University of Chicago Law School
Chicago Unbound

Public Law and Legal Theory Working Papers | Working Papers

2021

The Public Trust in Data

Aziz Z. Huq

Follow this and additional works at: https://chicagounbound.uchicago.edu/public_law_and_legal_theory

 Part of the Law Commons

Chicago Unbound includes both works in progress and final versions of articles. Please be aware that a more recent version of this article may be available on Chicago Unbound, SSRN or elsewhere.

Recommended Citation
Aziz Z. Huq, "The Public Trust in Data", Public Law and Legal Theory Working Paper Series, No. 765 (2021).

This Working Paper is brought to you for free and open access by the Working Papers at Chicago Unbound. It has been accepted for inclusion in Public Law and Legal Theory Working Papers by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

The Public Trust in Data

(forthcoming, *Georgetown Law Journal* -- (2021))

Aziz Z. Huq*

Abstract

Personal data is no longer just personal. Social networks and pervasive environmental surveillance via cellphones and the 'internet of things' extract minute-by-minute details of our behavior and cognition. This information accumulates into a valuable asset. It then circulates among data brokers, targeted advertisers, political campaigns, and even foreign states as fuel for predictive interventions. Rich gains flow to firms well positioned to leverage these new information aggregates. The privacy losses, economic exploitation, structural inequalities, and democratic backsliding produced by personal data economies, however, fall upon society at large.

This Article proposes a novel regulatory intervention to mitigate the harms from transforming personal data into an asset. States and municipalities should create "public trusts" as governance vehicles for their residents' locational and personal data. An asset in "public trust" is owed and managed by the state. The state can permit its use, and even allow limited alienation, if doing so benefits a broad public rather than a handful of firms. Unique among the legal interventions proposed for new data economies, a public trust for data allows a democratic polity to durably commit to public-regarding management of its informational resources, coupled to judicially enforceable limits on private exploitation and public allocation decisions. The public trust itself is a common-law doctrine of ancient roots. It was revived in the Progressive Era as an instrument to protect public assets against private exploitation. Both federal and state courts, including the U.S. Supreme Court, have since endorsed a variety of doctrinal formulations. The result today is a rich repertoire of rules and remedies for the management of common property. Personal data, usefully, has many similarities to assets long managed by public trust. And familiar justifications for creation of a public trust logically extend to personal data. Indeed, municipalities in the United States, Europe, and Canada have started to experiment with limited forms of a public trust in data. Generalizing from those experiences, this Article offers a 'proof of concept' for how personal data economies can be leashed through the public trust.

* Frank and Bernice J. Greenberg Professor of Law, University of Chicago Law School. Thanks to Lee Fennell, Sonia Kaytal, Aniel Kovvali, Randy Picker, and Eric Posner—as well participants in Chicago's Works-in-Progress workshop, for comments. All errors are mine. The Frank J. Cicero Fund provided support for research on this paper.

Electronic copy available at: https://ssrn.com/abstract=3794780

Table of Contents

Electronic copy available at: https://ssrn.com/abstract=3794780

Introduction

Personal data is no longer just personal. Social networks, websites, cellphones, and an 'internet of things' extract minute-by-minute details of our behavior and cognition. These are warehoused and circulated among data brokers, advertisers, political campaigns, and even foreign states. As it moves, this data accumulates into a valuable asset. It feeds the machine-learning algorithms that allow Amazon to predict purchases, Netflix to estimate views, and governments to anticipate crime. Its predictions drive interventions such as targeted advertising, prompts to digest political disinformation, or decisions to arrest suspects, bail denial to some, and keep yet more behind bars. Rich rewards flow to firms well positioned to leverage these new information aggregates. Dominant social platforms in the United States today have a market capitalization of more than four trillion dollars.[1] But these new affordances come with a price. The personal data economy's toll is felt in lost privacy, economic vulnerability for workers, swelling structural inequality at the social level, and a drip-fed corrosion of democratic values. The solutions proposed to ameliorate these harms include the creation of individual property rights to personal data and the reinvigoration of antitrust law. But all solutions to date are necessarily partial in ambition. None decisively rewire the growing concentration of wealth and income in dominant firms. None clearly redound to the benefit of all users creating value at the front end.

This Article proposes a novel regulatory intervention to mitigate the harms of personal data economies and to advance the public's privacy, equality, and economic interests. States and municipalities, it contends, should create "public trusts" as governance vehicles for their residents' personal data. An asset in "public trust" is owed and managed by the state subject to judicially enforceable controls on use and alienation. The asset can both be used by the general public or made available for controlled commercial exploitation. Either way, it remains subject to the state's supervening obligation "to protect the people's common heritage"[2] and to ensure it remains in good condition.[3] Uses that yield concentrated returns to a small coterie of individuals or firms are disfavored.[4] When created by legislation or state constitutional provision, a public trust provides a unified vehicle for democratic decision-making over common resources that simultaneously addresses both the risks of private and public abuse. It can also be coupled to judicial enforcement, either by a trustee or a lawyer for the state, to ensure that democratic decisions are durable. While

---

[1] *Stigler Committee on Digital Platforms: Final Report*, U. Chi. Booth Sch. Bus., at 7 (2019), https://www.chicagobooth.edu/-/media/research/stigler/pdfs/digital-platforms---committee-report---stigler-center.pdf [https://perma.cc/4YYG-PS9C] [hereinafter *Stigler Committee Final Report*].

[2] *Nat'l Audubon Soc'y v. Superior Court*, 33 Cal. 3d 419, 441, 658 P.2d 709, 724 (1983).

[3] *See* Joseph L. Sax, *The Public Trust Doctrine in Natural Resources Law: Effective Judicial Intervention*, 68 MICH. L. REV. 471, 477 (1970) (enumerating specific limits in the use of a public trust asset). Sax's article is widely recognized as marking a sea-change in scholarly understandings of the public trust doctrine. "Until it was revived and re-invented by Sax, the doctrine held that some resources, particularly lands beneath navigable waters or washed by the tides, are either inherently the property of the public at large, or are at least subject to a kind of inherent easement for certain public purposes." Carol Rose, *Joseph Sax and the Idea of the Public Trust*, 25 ECOLOGY L. Q. 351, 352 (1998) [hereinafter "Rose, *Idea of the Public Trust*"]. It became instead "a vehicle for insisting that public bodies pay attention to--and adequately vindicate--the changing public interest in diffuse resources." *Id.* at 355.

[4] Erin Ryan, *A Short History of the Public Trust Doctrine and Its Intersection with Private Water Law*, 38 VA. ENVTL. L.J. 135, 161 (2020) ("The public trust's doctrinal infrastructure shows that it doesn't just protect the public nature of these common resources--it also assigns responsibility for their protection--specifically, to the government.").

Electronic copy available at: https://ssrn.com/abstract=3794780

discrete pieces of regulation can be enacted outside the public trust framework, the latter has the advantage of simultaneously creating remedies for private and public abuse, and doing so in a durable fashion.

A public trust for data, I propose, might cover information generated by locational apps, sensing devices, or geotagged social-media platforms within a jurisdiction. That data would not need to be maintained within the jurisdiction or held in government databases. It would, however, be subject to that jurisdiction's regulation. The ensuing trust could permit commercial use on the payment of a user fee, which would then be used for the benefit of the population creating the data. The trust could forbid certain uses of the data—such as the use of photographic images to train facial recognition instruments. And the trust could impose obligations to create epistemic public goods with the data: For example, locational data could be mined for insight for epidemiological purposes against contagious diseases. Or it could be used to improve access for those with disabilities. Finally, public use of data subject to the trust would be constrained by limits designed to maintain individual privacy and public trust—much as data collected by social security and tax authorities is constrained.

A public trust in data (of a sort) has already been implemented by cities around the world:

- The Spanish city Barcelona uses a platform called "Decidem" as a vehicle for the governance of personal data.[5] For example, a company wishing to operate a service that creates and uses personal locational data—say, a bike sharing firm—must agree to give their data to Decidem, where its uses will be subject to public debate and decision.[6] Decidem is being adopted by other European cities such as Amsterdam. It aims to create "new types of local data commons where people are empowered to collect and share data in response to local challenges."[7]

- Since January 2019, New York City has mandated that ride-sharing companies such as Uber and Lyft disclose operational data on "the date, time, and location of pickups and drop-offs (at least down to the intersection), the vehicle's license number, the trip mileage, itemized trip fare, route (including whether the vehicle entered traffic-choked Midtown), and how much the driver was paid" as a condition of operating.[8] By bringing this data into public hands, New York City takes a crucial step toward making locational data a matter of public trust.

- On the other side of the country, the Silicon Valley Data Trust brings together streams of information from benefits agencies, child protection bureaus, schools, and education technology companies to create a "well-managed regional data trust [and] provide a

[5] Amy Lewin, *Barcelona's Robin Hood of Data*, SIFTED (Nov. 16, 2018), https://sifted.eu/articles/barcelonas-robin-hood-of-data-francesca-bria/
[6] *Id.*
[7] Theo Bass and Rosalyn Old, *Common Knowledge: Citizen-led Governance for Better Cities* 24 (Jan. 2020), https://media.nesta.org.uk/documents/DECODE_Common_Knowledge_Citizen_led_data_governance_for_better_cities_Jan_2020.pdf/.
[8] Aarian Marshall, *NYC Now Knows More Than Ever About Your Uber and Lyft Trips*, WIRED (Jan. 31, 2019), https://www.wired.com/story/nyc-uber-lyft-ride-hail-data/ [hereinafter "Marshall, *NYC*"].

Electronic copy available at: https://ssrn.com/abstract=3794780

comprehensive understanding of factors contributing to student failure and success."[9] Hosted by Santa Clara County, the Data Trust's primary purpose is to allow research to "improve service and educational outcomes, especially for children of poverty."[10] It thus allows for the production of a public good that would otherwise be untapped.

All these initiatives blend together private and public data, impose democratically determined use rules, and vindicate policy goals that would otherwise go unrealized. They are all ways to ensure, as Rana Foroohar of the *Financial Times* has put it, that firms are not "mining our biggest natural resource for free."[11]

This Article extends these emergent models of personal-data governance by a public body. It offers a 'proof of concept' for a new legal strategy, albeit one with deep common-law roots, tailored to a novel resource. The asset I'm concerned with here is "personal data." (By this, I mean information that "singles out a specific individual from others," and also "when specific identification is " possible, [if] not a significantly probable event."[12]) The paper's core intuition is to view this data not as an aspect of individual action or a particular firm's ingenuity, but as a shared asset—one realized through the entangled social interactions of the many, and one vindicated by recognizing the many rather than the one as pivotal. In this regard, its motivating impulse is Karl Polanyi's injunction to "transcend the self-regulating market by consciously subordinating it to a democratic society."[13]

Why look to the ancient common law for a solution to a distinctively modern problem of personal data?[14] Familiarity no doubt eases the transitional costs of adoption. But more importantly, the public trust is already well suited to address the harms flowing from personal data economies. The doctrine was first developed in Roman and English common law as a governance tool for common resources such as fisheries and shared navigable waters. In the late nineteenth century, the U.S. Supreme Court embraced it as an instrument for managing common assets at risk of abuse by powerful interest groups. The Court's analysis suggested that such an asset should benefit a broad public, not just powerful firms. It also identified a risk that state bodies such as legislatures might be captured, and thus dispose of the asset in ways that contravened the public interest.[15] To prevent this, a Progressive Era public trust came with a thicket of substantive and procedural safeguards. In this way, the public trust took flight in America's first gilded age of inequality and corporate dominance as a doctrinal shield for the public against the abuse of

[9] *The Silicon Valley Data Trust* (Jan. 12, 2021), https://www.svrdt.org/.
[10] *Id.*
[11] RANA FOROOHAR, DON'T BE EVIL: THE CASE AGAINST BIG TECH 275 (2019).
[12] Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and A New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814, 1877-78 (2011); *cf.* NEIL RICHARDS, INTELLECTUAL PRIVACY: RETHINKING CIVIL LIBERTIES IN THE DIGITAL AGE 8 (2015) (describing privacy as "the ability to control information about yourself, which may captured many, but not most, use of that term").
[13] KARL POLANYI, THE GREAT TRANSFORMATION: THE POLITICAL AND ECONOMIC ORIGINS OF OUR TIME 242 (2001).
[14] Even more familiar applications of public trust principles are subject to criticism. For a cogent argument on democracy-related grounds, William D. Araiza, *The Public Trust Doctrine As an Interpretive Canon*, 45 U.C. DAVIS L. REV. 693, 696 (2012) (noting the criticism that "courts have neither the legal authority nor the expertise:" to implement public trusts).
[15] *Cf.* Sax, *supra* note 3, at 521 ("The 'public trust' doctrine has no life of its own and no intrinsic content. It is no more—and no less—than a name given by courts to their concerns about the democratic process.").

