSCH. 636, 636 (2018) (finding that independent hotel revenue has increased as more online review information has become available). A range of social institutions have allowed market transactions to take place even when individuals do not know each other well enough to have longstanding relationships of trust. Word of mouth, gossip, and status signals help create enough of an incentive for trade partners to cooperate and enough of a disincentive against deception. *Cf.* Benjamin Klein & Keith B. Leffler, *The Role of Market Forces in Assuring Contractual Performance*, 89 J. POL. ECON. 615, 616 (1981) ("[E]conomists . . . have long considered 'reputations' and brand names to be private devices which . . . assure contract performance . . . ."); Hongbin Cai, Ginger Zhe Jin, Chong Liu & Li-an Zhou, *Seller Reputation: From Word-of-Mouth to Centralized Feedback*, 34 INT'L J. INDUS. ORG. 51, 64 (2014) (explaining that "seller reputation," which was facilitated by word of mouth before the availability of centralized feedback, "is one of the most important incentives for trade and cooperation"); Carl Shapiro, *Premiums for High Quality Products as Returns to Reputations*, 98 Q.J. ECON. 659, 660 (1983) ("The premiums that reputable firms earn . . . serve a crucial role in inducing such sellers to *maintain* their reputations.").

[143] Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 4-5 (2014).

[144] Sorrell v. IMS Health Inc., 564 U.S. 552, 564-65 (2011); Bambauer, *supra* note 84, at 60-61.

7. *For Personalizing or Targeting Speech, Even Without the Direct Participation of the Data Subject*

Finally, lawmakers should seriously consider recognizing a safe harbor for data collection and processing that is performed for the purpose of tailoring the creation or delivery of speech. What I am talking about here includes some exciting and high-value applications, such as personalized diagnosis and recommendation tools that are emerging in Health AI. [145] But it also includes some of the more controversial practices in the Big Data economy—behavioral advertising and hyperpersonalized social media newsfeeds—which have long attracted the attention and ire of regulators.[146] Yet these practices deserve protection from the threat of litigation outside especially harmful circumstances. First, the use of personal data to tailor or target messaging—including marketing—is fully protected speech under the First Amendment.[147] Also, services often depend on meeting the niche demands of their audiences or on the extra income coming from targeted advertising in order to subsidize the zero-price goods and services that internet users have come to love.[148] A law that makes these popular business models illegal will diminish the quantity and quality of content, perhaps in ways that users cannot fully appreciate. The case that the *targeted* quality of advertising or recommender systems has caused more harm than good to consumers generally has not been substantiated, with the possible exception of the wide-ranging (but also mixed) evidence about special harms of social media to adolescents.[149]

[145] For an overview of applications emerging in Health AI, see PETER LEE, CAREY GOLDBERG & ISAAC KOHANE, THE AI REVOLUTION IN MEDICINE: GPT-4 AND BEYOND (2023). To be clear, the speech itself can be the subject of regulation or litigation when it causes foreseeable harm, as when AI-generated recommendations advise a user to do something dangerous. But the specific acts of collecting and using data for the purpose of providing advice should not be the basis of liability or penalty.

[146] Trade Regulation Rule on Commercial Surveillance and Data Security, 87 Fed. Reg. 51273 (proposed Aug. 22, 2022) (to be codified at 16 C.F.R. ch. 1). For a description of European laws that restrict targeted advertising, see Lex Zard & Alan M. Sears, *Targeted Advertising and Consumer Protection in the European Union*, 56 VAND. J. TRANSNAT'L L. 799, 821-26 (2023).

[147] *Sorrell*, 564 U.S. at 552.

[148] Goldfarb & Tucker, *supra* note 65, at 3, 20; James C. Cooper, Jane R. Bambauer, Joshua D. Wright & John M. Yun, Comment Letter on Accountable Tech Petition for Rulemaking to Prohibit Tailored Advertising (Jan. 26, 2022), https://ssrn.com/abstract=4019697 [https://perma.cc/J6WZ-M3JN].

[149] Indeed, for the general population, there is some evidence that recommender systems help get people out of the harmful or self-destructive information rabbit holes that they would otherwise pursue. *See* Martin Hilbert, Arti Thakur, Feng Ji, Pablo M. Flores, Xiaoya Zhang, Jee Young Bhan & Patrick Bernhard, *8%-10% of Algorithmic Recommendations Are 'Bad', But . . . An Exploratory Risk-Utility Meta-Analysis and Its Regulatory Implications* 16 (Sept. 11, 2023) (unpublished manuscript), https://ssrn.com/abstract=4426783 [https://perma.cc/

* * *

Privacy law (and society at large) would benefit from the delineation of safe harbors. Establishing safe harbors will spur activity and innovations in these zones of liberty and will reduce the economy-wide costs of legal uncertainty and consent rituals. The set of per se nonviolations I have recommended here range from the banal (e.g., for the direct benefit of the data subject) to the more controversial (e.g., for cooperation with law enforcement, or for targeted content and marketing), but all have some claim to legitimacy based on logic, tradition, or constitutional protection.