Electronic copy available at: https://ssrn.com/abstract=3794780

concentrated private power. It is an animating logic that can be transposed seamlessly to the present day.

The public trust is also a malleable doctrinal tool. State courts have adapted and refined the public trust doctrine to fit widely divergent asset types.[16] It has been used to cover oyster beds, navigation rights in lakes and rivers, parklands, groundwater, and the littoral beaches of the Atlantic seaboard.[17] The environmental resources to which the doctrine is canonically applied are natural rather than man-made. But still, they have strikingly similarities to personal data. Some were created through the contributions of many; most are valuable when aggregated rather than when divided; they have a borderless, open-ended quality; and they present the need to balance both public and private uses. Personal data, in short, is well-suited to the public trust doctrine in terms of its form. For each kind of asset falling within the doctrine, judges have fashioned a distinct set of governance rules. These can include easements for the general public, limitations on alienation and uses, even something akin to administrative law's 'hard look' doctrine.[18] This sheer range of doctrinal ingenuity reflects common-law judges' ingenious efforts to balance diverse exploitation and spoilage risks. Today, it means that the public trust doctrine offers a rich repertoire of doctrinal tools for managing personal data as a common asset.

The most promising venue for the creation of a public trust is municipal government. Indeed, this is where it's already happening. Locational data of the sort collected by Decidem provides the most ready target for a public trust. Cities could also extend pubic trusts in data to protect the personal data their residents generate on platform economies such as Facebook, Twitter, and Google, and on the internet more generally. Regulation by even a small number of large American cities has the potential to force significant, public-regarding changes to personal data economies.[19] Once established by a legislature or state constitution-making body, a public trust also has democratic credentials: It enables ongoing public deliberation and determination of how a common asset is employed, with the courts working as a back-stop against interest-group capture. It could be used to stymie the harmful uses of personal, while promoting public-regarding ones, and preventing the concentration of profits in a small number of firms to the exclusion of those who create data in the first instance. The result would be a more democratic and less regressive data-based economy, one that was less inimical to interests in privacy and economic desert.

The argument has four elements. Part I summarizes the basic economic logic of commodification, circulation, and use animating personal-data economies. To establish the need for a new regulatory intervention, Part II documents, and then supplements, a catalog of normative objections to personal data economies. The leading proposed responses—including the creation of a private property interest in data and the aggressive application of structural antitrust remedies—are considered in Part III. While valuable, none of these interventions covers the waterfront of harms documented in Part II. The fourth, most important, part of the Article proposes and defends the possibility of a public trust in data as a generally applicable vehicle for addressing harms of new personal data economies.

[16] *Id*. at 509 (describing the doctrine as a "technique").
[17] *See infra* Part IV.A.
[18] *See infra* Part IV.B.
[19] *See infra* Part IV.C.

Electronic copy available at: https://ssrn.com/abstract=3794780

I. Our Data and the Economies it Makes

The case for a public trust in personal data stands upon first a factual and then a normative foundation. This Part takes up the factual part of that case by setting out how personal data becomes an asset. It historicizes the emergence of data economies, and outlines in more detail three circuits through which personal data passes to accrue commercial value.

A. The Pre-History of our Data Economies

Economies of personal data are more than a century old.[20] In 1903, the New York Life insurance company adopted the nation's first insurance rating system drawing on demographic and health data.[21] Two years later, Pennsylvania enacted a law requiring the collection of vital statistics, greatly expanding the empirical scope for actuarial calculation.[22] Midcentury advances in computing stimulated new data uses. In 1953, the AT&T Company purchased IBM's Univac system to manage its 100,000-person Bell System Employee Attitude Survey,[23] while the Nielsen Company started gathering data on what TV American families watched.[24] In 1959, the Simulmatics corporation broke new ground. It sieved polling data through a scrim of midcentury behavioral science to create predictions for sale to advertising agencies and political campaigns.[25] By 1972, the National Academy of Sciences would document fifty-five large "computing organizations" collecting personal data. These ranged from the Bank of America and the Mutual of Omaha to the Massachusetts Institute of Technology and the Church of the Latter-Day Saints.[26] Yet the Academy found "no radical departures" from pre-computer information practices.[27] Data may have been everywhere, but it was also almost always economically sterile.

Had technological and organizational constraints held fast, the best governance regime for personal data would be a question of only passing interest. But the 1980s saw great advances in machine-learning algorithms for discerning relationships and making predictions using large datasets.[28] Then in the 1990s, large corporations found themselves accumulating "tremendous"

---

[20] The creation of "data doubles" by "states, corporations, and voluntary organizations" to further a range of "governing ambitions" dates back to the post-Civil War era. Dan Bouk, *The History and Personal Economy of Personal Data over the Last Two Centuries in Three Acts*, 32 OSIRIS 85, 89 (2017); *see also* JAMES C. SCOTT, SEEING LIKE A STATE: HOW CERTAIN SCHEMES TO IMPROVE THE HUMAN CONDITION HAVE FAILED 2 (1992) (describing aspiration to render governed populations "legible").

[21] JAMES R. BENIGER, THE CONTROL REVOLUTION: TECHNOLOGICAL AND ECONOMIC ORIGINS OF THE INFORMATION SOCIETY 422 (1986); *see also* Bouk, *supra* note 20, at 97 ("By making people into 'statistical individuals' it became possible to sort them according to the futures the statistics predicted for them.").

[22] By 1929, all but two states, South Dakota and Texas, had followed suit. COLIN KOOPMAN, HOW WE BECAME OUR DATA 47-48 (2019).

[23] Frederick F. Stephan et al., *The machine revolution in the processing of data*, 21 PUB. OP. Q. 410, 411 (1957).

[24] Bouk, *supra* note 20, at 102.

[25] JILL LEPORE, IF THEN: HOW THE SIMULMATICS CORPORATION INVENTED THE FUTURE (2020).

[26] SARAH E. IGO, THE KNOWN CITIZEN: A HISTORY OF PRIVACY IN MODERN AMERICA 246 (2018).

[27] *Id. But see* Rob Lucas, *The Surveillance Machine*, 121 NEW LEFT REV. 132, 141 (2020) (describing collection activities of TRW corporation during the 1970s).

[28] ETHEM ALPAYIN, MACHINE LEARNING 28 (2016) (describing the emergence of machine learning alongside "the capacity to build parallel hardware containing thousands of processors"); MELANIE MITCHELL, ARTIFICIAL INTELLIGENCE: A GUIDE FOR THINKING HUMANS 39-40 (2019) (identifying 1980s as the first majority advance of artificial intelligence. A seminal paper describing the backpropagation method used in neural networks was published in 1986. *See* David E. Rumelhart, Geoffrey E. Hinton & Ronald J. Williams, *Learning representations by*

Electronic copy available at: https://ssrn.com/abstract=3794780

amounts of data, often across disparate and irreconcilable databases.[29] In response, techniques of "data warehousing" in ever "more comprehensive database[s]" emerged.[30] Improvements in computing power captured in Moore's famous law, the commodification of cheap data storage, and the creation of new data-flows cracked commercial frontiers.[31] And then, in the 2010s, improvements in the processing of natural-language corpuses opened a new, yet larger realm of speech and internet interaction for analysis and enclosure.[32]

There are data aplenty to analyze now. A 2003 study reported that whereas humanity had accumulated about 12 exabytes of data before the commodification of computers, in 2003 alone five exobytes were created.[33] Some of the first commercial surveillance tools, such as IBM's EasiOrder and Firefly Network's "intelligent agent" software, emerged just as firms were grappling with the problem of how to exploit this new resource.[34] The upward arc of data production continues. Data generated by the Internet is predicted to reach some 3.3 zetabytes in 2021.[35] The Internet facilitates, while being enabled by, personal data economies. The more it is used, the more data is produced about its users' habits, preferences, and behaviors—and the more apps and services can be built.[36] There is also an ongoing dispersion of sensors into cellphones, vehicles, appliances, and physical infrastructure—the so-called internet of things—that will make the familiar internet but one tributary of an increasingly engorged "exaflood" of data.[37]

The interaction between new machine-learning tools and this exaflood is reworking the economy—from manufacturing to logistics to retail to human resources to marketing. Many of the ensuing circuits of acquisition, processing, and use concern personal data. And the line between personal data economies and other data economies can be fuzzy. Financial institutions, for example, use machine-learning tools to conduct algorithmic trading instruments and manage compliance with capitalization regulations.[38] At the same time, many of the same institutions also use the same tools to assign credit risk and to identify fraudulent activity by employees, perhaps using the same data.[39]

Upon this fertile ground, data is reimagined as something that is not barren or static. Instead, it is "a raw material of business" and "a new form of value."[40] It is, in short, an asset—

---

back-propagating errors, 323 NATURE 533 (1986); *see also* James Somers, *Is AI Riding a One –Trick Pony?*, 120 MIT TECH. REV. 29, 31 (2017) (explaining the historical emergence of contemporary forms of machine learning).

[29] JOHN D. KELLEHER AND BRENDAN TIERNEY, DATA SCIENCE 8 (2018).

[30] *Id.* at 8-9.

[31] *Id.* at 30-31.

[32] NICK POLSON AND JAMES SCOTT, AIQ: HOW ARTIFICIAL INTELLIGENCE WORKS AND HOW WE CAN HARNESS ITS POWER FOR A BETTER WORLD 130-32 (2018).

[33] LUCIANO FLORIDI, INFORMATION: A VERY SHORT INTRODUCTION 6 (2010).

[34] Sarah Myers West, *Data capitalism: Redefining the logics of surveillance and privacy*, 58 BUS. & SOC. 20, 26 (2019).

[35] CARL BENEDIKT FREY, THE TECHNOLOGY TRAP: CAPITAL, LABOR, AND POWER IN THE AGE OF AUTOMATION 303-04 (2019). A zettabyte is 2 to the 70th power bytes.

[36] *Id.* at 304.

[37] FLORIDI, *supra* note 33, at 6.

[38] Larry D. Wall, *Some financial regulatory implications of artificial intelligence*, 100 J. ECON. & BUS. 55, 56-61 (2018).

[39] *Id.*

[40] VIKTOR MAYER-SCHÖNBERGER & KENNETH CUKIER, BIG DATA: A REVOLUTION THAT WILL TRANSFORM HOW WE LIVE, WORK, AND THINK 5 (2013).

Electronic copy available at: https://ssrn.com/abstract=3794780

even if it directly touches on, or indirectly could be used to reveal, information most people would consider properly subject their own exclusive control.

B. Three Contemporary Economies of Personal Data

Personal data can create economic rent in many ways. This section identifies three of the most important means by which such data circulates and accrues value. All three domains share the same three-step sequential logic of "dragnets, scores, and interventions."[41] First comes sweeping collection of data regardless of a firm's "imaginative reach or analytic grasp," in the hope it will "eventually be useful."[42] The second step involves nesting data within a classification system. Through the extraction of standardized "features,"[43] classification makes available "various scoring, grading, and ranking methods."[44] In the final step, data is treated as an asset to be refined to fit the needs of advertisers and others.[45] This is the point at which "data rents," or "revenues that can be derive from ownership and control rights over personal data (as an asset)" are derived.[46]

The combination of dragnets, scoring, and interventions characterizes various data economies. For present purposes, the most important of these are *platforms economies* such as Facebook and Google; the business-facing clearing houses called *data brokerages*; and an emerging economy of *sensing nets*—spatially distributed devices gathering data on individuals' behavior. These three nodes do not exhaust contemporary personal data economies. But they are the most important. As such, their workings bear directly on the question of appropriate governance regime for personal data.

1. Platforms Economies

A platform is a model for organizing transactions in the data economy. Platforms use "technical protocols and centralized control to define networked spaces in which users can conduct a heterogeneous array of activities."[47] A platform can perform several different functions. It can

---

[41] Marion Fourcade and Kieran Healy, *Seeing Like a Market*, 15 SOCIO-ECONOMIC REV. 9, 13 (2017).
[42] *Id.*
[43] David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653, 700–01 (2017) (describing feature selection).
[44] Fourcade and Healy, *supra* note 41, at 14. Features and classificatory systems can be opaque and, for all practical purposes, impenetrable to human understanding. Machine learning in particular can entail a transformation of data through analytic tools that cannot be reproduced in a form comprehensible to a human. Jenna Burrell, *How the machine 'thinks': Understanding opacity in machine learning algorithms*, 3 BIG DATA & SOC. 1, 10 (2016). Nevertheless, the classification system's results can have commercial value.
[45] Thomas Beauvisage and Kevin Meller, *Datasets: Assetizing and Marketing Personal Data*, *in* TURNING THINGS INTO ASSETS (2020) (draft at 9).
[46] Kean Birch, Margaret Chiappetta, and Anna Artyushina, *The problem of innovation in technoscientific capitalism: data rentiership and the policy implications of turning personal digital data into a private asset*, 41 POL. STUD. 468, 475 (2020); *see also* Paul Langley and Andrew Leyshon, *Platform capitalism: the intermediation and capitalisation of digital economic circulation*, 3 FIN. & SOC. 11, 13 (2017) ("[T]he revenues prescribed by the platform business model amount to the extraction of 'rents' from circulations and associated data trails.").
[47] JULIE E. COHEN, BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTION OF INFORMATIONAL CAPITALISM 41 (2019). Another definition of a "platform" is "a discrete and dynamic arrangement defined by a particular combination of socio-technical and capitalist business practices." Langley & Leyshon, *supra* note 46, at 14.

Electronic copy available at: https://ssrn.com/abstract=3794780

"replace" and "rematerialize" a market—think Amazon, Uber, or AirBnB.[48] Often, a platform creates two connected markets: one facing consumers from whom data is extracted, and the other facing other businesses that purchase either the data or a good into which the value of acquired data is impounded (e.g., advertising slots). On the consumer-facing side, a platform can be a substitute for social and cultural connectivity. Or it can become the democratic public sphere. On the business-to-business side, a platform's data can help to target messaging and products; to manage and control behavior (e.g., worker productivity); to model probabilities; or to grow the value of other physical assets (e.g., by tailoring their deployment to the emergence of new consumer demands).[49]

Leading platforms interact with staggering numbers of individuals. Because it is the largest such platform, Facebook serves as a useful example. Starting in 2004 at Harvard, Facebook "leapfrogged from campus to campus," engorging itself along the way.[50] As it grew, its allure remained the sentimental leveraging of "the need to see what old friends or family were up to without the burden of talking to them."[51] In return for this connectivity, users "hand[ed] over a treasure trove of detailed demographic data."[52] This data could be used to target advertisements, which generate the overwhelming share of platform revenues. In its 2019 regulatory filings, Facebook reported having some 2.6 billion users.[53] Two-third of Americans use Facebook every day.[54] This creates a surge of detailed data on items, including "the amount of time you hover your mouse over a particular button and the number of days an item sits in your shopping basket, to every location you've visited with your phone and how you psychologically react to different posts and words."[55] Yet in Facebook's "data-for-payment" model, consumers may or may not understand the extent to which they are paying for a service through their disclosures of their own (and perhaps also others') personal data.[56]

In addition to the data acquired through user interactions on the social network, Facebook also derives user and nonuser information from the millions of independent websites and apps integrating Facebook's Like button.[57] "Facebook uses plug-ins to track users' browsing histories when they visit third-party websites, and then compiles these browsing histories into personal

---

[48] COHEN, *supra* note 47, at 42; *see also* Orly Lobel, *The Law of the Platform*, 101 MINN. L. REV. 87, 94 (2016) (focusing on the role of a "platform company … as an online intermediary between buyers and sellers of goods and services" aided by new digital technologies).
[49] Jathan Sadowski, *When data is capital: Datafication, accumulation, and extraction*, 6 BIG DATA & SOCIETY 1, 5-6 (2019).
[50] TIM WU, THE ATTENTION MERCHANTS: THE EPIC SCRAMBLE TO GET INSIDE OUR HEADS 296-97 (2016).
[51] *Id.* at 296.
[52] *Id.* at 301.
[53] Form 10-Q, United States Sec. & Exch. Comm'n (Mar. 31, 2020), http://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/bfe31518-2e18-48fb-8d98-5e8b07d94b2a.pdf.
[54] Aaron Smith & Monica Anderson, Pew Research Ctr., *Social Media Use in 2018*, at 2 (2018), https://www.pewinternet.org/wp-content/uploads/sites/9/2018/02/PI_2018.03.01_Social-Media_FINAL.pdf.
[55] Lina M. Khan, *Sources of Tech Platform Power*, 2 GEO. L. TECH. REV. 325, 329 (2018). When this facility was rolled out, Facebook "Facebook induced websites to install Facebook plug-ins by representing that the company would not use this installed code to channel user data to its advertising business. Lina M. Khan, *The Separation of Platforms and Commerce*, 119 COLUM. L. REV. 973, 1005 (2019) [hereinafter "Khan, *Separation*"].
[56] Stacy-Ann Elvy, *Paying for Privacy and the Personal Data Economy*, 117 COLUM. L. REV. 1369, 1384 (2017).
[57] *Websites using Facebook Like Button*, Builtwith.Com (2018), https://trends.builtwith.com/websitelist/Facebook-Like-Button.

Electronic copy available at: https://ssrn.com/abstract=3794780

profiles which are sold to advertisers to generate revenue."[58] When installed on an Android mobile phone, Facebook's app also captures call history and messaging activity.[59] All this data helps Facebook sell "impression-targeted ads" and "action-based ads."[60] And there is no opt-out. For its first six years, user surveillance was not among Facebook's mandatory terms. But since an initial public offering, user surveillance has been mandatory[61] and legally largely unconstrainted.[62]

The collection of personal data as a collateral, often unwitting, side-effect of platform use is central to the business model of other platforms. Take Google. Every use of Google Maps or Search "saves certain information about a user's activity" for the company.[63] As a result, Google's AdWords, which appear as text alongside search results, are "wildly successful as a means for monetizing the company's search business."[64] More than four-fifths of Alphabet's revenue derives from the sale of advertisements targeted using this data.[65]

A novel platform economy with profound and unexplored ramifications is the consumer genomics market. This market did not exist before the early 2000s.[66] Once the human genome had been sequenced (in 2003), advances in sequencing tools reduced costs to the point where it had become a multibillion-dollar *consumer* industry by the mid-2010s.[67] Like social media networks and search, consumer genomics is a two-sided market. It serves both individuals seeking genetic information and also pharmaceutical firms and research laboratories seeking data stocks.[68] With rare exceptions, genomic data from a single individual is not particularly valuable. Hundreds or thousands of samples are required for effective medical exploitation.[69] Genetic datasets, however,

---

[58] *In re Facebook, Inc. Internet Tracking Litig.*, 956 F.3d 589, 596 (9th Cir. 2020).
[59] Alex Hern, *Facebook Logs SMS and Calls, Users Find as They Delete Accounts*, THE GUARDIAN (Mar. 26, 2018), https://www.theguardian.com/technology/2018/mar/25/facebook-logs-texts-and-calls-users-find-as-they-delete-accounts-cambridge-analytica.
[60] Dina Srinivasan, *The Antitrust Case Against Facebook: A Monopolist's Journey Towards Pervasive Surveillance in Spite of Consumers' Preference for Privacy*, 16 BERKELEY BUS. L .J. 39, 42–43 (2019).
[61] *Id.* at 44-45; Jennifer Shore & Jill Steinman, *Did You Really Agree to That? The Evolution of Facebook's Privacy Policy*, TECH. SCI. (2015), https://techscience.org/a/2015081102 (documenting decline in the strength of Facebook's privacy promises).
[62] But in April 2020, the Court of Appeals for the Ninth Circuit held that plaintiffs had standing and a right of action under California privacy law and federal wiretap statute to challenge Facebook's use of plug-ins to track logged-out users' browsing histories when they visited third-party websites. *In re Facebook, Inc. Internet Tracking Litig.*, 956 F.3d 589, 596 (9th Cir. 2020).
[63] *In re Google Location History Litig.*, 428 F. Supp. 3d 185, 188 (N.D. Cal. 2019) (citation and quotation marked omitted).
[64] West, *supra* note 34, at 32.
[65] Alphabet Inc., Annual Report (Form 10-K) 7 (Feb. 4, 2019), https://www.sec.gov/Archives/edgar/data/1652044/000165204419000004/goog10-kq42018.htm
[66] JOHN ARCHIBALD, GENOMICS: A VERY SHORT INTRODUCTION 14-26 (2018).
[67] Susi Geiger and Nicole Gross, *A tidal wave of inevitable data? Assetization in the consumer genomics testing industry*, BUS. & SOC. 1, 10 (2019).
[68] *Id.* at 11.
[69] Laura M. Beskow, *Lessons from HeLa cells: the ethics and policy of biospecimens*, 17 ANN. REV. GENOMICS & HUMAN GENETICS 395, 397 (2016).

Electronic copy available at: https://ssrn.com/abstract=3794780

fall outside the 1996 Health Portability Accountability Act, and other regulatory regimes.[70] Its commercial sale, purchase, and use is therefore largely unregulated.[71]

2. Data Brokers

If social media and search engines are the store front of the personal data economy, data-brokers are its back office. Data brokers engage in "information arbitrage" by "buying, repackaging, and selling consumer data across various markets."[72] They help create telemarketing lists, aid debt collectors, screen employees, and select targets for credit offers.[73] In 2017, the data-broker Axiom offered to sell up to three thousand attributes on *each* of the seven hundred million individuals in their records; in 2018, it could offer up to ten thousand items on some 2.5 billion people.[74] These numbers are likely bigger now.

The data brokerage industry is opaque.[75] In 2014, the Federal Trade Commission promulgated a much-remarked report listing nine firms.[76] In May 2018, Vermont enacted legislation requiring entities collecting third-party data for commercial purposes to register.[77] More than 120 did. These ranged from long-established credit reporting agencies such as Experian and Axcion to novel online search engines such as Spokeo, and smaller niche actors catering to landlords and insurance companies.[78] Some are behemoths. In 2018, Axcion reported operating revenues of $917 million.[79] And the Vermont registry is incomplete. It excluded companies that exploit or trade data related to their own customers, including platforms such as Facebook and

---

[70] Kelsey Russo, *The Digital Life of Henrietta Lacks: Reforming the Regulation of Genetic Material*, 38 J. LEGAL MED. 449, 460–61 (2018).