C. *The Messy Middle*

Between the per se violations and safe harbors resides an indeterminant middle category where a wide range of factors will have to be used to determine whether a particular practice, carried out in its real-world context, causes unjustified harm. In this zone, reasonable minds may disagree on whether a data practice is appropriate, just as they will disagree in difficult cases of negligence about whether the defendant behaved reasonably. In other words, these are the hard cases. But before diving into the factors that will determine whether a privacy violation has occurred, let us first reflect on how much has already been resolved. Many of the data practices that provoke privacy debates—from behavioral advertising to identity theft—have already been determined as either automatically permissible or automatically impermissible. What is left are practices that are not terribly common or that have not yet emerged, and that do not obviously belong in one per se category or the other.

For these hard cases, most privacy experts would home in on a few factors that cut either for or against the recognition of a privacy harm. The analysis of "highly offensive" in the *In re Facebook* case provides a starting place for the factors that should be relevant—the motivation of the data processor, the risk of harm to the data subject, and the impact on third parties, among others.[150] Other

4PSD-G92N]. For the impact on youth, see Kaitlyn Tiffany, *No One Knows Exactly What Social Media Is Doing to Teens*, ATL. (June 13, 2023), https://www.theatlantic.com/technology/archive/2023/06/social-media-teen-mental-health-crisis-research-limitations/674371 [https://perma.cc/735V-YU5C]; and Stuart Ritchie, *Don't Panic About Social Media Harming Your Child's Mental Health—The Evidence Is Weak*, INEWS (Mar. 25, 2023), https://inews.co.uk/news/technology/dont-panic-about-social-media-harming-your-childs-mental-health-the-evidence-is-weak-2230571 [https://perma.cc/A6L8-NP5C] (describing and evaluating the key studies).

[150] *In re* Facebook, Inc. Internet Tracking Litig., 956 F.3d 589, 606 (9th Cir. 2020).

relevant factors include whether the data processor overcollected (recording more data than is likely to be useful for legitimate primary and secondary purposes),[151] used security best practices,[152] and provided salient forms of notice.[153] The costs of obtaining consent (both financial costs and the costs to the utility of the data for the particular purpose) should also be considered, as should the actual and perceived costs and benefits to the individual, to the data processor, and to third parties. Identifying a privacy violation requires a consideration of the *totality* of these factors. The hallmark of a violation is when the privacy risks cannot be justified by the benefits of a practice.[154]

CONCLUSION

This Essay has argued that privacy law, in order to be meaningful and workable in a technologically advanced environment, must be crafted around principles of risk-mitigation rather than data ownership. Data processing should operate in a general zone of permissiveness, with limitations based on the foreseeable risks that a particular practice will create.

This proposal will be foreign to a privacy culture accustomed to advocating for data-subject control, and some may worry that a risk-based privacy regime puts too much faith in the companies that collect and use data—allowing *them* to decide what to do with it instead of the data subject. Viewed this way, a risk-based approach could be mistaken as little more than codified self-regulation. But this is not so. A risk-based privacy framework would *not* leave the scope and meaning of privacy protection to the priorities and whims of data users. Instead, a neutral intermediary—a judge or a federal agency, for example—would craft rules for safe harbors and per se violations, would determine the propriety of other practices in a case-by-case manner, and would assess new data practices and business models as they emerge. Data processors would be no more in control over the definition of "privacy" than manufacturers are in control of the definition of "negligent design."

Nevertheless, there is a deep question hidden in the objection. Who *should* be considered knowledgeable enough and reasonable enough to define which

[151] *See The OECD Privacy Framework*, ORG. FOR ECON. COOP. & DEV. 14 (2013), https://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf [https://perma.cc/Y9J5-SA4B].

[152] *See* Derek E. Bambauer, *Privacy Versus Security*, 103 J. CRIM. L. & CRIMINOLOGY 667, 682-83 (2013).

[153] *See* HEW Report, *supra* note 33, at 57-58.

[154] *See* Letter from Michael Pertschuk, Chairman, Fed. Trade Comm'n, to Wendell H. Ford, U.S. Sen., & John C. Danforth, U.S. Sen. (Dec. 17, 1980), https://www.ftc.gov/legal-library/browse/ftc-policy-statement-unfairness [https://perma.cc/2CEY-7TCJ].

practices are harmful, beneficial, or more-or-less a wash? And on what basis could we feel confident that the decision maker has all the necessary information? There are viable arguments in favor of common-law courts, of legislatures, and of expert agencies like the Federal Trade Commission. I am not sure which has the best case. But current and future debates about *who* should craft privacy rules of reasonable care should not let the main point be obscured: data processing is a presumptively valid, net-positive activity. Safe harbors should be ample enough to cover nearly every low-risk activity involving the collection and use of information. Close cases should require the decision makers—whichever branch of government they are in—to take an accounting of not only the data subject's interests but the interests of processors and third parties, too. And unnecessarily risky practices involving personal data should be forbidden and deterred, no matter how many unthinking click-through consents the data user may have been able to collect.

Brechner Eminent Scholar and Professor of Law, University of Florida. I am tremendously grateful for comments and feedback on an early draft from CJ Pommier, Derek Bambauer, Neil Chilson, Jorge Contreras, James Cooper, Erika Douglas, Barbara Evans, Dan Gilman, Jim Harper, Gus Hurwitz, Ginger Jin, DeBrae Kennedy-Mayo, Geoff Manne, Will Rinehart, Eugene Volokh, Peter Winn, Andrew Woods, Christopher Yoo, and John Yun. I also benefitted greatly from the patience and wisdom of Dorrin Akbari and the editors of the Yale Law Journal Forum.