[71] Academic research using genomic data is, however, governed by the Common Rule. Beskow, *supra* note 69, at 399-400.

[72] Matthew Crain, *The Limits of Transparency: Data Brokers and Commodification*, NEW MEDIA & SOC. 1, 11 (2016); *see also* West, *supra* note 34, at 29-30 (characterizing data brokers as "an industry that quickly grew around the collection of on-line data, forming a market ecosystem that treated data as a commodity to be sold and circulated").

[73] Leanne Roderick, *Discipline and Power in the Digital Age: The Case of the U.S. Consumer Data Broker Industry*, 40 CRITICAL SOC. 729, 732 (2014).

[74] Steven Melendez, *Here are the Data Brokers Quietly Buying and Selling Your Data*, FAST COMPANY (Mar. 8, 2019), https://www.fastcompany.com/90310803/here-are-the-data-brokers-quietly-buying-and-selling-your-personal-information [hereinafter "Melendez, *Here are the Data Brokers*"]

[75] *See* COHEN, *supra* note 47, at 62 (noting refusal of data brokers to testify before Congress); West, supra note --, at 30 (noting how data brokers "remain under the radar").

[76] Fed. Trade Comm'n, *Data Brokers: A Call for Transparency and Accountability* ii (2014), , http://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf (listing Acxiom, Corelogic, Datalogix, eBureau, ID Analytics, Intelius, PeekYou, Rapleaf, and Recorded Future).

[77] *See* H. 764, 2017-2018 Gen. Assemb. (Vt. 2018), https://legislature.vermont.gov/bill/status/2018/H.764 [https://perma.cc/QYX3-CEWX] (reporting that the bill was enacted without the governor's signature on May 22, 2018). The act defines "data broker" as "a business ... that knowingly collects and sells or licenses to third parties the brokered personal information of a consumer with whom the business does not have a direct relationship." Vt. Stat. Ann. tit. 9, § 2430 (4)(A).

[78] Steven Melendez, *A landmark Vermont law nudges over 120 data brokers out of the shadows*, FAST COMPANY (Mar. 2, 2019), https://www.fastcompany.com/90302036/over-120-data-brokers-inch-out-of-the-shadows-under-landmark-vermont-law.

[79] Axcion Annual Report at F2 (2009), at https://www.annualreports.com/HostedData/AnnualReports/PDF/NASDAQ_ACXM_2018.pdf

Electronic copy available at: https://ssrn.com/abstract=3794780

search engines such as Google.[80] Its 120 registered firms are only a "fraction" of the broader set of back-end processors and vendors of personal data.[81] Over time, the data brokerage industry has had "movements of consolidation" resulting in "large multi-purpose data brokers."[82] But a recent estimate still counts between 2,500 and 4,000 data-brokering firms in the United States.[83]

3. Sensing Nets

A third personal data economy involves the collection, classification, and application of digital traces produced by physical devices. Small, low-cost, wireless, and energy-efficient sensors can be installed in a range of objects to generate geolocated digital signatures.[84] Vehicles, home appliances, Fitbits and other portable devices, home and office security systems, and even medical devices produce a constant stream of physical, behavioral, locational, and biometric information.[85] Almost all American vehicles, for example, contain an event data recorder that "continuously measure[s] information on a car's speed, braking, acceleration, angular momentum, and other similar data.[86] In the home, "Siri" is now in active use on more than a half billion devices globally.[87] In the United States, as of 2019 some 69 percent of U.S. homes were using "smart" devices, including home networking, home security, smart thermostats, smart lighting, or video doorbells.[88] Closer to the bone, medical devices such as the artificial pancreas, used by diabetics to substitute for their inadequate insulin supply, pipe out a stream of information about somatic operations that are unavailable even to the person in question.[89]

All these digital traces enable a wide range of inferences about behavior, habits, and bodies. Together, they operate as a "sensing net" covering a large and varied slice of human behavior.[90]

---

[80] Issie Lapowsky, *How Tim Cook's Data Registry Might Actually Work*, WIRED (Jan. 23, 2019), https://www.wired.com/story/tim-cook-data-broker-registry/.
[81] Melendez, *Here are the Data Brokers*, *supra* note 74. In 2014, journalist Julia Angwin reported that she had identified 212 data brokers by searching for entities that held her own information; her efforts to scrub personal data from them were only partially successful. JULIA ANGWIN, DRAGNET NATION: A QUEST FOR PRIVACY, SECURITY, AND FREEDOM IN A WORLD OF RELENTLESS SURVEILLANCE 161-63 (2014).
[82] Beauvisage and Meller, *supra* note 45, at 11.
[83] Paul Boutin, *The Secretive World of Selling Data About You*, NEWSWEEK (May 30, 2016), http://www.newsweek.com/secretive-world-selling-data-about-you-464789.
[84] Fed. Trade Comm'n, *Internet of Things: Privacy & Security in a Connected World* 5-6 (2015), https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-staff-report-november-2013-workshop-entitled-internet-things-privacy/150127iotrpt.pdf [https://perma.cc/W8YD-SGA9].
[85] An early and prescient taxonomy is Scott R. Peppet, *Regulating the Internet of Things: First Steps Toward Managing Discrimination, Privacy, Security, and Consent*, 93 TEX. L. REV. 85, 98-117 (2014). For examples, see Maggie Astor, *Your Roomba May Be Mapping Your Home, Collecting Data That Could Be Shared*, N.Y. TIMES (July 25, 2017) (vacuum cleaner); Ry Crist, *Here's What's Next for Samsung's Family Hub Smart Fridge*, CNET (Jan. 7, 2018, 2:00 PM PST) (fridge); Chris Matyszczyk, *Samsung's Warning: Our Smart TVs Record Your Living Room Chatter*, CNET (Feb. 8, 2015, 2:10 PM PST) (television).
[86] Daniel Harper, *Automobile Event Data Recorders, and the Future of the Fourth Amendment*, 120 COLUM. L. REV. 1255, 1255–56 (2020).
[87] *HomePod arrives February 9, available to Order this Friday*, APPLE (Jan. 23, 2018), https://www.apple.com/newsroom/2018/01/homepod-arrives-february-9-available-to-order-this-friday/
[88] Chuck Martin, *Smart Home Technology Hits 69% Penetration in U.S.,* MEDIAPOST (Sept. 30, 2019), https://www.mediapost.com/publications/article/341320/smart-home-technology-hits-69-penetration-in-us.html/
[89] For a thoughtful mediation on this technology, see Mark C. Taylor, *A.I. and I*, N.Y. TIMES (Dec. 14, 2020), https://www.nytimes.com/2020/12/14/opinion/AI-human-body.html.
[90] COHEN, *supra* note 47, at 57-58.

Electronic copy available at: https://ssrn.com/abstract=3794780

The end result will be one "in which all software needs to be spatially aware."[91] One especially ambitious proposal involves the dispersion "smart dust," or "nanosensors—scattered micro devices that are smaller than grains of rice" that would be "laced ubiquitously" through urban physical spaces.[92]

The sensing net torques economic logic of familiar goods. In 2017, the American appliance manufacturer Whirlpool sought tariffs on Korean competitors LG and Samsung for flooding the U.S. market with cheap smart devices. The Korean companies' strategies reflected its belief that "in a data-driven business," the best way to expand market share is "to push prices as low as possible in order to build your customer base, enhance data flow, and cash in" later.[93] Companies can use the data produced by, say, a fridge by "selling you recipe subscriptions, maybe, or getting a cut of your food order," which it places automatically when you run low.[94] The physical appliance at least "becomes (perhaps primarily) a means of producing data."[95] At some point, personal data created by appliances will be more valuable than the appliance itself.[96] Like consumer genetic data, these flows are now largely unregulated.[97] Indeed, roughly two-fifth of firms offering paid mobile applications or devices that monitor consumer health lack even a privacy policy.[98] As a result, there are few legal constraints on the manner in which the resulting data personal can be transformed into assets.

* * *

The technological and material foundations of personal data economies are barely twenty years old. Yet these two decades have witnessed an explosion of new tools for social connectivity, information acquisition, and more. These generate, and sometime rely on, new flows of personal data. This is treated as an asset, either for exploitation within a firm or for resale. The next result is a robust and interlinked series of personal data economies—the most important of which have been detailed here.

II. The Discontents of Personal Data Economies

Personal data economies enhance individual welfare through connectivity, search, and the personalized tailoring of goods and services. Yet there is also widespread discontent about their individual and social effects. The ambition to constrain some or all of these ills motivates many proposals for new governance regimes. To evaluate governance regimes for the data economy, it is useful to begin with its critiques.

---

[91] SHASHI SHEKHAR AND PAMELA VOLD, SPATIAL COMPUTING 23 (2019).
[92] CARLO RATTI AND MATTHEW CLAUDEL, THE CITY OF TOMORROW: SENSORS, NETWORKS, AND THE FUTURE OF URBAN LIFE 48 (2016).
[93] Adam Davidson, *A Washing Machine that tells the Future*, NEW YORKER (Oct. 23, 2017), https://www.newyorker.com/magazine/2017/10/23/a-washing-machine-that-tells-the-future.
[94] *Id.*
[95] Sadowski, *supra* note 49, at 7.
[96] *See, e.g.,* Matt McFarland, *Your Car's Data May Soon Be More Valuable Than the Car Itself*, CNN: TECH (Feb. 7, 2017 9:05 AM), http://money.cnn.com/2017/02/07/technology/car-data-value/index.html
[97] Three states have enacted legislation respecting biometric data. *See* 740 Ill. Comp. Stat. Ann. 14/15 (2018); Tex. Bus. & Com. Code Ann. § 503.001 (2017); Wash. Rev. Code Ann. § 19.375.010 (2017).
[98] Stacy-Ann Elvy, *Commodifying Consumer Data in the Era of the Internet of Things*, 59 B.C. L. REV. 423, 439 (2018).

Electronic copy available at: https://ssrn.com/abstract=3794780

This Part offers a taxonomy of normative critiques leveled against personal data economies. They fall into six broad categories: privacy, autonomy, exploitation, economic inequality, democratic backsliding, and state domination. To this half-dozen found in the literature, I add a seventh: The failure to generate beneficial public goods. Boundaries between these categories are sometimes porous. Problems of privacy, for example, can sometimes be rephrased as matters of autonomy, as can critiques of exploitation. Economic inequality can also be understood as a mere aggregation of worries about individuals' exploitation. Despite these blurred lines, the taxonomy is helpful because it clarifies the stakes, and provides a way to sort among different normative priorities. From the taxonomy, a general theme also bubbles up: the most plausible and forceful lines of critique largely (although not perfectly) converge in a consequentialist concern about the way in which platform economies, data brokers, and sensing nets exacerbate structural inequality in society at large.

A. Privacy

Economies of personal data raised privacy concerns from their conception. Even early information-economy boosters recognized that a firm alchemizing personal data into economic rents could impose privacy spillovers.[99] Privacy concerns are either retail—i.e., specific to certain affordances, firms, or sectors—or wholesale—i.e., applicable to the whole data economy. I sketch three retail lines of criticism, and then summarize the leading global critique.

First, a firm might be criticized for failing to abide by its own privacy regulations. Data sharing among platform economies is an example. In December 2018, a set of special arrangements between Facebook on the one hand, and Amazon, Bing, Spotify, and Yahoo (among others) allowed those counterparties access to users' personal data through undisclosed exemptions to privacy policies.[100] Harms arise because users did not agree to, and so cannot stop, these transfers.

Second, a firm might fail to secure personal data from external misappropriation. Again, harm arises from disclosures without consent and beyond user control. Data breaches are said to cause "an increased risk of identity theft, fraud, and reputational damage," as well as immediate "[e]motional distress."[101] The precise rate of personal data breaches is unknown because not all are reported.[102] Between 2000 and 2010, however, one study found more than 230 data-breach suits in federal court.[103] Risk of a breach unsurprisingly rises as the volume of information being

[99] Viktor Mayer-Schoenberger and Kenneth Cukier, for example, called for a new "caste of big-data auditors" to deal with privacy concerns. MAYER-SCHÖNBERGER & CUKIER, *supra* note 40, at 184. Informational privacy is "concerned with the use, transfer, and processing of the personal data generated in daily life." Paul M. Schwartz, *Property, privacy, and personal data*, 117 HARV. L. REV. 2056, 2058 (2003) [hereinafter "Schwartz, *Property*"].
[100] Gabriel J.X. Dance, Michael LaForgia, and Nicholas Confessore, *As Facebook raised a privacy wall, it carved an opening for tech giants*, N.Y. TIMES (Dec. 18, 2018), https://www.nytimes.com/2018/12/18/technology/facebook-privacy.html.
[101] Daniel J. Solove & Danielle Keats Citron, *Risk and Anxiety: A Theory of Data-Breach Harms*, 96 TEX. L. REV. 737, 745 (2018).
[102] Sasha Romanosky & Alessandro Acquisti, *Privacy Costs and Personal Data Protection: Economic and Legal Perspectives*, 24 BERKELEY TECH. L.J. 1061, 1101 (2009).
[103] Sasha Romanosky, David Hoffman & Alessandro Acquisti, *Empirical Analysis of Data Breach Litigation* 11 J. EMPIRICAL LEGAL STUD. 74, 74–75, 93 (2014).

Electronic copy available at: https://ssrn.com/abstract=3794780

held grows.[104] A breach, moreover, may trigger tort liability under state or federal law. William McGeveran has identified fourteen different legal regimes covering data breaches, which together create "a common set of standards for data security in the United States."[105]

A third possible retail concern is that certain technological affordances built into an application or an appliance create pervasive and unavoidable privacy risk. Consider two examples. The first comes from the emergent sensing net: Digital assistants such as Siri and Alexa do not just record and transmit ambient conversation (as smart-phones do), they also detect and map movement and behavior using "lidar."[106] Google's home alarm system initially recorded ambient noise—without disclosing this to purchasing home owners.[107] The decision to build these affordances into the device arguably presents a distinct challenge to privacy norms.[108]

Second, Facebook "leverage[s] the code on third-party sites and apps used to deliver other Facebook products--Like buttons, Login buttons, conversion tracking pixels, retargeting pixels, and the Facebook software development kit--for the additional new purpose of tracking users" as they interact with other web sites.[109] As Dina Srinivasan explains, "when a consumer visited a website with a Facebook plugin, Facebook piggybacked onto the requests and responses necessary to simply display the plugins, to now also surveil the users of competitor ad sellers" whether or not she uses Facebook.[110] This data then "augment[s]" Facebook's ad targeting.[111] Here, there is plausibly a worry not just about what is disclosed, but also about the changed relation between firms and consumers. The marginal increase in information revelation is substantial, with the firm gaining a whole new kind of insight into millions of its customers without a reciprocal advantage. Worry about privacy bleeds into concern about unequal power.

Beyond these localized, retail worries, there is also a most ambitiously gauged privacy argument against the very enterprise of extracting economic rents from personal data in the first instance. A version of this categorical argument is offered by Shoshana Zuboff. She argues that technologies that acquire personal data operate as "one-way mirror" erasing the possibility of

---

[104] Benjamin Edwards, Steven Hofmeyr & Stephanie Forrest, *Hype and Heavy Tails: A Closer Look at Data Breaches*, 2 J. CYBERSECURITY 3, 4–6 (2016). A variation on the data breach problem is when companies "charge a premium price and deliver a bargain-basement service that falls below industry standards where data security is concerned." Lior Jacob Strahilevitz, *Data Security's Unjust Enrichment Theory*, 87 U. CHI. L. REV. 2477, 2491 (2020).

[105] William McGeveran, *The Duty of Data Security*, 103 MINN. L. Rev. 1135, 1139 (2019).

[106] *How Creepy is Your Smart Speaker?*, THE ECONOMIST, May 11, 2019, https://www.economist.com/leaders/2019/05/11/how-creepy-is-your-smart-speaker.

[107] Taylor Telford, *Google Failed to Notify Customers It Put Microphones in Nest Security Systems*, WASH. POST (Feb. 20, 2019, 11:41 AM EST).

[108] A related concern is that "the accumulation of data, including personal data, by dominant firms [in ways that] entrenc[h] their dominant positions." Giuseppe Colangelo & Mariateresa Maggiolino, *Data accumulation and the privacy–antitrust interface: insights from the Facebook case*, 8 INT'L DATA PRIVACY L. 224, 225 (2018). In February 2019, for instance Germany's competition authority the Bundeskartellamt prohibited Facebook from combining users' WhatsUp, Instagram, and Facebook data streams without their consent. Bundeskartellamt, *Bundeskartellamt prohibits Facebook from combining user data from different sources: Background information on the Bundeskartellamt's Facebook proceeding*, Feb. 7, 2019.

[109] Srinivasan, *supra* note 60, at 71.

[110] *Id.* at 72.

[111] *Id.* Apple has recently applied IPhone users to opt out of much of this tracking, to facebook's manifest dismay.

Electronic copy available at: https://ssrn.com/abstract=3794780

interiority by rendering the psychologically internal into a digital feed.[112] What is lost as a result, she argues, is "the sanctity of the individual, the ties of intimacy, the sociality that binds us together in promises and the trust they breed."[113] Surveillance capitalism, she argues, "rob[s] us of the life-sustaining inwardness, born in sanctuary, that finally distinguishes us from machines."[114] Although Zuboff is not entirely clear on what reform follows from this, her analysis is at the least to consistent with a call for a radical contraction in personal data economies.

The wholesale critique of personal data economies offered by Zuboff prickles with difficulties.[115] Her claim that "intimacy" and "sociality" are fatally comprised by social networks is overdrawn. Connections created or sustained during the pandemic through platform economies such as Facetime and Zoom belie her cynicism. Gauging concerns about privacy in such sweeping terms also risks losing sight of more specific ways in which data economies can conduce to disclosure-based harms. Moreover, Zuboff importantly misses the way in which platform economies and sensing nets can create valuable public goods. In proceeding, therefore, I will focus on retail rather than wholesale privacy critiques.

B. Autonomy

A second common theme in critiques of personal data economies concerns the influence gained by platforms and other firms over individuals' autonomy or agency. Both terms are important yet frustratingly imprecise.[116] Autonomy, for example, might be glossed narrowly "as abstract rationality and responsibility attributed to an individual."[117] Alternatively, Margaret Radin has argued, a person might "be bound up with an external 'thing' in some constitutive sense," so as to make "control over that 'thing'" a necessary part of their "autonomy" that the law should recognize.[118] Radin applied this insight to residential property, personal vehicles, and wedding rings.[119] Her logic can be extended to personal data economies.

Several scholars have tried to develop the connection between data and autonomy. None leans on Radin's seminal work. But her approach is echoed in John Cheney-Lippold's claim that "datafied lives … increasingly define who we are and who we can be."[120] His account emphasizes not just the acquisition of data but also the extraction of commercial value via predictions. This ability to intimate future behavior is cast as a deprivation of human autonomy. In a similar vein, Oxford philosopher James Williams argues that platform economies "threaten to frustrate one's authorship of one's own life," such that "the operation of the will … has also been short-circuited and undermined."[121] Zuboff, again, argues that it is a per se wrong when "human nature is scraped,

---

[112] SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER 81 (2018)
[113] *Id.* at 516.
[114] *Id.* at 492.
[115] *See* Mariano-Florentino Cuéllar and Aziz Z. Huq, *Economies of Surveillance*, 133 HARV. L. REV. 1280 (2019) (criticizing Zuboff's analysis).
[116] *See* Aziz Z. Huq, *A Right to a Human Decision*, 106 VA. L. REV. 611 (2020).
[117] Margaret Jane Radin, *Property* and *Personhood*, 34 STAN. L. REV. 957, 960 (1982).
[118] *Id.*
[119] *Id.* at 987, 98, 992-93, 1000.
[120] JOHN CHENEY-LIPPOLD, WE ARE DATA: ALGORITHMS AND THE MAKING OF OUR DIGITAL SELVES 19 (2017).
[121] JAMES WILLIAMS, STAND OUT OF OUR LIGHT: FREEDOM AND RESISTANCE IN THE ATTENTION ECONOMY 88 (2019).

Electronic copy available at: https://ssrn.com/abstract=3794780

torn, and taken, for another century's market product" because users are being transformed into "means to others' ends."[122] Zuboff uses forceful language redolent of Radin's to critique the use of personal data from platform economies for behavioral predictions.[123] These defy "millennia of human contest and sacrifice" by denying "the freedom of the will."[124] In the legal academy, Tal Zarsky has defined "manipulation" in terms that cover all predictive applications of personal data to which users are "oblivious."[125] Stanching such manipulation means changing the terms of the personal data economy, since "personal data is the fuel of the manipulation process."[126]

Like the wholesale privacy critique lodged by Zuboff, the argument from autonomy seems overbroad. A key question is why the interventions powered by personal data are so intrusive or offense as to violate individual autonomy. Facebook and Google monetize their personal data streams through the sale of digital advertising.[127] To be sure, these manipulate their audience; but the same is true of *all* advertising. Absent concerns about capacity (e.g., for minors), it is hard to see how even precisely targeted digital advertising is a fatal attack on human autonomy. They simply do not all comprehensively "structure users' conditions of possibility."[128] Further, the vague intuition that data-based interventions are 'too effective' does not cash out as a clear line that can be used to distinguish beneficial forms of targeted advertisement from improper manipulation. More modestly, and more plausibly, concerns about autonomy are probably best glossed as worries about the ability of platform economies and data brokers to seize a disproportionate share of the economic surplus created by personal data economies. Again, the question is at bottom one of power and distribution.

C. Retail economic exploitation

Adjacent to these autonomy concerns, and often expressed in the same terms, is a worry about the economic exploitation of individual users and contributors. This objection runs against both the "dragnet" and the "intervention" stages of data economies.[129] I offer here three examples of arguable exploitation. The first two touch on how personal data is extracted by platform economies. The third relates to how first-degree price discrimination arises at the back-end.

A first problem arises if users do not understand that platform economies or sensing nets are acquiring their data. If that happens, "a subsidy is given to those data-processing companies

---

[122] ZUBOFF, *supra* note 112, at 94.
[123] Her argument presumably would apply with greater force to sensing nets.
[124] *Id.* at 331-32; *see also* Daniel Susser, Beate Roessler, and Helen Nissenbaum, *Online Manipulation: Hidden Influences in A Digital World*, 4 GEO. L. TECH. REV. 1, 3–4 (2019) (suggesting that "[t]he information we volunteer and shed about our interests, preferences, desires, emotional states, beliefs, habits, and so on, provides everything a would-be manipulator needs to know about how to subvert our decision-making").
[125] Tal Z. Zarsky, *Privacy and Manipulation in the Digital Age*, 20 THEORETICAL INQUIRIES L. 157, 169 (2019).
[126] *Id*. at 186. The objection that Zuboff and Zarsky advance seems to be reflect a concern that data, even if not consciously disclosed, "such as the web browser we use and who we call on the phone, has constitutive effects on our love." CHENEY-LIPPOLD, *supra* note 120, at 195. But why are such "constitutive effects" different in kind from other viscidities of good or bad fortune beyond an individual's control? Such critiques help themselves to unwarranted assumptions about the baseline extent of human agency.
[127] *See, e.g.*, *Bass v. Facebook, Inc.*, 394 F. Supp. 3d 1024, 1028 (N.D. Cal. 2019) (noting that 96% of Facebook's revenue comes from targeted advertising).
[128] CHENEY-LIPPOLD, *supra* note 120, at 169.
[129] Fourcade & Healy, *supra* note 41, at 11.

Electronic copy available at: https://ssrn.com/abstract=3794780

that exploit personal data. As a result, these organizations are not charged the true 'cost' of the personal data they acquire."[130]

A second problem arises because—notwithstanding their sheen of novelty and innovation—platform economies reproduce gendered divisions of labor and reward that have long plagued industrial capitalism. In an earlier era, it was believed that "[t]he male head of the household would be paid a family wage, sufficient to support children and a wife-and-mother, who performed domestic labor without pay."[131] While this "family wage" concept has collapsed, a similar gendered division of labor persists in personal data economies. There, (usually male) architects of a platform's code tend to be highly remunerated, whereas content contributors go without any financial reward. In the dominant discourse employed in the tech sector, this "primacy of the platform" is justified by the assumption that content creation—whether on review services such as Yelp or postings on Etsy or another social media platform—just "isn't work."[132] The labor of content providers, "a female-dominated sector of the economy," are instead equated to "domestic, especially female, labor"[133] that the family-wage model treats as unpaid. Platform economies thus are a "reinvention of the family as an instrument for distributing wealth and income."[134]

Social networks that supply connectivity rather than content operate on a similarly gendered logic. Sociologists Marion Fourcade and Daniel Kluttz argue that these networks' acquisition of personal data rely upon preexisting social structures of trust, consent, and gift giving.[135] Platform economies subtly exploit a "natural compulsion to reciprocate" and "existing solidaristic bond[s]" to generate a circulatory system of interactions ripe with personal data.[136] The "affective labor" that is the ordinary work of human contact and interaction is seized and transformed into an informational asset.[137] The marketing of consumer genetics similarly appeals to altruism. It asks consumers to "give back" and to participate in a "crowdsourced healthcare revolution."[138] As such, platform economies rely on subtle (and gendered) emotional manipulation to commercial ends.

Third, at the last step of its economic cycle, data can be used to enable first-degree price discrimination by which different consumers are presented with variable, individualized prices for

---

[130] Schwartz, *Property*, *supra* note 99, at 2079.
[131] NANCY FRASER, FORTUNES OF FEMINISM: FROM STATE MANAGED CAPITALISM TO NEOLIBERAL CRISIS 111 (2013).
[132] ADRIAN DAUB, WHAT TECH CALLS THINKING 50-51 (2020).
[133] *Id*.
[134] MELINDA COOPER, FAMILY VALUES: BETWEEN NEOLIBERALISM AND THE NEW SOCIAL CONSERVATISM 17 (2017). Cooper argues that such reinventions are "periodic," and endemic to capitalism as an economic arrangement. *Id.*
[135] Marion Fourcade and Daniel N. Kluttz, *A Maussian bargain: Accumulation by gift in the digital economy*, 7 BIG DATA & SOC. 1, 10 (2020).
[136] *Id.*
[137] Kim Doyle, *Facebook, Whatsapp and the commodification of affective labour*, 48 COMMUNICATION, POL. & CULT. 51, 61-62 (2015) ("Facebook attracts and maintains users through the affective investments users make with the platform to the monetarization of these platforms.").
[138] Geiger & Gross, *supra* note 67, at 18-19.

Electronic copy available at: https://ssrn.com/abstract=3794780

the same product.[139] This can harm users. For example, data brokers crunch the information voluntarily or unwittingly supplied by consumers to offer financial goods, such as "subprime credit," that may be inflict a heavy toll.[140] Individualized harms might also be imposed by the sharing of erroneous data that serves to limit a person's access to credit or goods.[141] And while first-degree price discrimination can reflect consumer preferences, it also allows sellers to "extract[] the entire surplus by setting a price that is just below each consumer's [willingness to pay]."[142] Further, where such discrimination takes advantage of consumer misperceptions, consumers can end up strictly worse off.[143] Even apart from welfare effects, these dynamics are "a profound challenge to the distribution of wealth between producers and consumers."[144] Even if Pareto optimal, that is first-degree price discrimination conduces to objectionable forms of inequality—a possibility that leads us to the next critique.

D. Structural economic inequality

Concerns about the equities of interaction between platforms and users or content-providers can be scaled-up into a systemic, macroeconomic concern about the distribution of economic rents from data economies and about their dynamic effects upon labor markets. The link between data economies and inequality is often framed in vague and suggestive terms, using economy-wide trends.[145] We can be more precise. I chart briefly here three specific pathways by which data economies exacerbate aggregate economic inequalities or generate new forms of inequitable hierarchy.[146]

To begin with, personal-data economies is facilitating new Taylorite strategies of automated scheduling, task redefinition, loss- and risk-prediction in the workplace. These can have major dignitary, distributive and economic consequences.[147] A California company called Humanyze, for example, offers employers technologies to track what their workers sit, where they

---

[139] Alessandro Acquisti, Curtis Taylor, and Liad Wagman, *The Economics of Privacy*, 54 J. ECON. LIT. 442, 466 (2016) ("Tracking and measurability, in addition to websites' ability to dynamically update and personalize prices for each visitor, are bringing online markets closer to the theoretical scenario of first-degree price discrimination."). For a discussion of possible first-degree price discrimination in the ride-sharing context, see Ryan Calo & Alex Rosenblat, *The Taking Economy: Uber, Information, and Power*, 117 COLUM. L. REV. 1623, 1659 (2017).

[140] Roderick, *supra* note 73, at 732.

[141] *See* John Lucker, Susan K. Hogan & Trevor Bischoff, *Predictably Inaccurate: The Prevalence and Perils of Bad Big Data*, 21 DELOITTE REV. (July 31, 2017), https://www2.deloitte.com/us/en/insights/deloitte-review/issue-21/analytics-bad-data-quality.html (describing "the potential prevalence and types of inaccurate data from US-based data brokers").

[142] Oren Bar-Gill, *Algorithmic Price Discrimination When Demand Is A Function of Both Preferences and (Mis)perceptions*, 86 U. CHI. L. REV. 217, 220–21 (2019).

[143] *Id.* at 221. For a similar analysis of "behavioral discrimination," see ARIEL EZRACHI AND MAURICE E. STUCKE, VIRTUAL COMPETITION: THE PROMISE AND PERILS OF THE ALGORITHM-DRIVEN ECONOMY 31-32, 117-30 (2016).

[144] Ramsi A. Woodcock, *Big Data, Price Discrimination, and Antitrust*, 68 HASTINGS L.J. 1371, 1374 (2017).

[145] *See, e.g.,* Amy Kapczynski, *The Law of Informational Capitalism*, 129 YALE L.J. 1460, 1477 (2020) (drawing on general economic trend data to suggest that "[i]nformation technologies plausibly accelerate such winner-take-all dynamic").

[146] In effect, I am specifying in more detail Katharina Pistor's claim that "big data … is the power to transform free contracting and markets into a controlled space that gives a huge advantage to sellers over buyers." Katherina Pistor, *Rule by Data: The End of Markets?*, 83 L. & CONTEMP. PROB. 100, 117 (2020).

[147] *See* Ifeoma Ajunwa, Kate Crawford & Jason Schultz, *Limitless Worker Surveillance*, 105 CALIF. L. REV. 73 5, 738-39 (2017); Pegah Moradi and Karen Levy, *The Future of Work in the Age of AI*, in THE OXFORD HANDBOOK OF ETHICS OF AI (Markus D. Dubber, Frank Pasquale, and Sunit Das, eds., 2020).

Electronic copy available at: https://ssrn.com/abstract=3794780

are in an office, how quickly they move, how much time they spend speaking to people of the same gender, and the amount of time they spend listening or speaking.[148] Its tool allows the extraction of a greater share of surplus value from workers, and hence a greater share of profits, without any concomitant benefit for workers.[149] Quite apart from its creepiness, it can also thwart collective action. This further inhibits employees from securing a greater share of economic rents.[150]

Second, and relatedly, the business model of certain platform-based applications may generate a surplus simply not because of data, but through the exploitation of gaps in the legal protections for workers. Loss-making ride-share companies such as Uber, for example, exploit the difference between the legal protection for employees and for contractors as a way to offer lower prices than taxi companies.[151] Obviously, this is to the considerable detriment of labor. Regulatory arbitrage may be only part of the business model of these platform economies.[152] But when coupled to the increasing control of the workplace, such platform-economy tools have the potential to dramatically alter how gains from commerce are distributed between employers and workers.

Third, and more broadly, data economies supply new predictive tools that substitute for human capital. These shape the overall labor market and drive economic inequality. Data-driven machine learning is a "general purpose technology."[153] It is deployed in lieu of human labor across many different workplaces. Advances in machine-learning since the 1980s have enlarged the slice of the labor market for which automation is a substitute. Moore's law means that the cost of such substitution falls over time. Some economists argue that the resulting loss of jobs likely will strike lower-income blue- and white-collar positions hardest.[154] Others attribute declining labor share to declining productivity and a global manufacturing capacity glut.[155] Either way, among the sectors of the labor market most likely to be undermined by automation are "[o]ffice and administrative support, production, transportation and logistics, food preparation, and retail jobs."[156] For instance,

---

[148] *There Will Be Little Privacy in the Workplace of the Future*, THE ECONOMIST (Mar. 28, 2018), https://www.economist.com/special-report/2018/03/28/there-will-be-little-privacy-in-the-workplace-of-the-future.
[149] *See also* ZUBOFF, *supra* note 112, at 409.
[150] Brishen Rogers, *The Law and Political Economy of Workplace Technological Change*, 55 HARV. C.R.-C.L. L. REV. 531, 542 (2020).
[151] DAUB, *supra* note 132, at 6-7. On Uber's losses, see Andrew J. Hawkins, *Uber reports $2.9 billion quarterly loss during pandemic,* THE VERGE, (May 7, 2020), https://www.theverge.com/2020/5/7/21251111/uber-q1-earnings-rides-loss-eats-delivery-coronavirus. For a slightly different critique of Uber as "treat[ing] drivers both like consumers and like workers," and exploiting that ambiguity, see ALEX ROSENBLATT, UBERLAND: HOW ALGORITHMS ARE REWRITING THE RULES OF WORK 206-07 (2018).
[152] Brishen Rogers, *The Social Costs of Uber*, 82 U. CHI. L. REV. DIALOGUE 85, 87 (2015) (arguing that Uber's "success is not based just on regulatory arbitrage" but "in having reduced … transaction costs").
[153] Erik Brynjolfsson & Tom Mitchell, *What Can Machine Learning Do? Workforce Implication*s, 358 SCIENCE 1530, 1530 (Dec. 22, 2017).
[154] FREY, *supra* note 35, at 298-99 ("The employment prospects for the middle class crucially hinge upon what computers can or cannot do."); *see also* David Autor & Anna Salomons, *Is Automation Labor-Displacing? Productivity Growth, Employment, and the Labor Share*, Brookings Papers on Econ. Activity 8 (Mar. 8, 2018), https://www.brookings.edu/wp-content/uploads/2018/03/1_autorsalomons.pdf [https://perma.cc/4XQ4-5NVA] (finding that "automation has become increasingly labor-displacing in recent decades, both at the industry level and in aggregate").
[155] Aaron Benanav, *Automation and the Future of Work—1*, 119 NEW LEFT REV. 5, 37-38 (2019).
[156] *Id.* at 320.

Electronic copy available at: https://ssrn.com/abstract=3794780

in 2018, Google announced its development of technology to replace call-center workers.[157] This exacerbates a secular trend of reduced demand for blue-collar and low-skill positions, coupled to greater demand for professional and managerial positions.[158] Slimmer opportunities for middle-skilled workers, coupled to greater demand for more skilled labor, polarizes incomes—at least without a social wage.[159] Income and wealth polarization have further ramifications. For example, it is increasingly the case that wealth provides a way to opt out of the personal-data economies run by platform economies and data brokers.[160]

* * *

In summary, all three dynamics surfaced here "set the terms that structure an engagement" in which information is first acquired.[161] The affective labor of social connection upon which Facebook rests, or the critical engagement with goods and services that are reflected in Yelp and Amazon reviews, yields little or no direct return.[162] Where participation in a data economy allow a firm like Facebook or Google to acquire information beyond an immediate consumer or user interaction—for example, by tracking activity on third-party websites—the payoff to consumers is slight. Instead, platforms, data brokers, and device manufacturers shape the conditions in which users generate personal data, their awareness of doing so, and their expectations about how such data will be used. This maximizes revenue for firms while suppressing payouts to users. All this allows firms to mold how and when consumers disclose personal data without demanding a payoff. An asymmetry in the distribution of rents is thereby baked into the architecture of personal data economies as presently structured.

Putting these three arguments together reveals an overarching structural dynamic that echoes Thomas Piketty's more general critique of contemporary economic arrangements. According to Piketty's now famous formulation, the return to capital under contemporary conditions has tended to exceed the growth rate. A result of this imbalance is a steady increase in the Gini coefficient thanks to accelerating increases in high-end wealth concentrations.[163] The three structural critiques of personal data economies echo Piketty's point that technology, among other factors, "influence[s] the relative power of different social groups."[164] Although his analysis places greater weight on nontechnological factors, the parallel suggests that personal data

[157] FREY, *supra* note 35, at 306-07.
[158] CARLES BOIX, DEMOCRATIC CAPITALISM AT THE CROSSROADS: TECHNOLOGICAL CHANGE AND THE FUTURE OF POLITICS 102-08 (2019).
[159] *Id.* at 118-23; *see also* Cynthia Estlund, *What Should We Do After Work? Automation and Employment Law*, 128 YALE L.J. 254, 280 (2018) ("Workers without in-demand skills will compete, and drive down wages, for the jobs that machines cannot do as well or as cheaply but that most humans can do. The winners--those who make or own the technology or whose scarce skills are augmented by technology--will win access to private enclaves of privilege fortified against the rage and resentment of the losers.")
[160] Joseph W. Jerome, *Buying and Selling Privacy: Big Data's Different Burdens and Benefits*, 66 STAN. L. REV. ONLINE 47, 48 (2013) (describing "social networks where users join for a fee and the rise of reputation vendors that protect users' privacy online").
[161] Crain, *supra* note 72, at 12.
[162] Consumers and users, it might be argued, particulate in the production of a collective epistemic or social good, and benefit insofar as they derive some value from that good.
[163] THOMAS PIKETTY, CAPITAL IN THE TWENTY-FIRST CENTURY 25 (Arthur Goldhammer trans. 2014).
[164] *Id.* at 305 (noting that "if the supply of skills does not increase at the same pace as the needs of technology, then groups whose training is not sufficiently advanced will earn less and be relegated to devalued lines of work").

Electronic copy available at: https://ssrn.com/abstract=3794780

economies are likely to follow, and inscribe more deeply, wealth hierarchies accreting during the last half century.

The structural critique developed here might be parried by noticing that firms in the data economy offer consumers services that are worth as much, or more, than the profits accruing to firms. Facebook's 2019 revenues exceeded 70 billion dollars.[165] Although Facebook does not charge an access fee, experimental tests of consumers' willingness to abandon the social network offer a glimpse of its value. One experimental study of American liberal arts college arts students and MTurk participants (all American) found participants asked for around $2,000 to deactivate Facebook for a year.[166] That study concluded that "the most conservative … estimates, if applied to Facebook's 214 million U.S. users, suggests an annual value of over $240 billion to users."[167]

Yet this does not defeat the arguments from structural economic inequality. To begin with, there is no consensus on how much value Facebook is creating for users. Studies, in practice, find a range of valuations, including ones that undermine the conclusion that Facebook creates net social benefits.[168] Some suggest that users' willingness to pay for Facebook is much smaller than their willingness to accept payoffs for deactivating Facebook.[169] Moreover, there is some evidence that quitting Facebook has positive effects on both physical and mental health.[170] It is not clear that studies of Facebook's self-reported value to users capture these spillovers. Pending better evidence, it seems better to view the welfare analysis as inconclusive.

More importantly, the arguments from structural economic inequality apply here even if data economies are Pareto efficient. Even if a particular data economy can be justified as a boost to overall welfare, it might still shift the distribution of entitlements in a society in undesirable ways. Creating a bit more wealth overall at the cost of substantially greater economic inequality is a tradeoff that many view as regrettable.[171]

E. Democratic backsliding

Can personal data economies erode democratic norms? Since the November 2016 election, much public and political concern has focused on the possibility that personal data generated by

---

[165] Facebook Reports Fourth Quarter and Year-End Results (Jan. 20, 2020), https://investor.fb.com/investor-news/press-release-details/2020/Facebook-Reports-Fourth-Quarter-and-Full-Year-2019-Results/default.aspx

[166] Jay R. Corrigan, et al., *How much is social media worth? Estimating the value of Facebook by paying users to stop using it*, 3.12 *Plos one* e0207101 (2018).

[167] *Id.*; *see also* Bodo Herzog, *Valuation of digital platforms: experimental evidence for Google and Facebook*, 6 INT'L J. FIN. STUD. 87 (2018) (finding a "weighted average willingness-to-pay (WTP) of 28.26 € for Facebook per week").

[168] Erik Brynjolfsson, Avinash Collis, and Felix Eggers, *Using massive online choice experiments to measure changes in well-being*, 116 PROC. NAT'L ACAD. SCI. 7250, 7252 (2019) (estimating willingness to accept two to three hundred dollars a year to abandon Facebook).

[169] Cass R. Sunstein, *Valuing Facebook*, 4 BEHAVIOURAL PUB. POL. 370, 370-71 (2020).

[170] Roberto Mosquera et al., *The economic effects of Facebook*, 23 EXPERIMENTAL ECON. 575, 592 (2020) (finding, based in an experimental study of more than 1,700 users that "a one-week Facebook restriction decreased feelings of depression and increased engagement in healthier activities").

[171] *See* ANTHONY ATKINSON, INEQUALITY: WHAT CAN BE DONE 11-23 (2015) (summarizing both instrumental and intrinsic reasons for concern about high levels of economic inequality within a nation).

Electronic copy available at: https://ssrn.com/abstract=3794780

platform economies can be used to foster polarization and political extremism.[172] Worries about "fake news" arise against a context of rising "partisan sectarianism," in which "out-party hate" has become the leading predictor of voting behavior.[173] The politics of enmity, engorged by misinformation-filled platform economies, is viewed as a serious destabilizing risk to a democratic system. As the January 2021 attack on the U.S. Capital graphically illustrated, if people believe that a loss at the polls will lead to the irreparable calamity of an opposition's victory, it will pursue extreme, anti-democratic measures to head off that prospect.

Platform economies arguably contribute to processes of democratic backsliding in three ways. First, Facebook and Twitter enable "junk news circulation."[174] After the 2016 presidential election, for example, the 569 sites known to reliably disseminate the highest number of false news stories received some 60 million Facebook engagements per month.[175] This dissemination undermines the epistemic predicate of effectual democratic choice. It can also exacerbate racist, xenophobic, and anti-Semitic ideas and movements.[176] Second, a platform that depends on user engagement has an economic incentive to promote polarizing content that induces users to spend more time on the site. Facebook, for example, is said to have refused to make its content less divisive for this reason.[177] Third, user data harvested from platforms can be analyzed to guide misinformation campaigns.[178] Precise, individual-level information, gathered without a user's knowledge, is used to guide "techniques reliant on subterfuge and opacity" with the aim of changing voting behavior.[179] Among the most notorious examples of political redlining is the use of Facebook data by the British firm Cambridge Analytica on behalf of the Trump campaign.[180]

---

[172] For summaries of these concerns, see COHEN, *supra* note 47, at 75-77, 83-89, 96; *see also* Michael J. Abramowitz, *Stop the Manipulation of Democracy Online*, N.Y. TIMES (Dec. 11, 2017), https://www.nytimes.com/2017/12/11/opinion/fake-news-russia-kenya.html

[173] Eli J. Finkel et al., *Political sectarianism in America*, 370 SCIENCE 533, 533 (2020).

[174] PHILIP N. HOWARD, LIE MACHINES: HOW TO SAVE DEMOCRACY FROM TROLL ARMIES, DECEITFUL ROBOTS, JUNK NEWS OPERATIONS, AND POLITICAL OPERATIVES 12 (2020); *see also* SIVA VAIDHYANATHAN, ANTISOCIAL MEDIA: HOW FACEBOOK DISCONNECTS US AND UNDERMINES DEMOCRACY 16 (2019) ("Facebook is … the most pervasive and powerful catalyst of information pollution and destructive nonsense.").

[175] Hunt Allcott, Matthew Gentzkow, and Y. Chuan, *Trends in the diffusion of misinformation on social media*, 6 RES. & POL. 1, 2 (2019).

[176] Hate groups use Facebook ads a "vehicle" for constructing narratives to justify denigration of violence against minority groups through the propagation of narratives of those groups as "separate and hostile." Megan Squire, *Network, Text, and Image Analysis of Anti-Muslim Groups on Facebook*, 2019 J. WEB SCIENCE 1-2 (2019), https://webscience-journal.net/webscience/article/view/77. About half of the misogynistic and racists posts that violate Facebook's community standards are not taken down, even when they are reported to the company. Caitlin Ring Carlson and Hayley Rousselle, *Report and repeat: Investigating Facebook's hate speech removal process*, 25 FIRST MONDAY (2020), https://doi.org/10.5210/fm.v25i2.10288.

[177] Jeff Horwitz and Deepa Seetharaman, *Facebook executives shut down efforts to make the site less divisive*, WALL ST. J. (March 26, 2020), https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499?mod=hp_lead_pos5.

[178] Platform-generated data need not be political in nature to reveal users' policy and partisan preferences. HOWARD, *supra* note 174, at 157 ("Our credit card data and city travels generate data for political inference, whether we intends to allow this or not.").

[179] Zeynap Tufekci, *Engineering the public: Big data, surveillance and computational politics*, 19 FIRST MONDAY (2014), https://doi.org/10.5210/fm.v19i7.4901

[180] Carole Cadwalladr & Emma Graham-Harrison, *Revealed: 50 Million Facebook profiles harvested for Cambridge Analytica in major data breach,* GUARDIAN (Mar. 17, 2018, 6:03 PM), https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election.

Electronic copy available at: https://ssrn.com/abstract=3794780

The scale and effect of these campaigns is not clear. Platforms are not the sole or most important driver of partisan sectarianism.[181] It is "tough to estimate" the causal effect of false news or ads on voting behavior.[182] The best empirical studies of the Cambridge Analytica's campaign suggest that its impacts "are likely exaggerated," and that the use of tailored online advertising to change voter behavior remains "primarily an act of faith."[183] At the same time, experimental evidence suggests that individual disengagement from social media is associated with both lower levels of political knowledge and reduced partisan polarization.[184] So perhaps the polity as a whole would be less informed but also less divided in the absence of platform economies. Moreover, domestic misinformation raises distinct First Amendment issues from foreign actions. If platforms are a particularly effective way of reaching citizens—say, in particular those who ordinarily do not vote—should its 'manipulative' potential be enough to place it beyond bounds for political campaigns? The case against online misinformation is more complex than first appears.

F. State domination

In a recent study of democracy's history, the political scientist David Stasavage has argued that autocratic forms of government have prevailed over nascent democracies "whenever new or improved technologies reduced the information advantage that members had over rulers."[185] The same dynamic is plausibly at work today: Personal data economies produce epistemic resources that can be leveraged by a state seeking to exercise undemocratic power over its citizens.

In the United States, worry about undemocratic power focuses on how policing agencies are leveraging personal data.[186] There has been a robust debate on how the content and metadata generated by telephone calls can be used by the federal government. To date, there has been no known repetition of the large-scale surveillance of domestic political opposition that characterized the 1960s.[187] Overseas, the exploitation of personal data economies for repression is more advanced. The Communist Party of China has proved adept at harnessing personal data. In the Western province of Xinjiang, pervasive digital surveillance is used to identify "the digital footprint of unauthorized Islamic practice."[188] Across the whole country, public surveillance

[181]The United States has experienced increasing out-party hate since the late 1980s, well before broad availability of platform economies. Finkel et al., *supra* note 173, at 534.
[182] HOWARD, *supra* note 174, at 110.
[183] YOCHAI BENKLER, ROBERT FARIS, AND HAL ROBERTS, NETWORK PROPAGANDA: MANIPULATION, DISINFORMATION, AND RADICALIZATION IN AMERICAN POLITICS 279 (2018); *id.* at 347-48 (underscoring the critical role of "professional,. Commercial, and nonprofit think tanks," especially on the political right, in amplifying political misinformation and thereby turning a "networked public sphere" into a "networked propaganda system")
[184] Hunt Allcott et al., *The welfare effects of social media*, 110 AM. ECON. REV. (2020).
[185] DAVID STASAVAGE, THE DECLINE AND RISE OF DEMOCRACY: A GLOBAL HISTORY FROM ANTIQUITY TO TODAY 97 (2020).
[186] The most recent line of work concerns spatially distributed forms of data acquisition. *See, e.g.,* Andrew Guthrie Ferguson, *Structural Sensor Surveillance*, 106 IOWA L. REV. 47, 49 (2020) (asking whether "smart city sensors [are] unconstitutional because they inadvertently allow for aggregated government collection of personal data without a probable-cause search warrant?"); Gabriel Bronshteyn, *Searching the Smart Home*, 72 STAN. L. REV. 455, 457 (2020) ("Law enforcement and intelligence agencies have already begun taking notice of the vast quantities of profoundly intimate data being generated from within the 'smart home.'").
[187] *Cf.* BRUCE SCHNEIER, DATA AND GOLIATH: THE HIDDEN BATTLES TO COLLECT YOUR DATA AND CONTROL YOUR WORLD 120-25 (2015); ANGWIN, *supra* note --, at 46-50 (expressing doubt on the utility of surveillance as a national security tool).
[188] Darren Byler, *China's Hi-tech War on its Muslim minority*, GUARDIAN, Apr. 11, 2019.

Electronic copy available at: https://ssrn.com/abstract=3794780

cameras are integrated with facial recognition and artificial intelligence to create "a vast and unprecedented national surveillance system."[189] China has also created a global market in the tools of digital totalitarianism.[190] Even democratic governments unlikely to purchase China's surveillance tools are exploiting personal data for state security. In India, the BJP government introduced in 2020 a Personal Data Protection bill that requires platform economies to engage in "data localization"—the storage of personal data within India.[191] The bill also exempts the Indian state from most limits on the acquisition and use of such data.[192] By combining data localization and unfettered state access to locally stored data, the bill pumps up the risk of state repression.[193]

The problem of state repression through personal data economies, in short, is intractably entangled with privacy, autonomy, exploitation, and inequality worries. The more the state regulates the flows and usages of personal data, the greater its ability to leverage its regulatory authority to repressive ends.

G. The Underproduction of Public Goods

Critical commentary on personal data economies focuses on potential harms. Little attention is paid to the mirror-image problem of socially valuable applications of personal data being foregone.[194] Personal data economies typically realize profits via the increasingly precision in targeted advertising. Data produced through platform economies and sending nets, however, might also be deployed to target resources and interventions to improve public health, to uphold environmental standards, to facilitate smooth traffic flows, and to identify and deliver much-needed public services to marginalized populations.[195] Making machine-learning instruments developed using personal data available for the benefit or use of the general public "could facilitate the production of genuinely innovative products on a relatively low budget."[196] And personal data economies could give the public a better grasp on how well its elected representatives are performing. For many things the state does, little is known about quality or efficacy. For example,

---

[189] Paul Mozur, *Inside China's Dystopian Dreams: A.I., Shame, and Lots of Cameras*, N.Y. TIMES, July 8, 2018.
[190] Paul Mozur, Jonah M. Kessel & Melissa Chan, *Made in China, Exported to the World: The Surveillance State* N.Y. TIMES, Apr. 24, 2019, https://www.nytimes.com/2019/04/24/technology/ecuador-surveillance-cameras-police-government.html.
[191] Manasi Gopalkrishnan, *India's personal data privacy law triggers surveillance fears*, DW, (Nov. 11, 2020), https://p.dw.com/p/3l8yr.
[192] *Id.*
[193] *Id.* (noting recent surveillance and intimidation of political activists by the Indian government).
[194] Indeed, it is more common to find the opposite claim that "databases [of personal information] have no significant public good characteristics" and instead "are the paradigmatic example of a good whose entire value is privately appropriable. Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject As Object*, 52 STAN. L. REV. 1373, 1388 (2000) [hereinafter "Cohen, *Examined Lives*"]. If this was true in 2000, I think it is no longer the case.
[195] For examples in each of these fields, see Muin J. Khoury and John Ioannidis, *Big data meets public health*, 346 SCIENCE 1045, 1054-55 (2014) (identifying epidemiological problems that can be addressed using personal data); Debra Lam & John Wagner Givens, *Small and Smart: Why and How Smart City Solutions Can and Should Be Adapted to the Unique Needs of Smaller Cities*, 12 NEW GLOBAL STUD. 21, 31-32 (2018) (discussing the use of sensor nets to monitor municipal water quality); SHEKHAR AND VOLD, *supra* note --, at 14-15 (discussing Los Angeles' use of spatial computing to mitigate traffic blockages); Blake E. Reid, *Internet Architecture and Disability*, 95 Ind. L.J. 591, 592 (2020) (discussing the provision of services to the disabled through "smart city" initiatives).
[196] Evgeny Morozov, *Digital Socialism: The Calculation Debate in the Age of Big Data*, 116 NEW LEFT REV. 33, 63 (2019). Release of the data itself would generate privacy objections.

Electronic copy available at: https://ssrn.com/abstract=3794780

information about the efficacy and the social costs of policing is scarce.[197] Personal data economies can be leveraged to fill these gaps. The failure to do so has steep opportunity costs.

The thwarted potential of socially beneficial uses of personal data is illustrated by the failure to use locational data generated by cell-phones to map the trajectories of Covid-19 infections. "COVID-19 moves too quickly through the population to be amenable to standard [manual] contact tracing methods."[198] In the time that manual tracers acquire information from an infected person, and then reach their contacts, the virus may have spread to hundreds more. To remedy his gap, digital contact tracing uses data generated by cellphones in two different ways. More ambitiously, personal data can be used to identify intersections between the movements of a specific infected person and others. It thus can facilitate the construction of a catalog of those who need to be warned of potential infection.[199] More modestly, locational data generated by cellphones can be used to model the population-level diffusion of the virus.[200] This allows public-health authorities to identify specific locations that act as high-frequency transmission nodes.[201] It can also help when calculating estimates of contagion rates with different combinations of closures and openings.[202] Digital contract tracing, however, has largely failed to take root in the United States.[203] Its failure is just one of the missed opportunities to exploit personal data economies to create important public goods.

Sharing data for the production of public good raises privacy concerns. Both Canada and the United Kingdom, though, have developed protocols for sharing data produced by state entities without compromising privacy. Under Canada's Statistics Act, for example, researchers can use a "Real Time Remote Access System" that enables data to be queried without being exposed.[204] The United Kingdom's Digital Economy Act creates protocols for the sharing of deidentified data with accredited researchers.[205] Generalizing from these examples, Lisa Austin and David Lie have

---

[197] Barry Friedman & Elizabeth G. Jánszky, *Policing's Information Problem*, 99 TEX. L. REV. 1, 33 (2020).
[198] Michael J. Parker et al., *Ethics of instantaneous contact tracing using mobile phone apps in the control of the COVID-19 pandemic*, -- J. MED. ETH. --, at *1 (2020); *see also* Michelle M. Mello and C. Jason Wang, *Ethics and governance for digital disease surveillance*, 368 SCIENCE 951, 952 (2020) ("Serious doubts have been raised about whether traditional methods of contact tracing can arrest the COVID-19 epidemic.").
[199] Ada Lovelace Institute, *Exit through the App Store? A rapid evidence review of the technical considerations and societal implications of using technology to transition from the COVID-19 crisis* 22 (Apr. 2020), https://www.adalovelaceinstitute.org/wp-content/uploads/2020/04/Ada-Lovelace-Institute-Rapid-Evidence-Review-Exit-through-the-App-Store-April-2020-2.pdf.
[200] Serina Chang, et al., *Mobility network models of COVID-19 explain inequities and inform reopening*, -- NATURE 1, 1 (2020).
[201] *Id.* at 4. This is a particularly effective intervention for Covid-19 because a small proportion of cases seem to responsible for a high proportion of infections. Dyani Lewis, *Why many countries failed at COVID contact-tracing-but some got it right*, 588 NATURE 384, 386 (2020). For a discussion of how advances in data science facilitate the identification of "hot spots" where contagion frequency rises, see SHEKHAR AND VOLD, *supra* note 195, at 177-81
[202] Joshua Graff Zivin and Nicholas Sanders, *The spread of COVID-19 shows the importance of policy coordination*, -- PROC. NAT'L ACADEMY OF SCI. – 1, 3 (2020).
[203] C. Aschwanden, *Contact tracing, a key way to slow COVID-19, is badly underused by the US*, SCIENTIFIC AMERICAN (Jul. 21, 2020), https://www.scientificamerican.com/article/contact-tracing-a-key-way-to-slow-covid-19-is-badly-underused-by-the-u-s/.
[204] Statistics Act, R.S.C. 1985, c S-19, §§ 5(3), 6(1), 17(1), 30 (laying out some of the rules and penalties for using Statistics Canada's data).
[205] *See* Digital Economy Act 2017, c. 30, pt. 5, ch. 5 (Eng.), http://www.legislation.gov.uk/ukpga/2017/30/contents/enacted (requiring, prior to disclosure, that information

Electronic copy available at: https://ssrn.com/abstract=3794780

proposed a system of "safe sharing" whereby "a party holding raw data with [personal data] could allow another party to analyze the data in select ways, while blocking them from viewing the raw data itself."[206] Their approach illustrates how epistemic public goods can be created without compromising privacy values.

Under-utilization has not been a focus of scholarship on personal data economies. But cities have started to grasp data as a "public good."[207] So-called smart cities that "collect and utilize an extensive range of personal and sensitive data" can be treated as "data stewards" responsible for wise use of that asset.[208] Some scholars have argued for cities to be subject to "fiduciary-like responsibilities to consider the ethical and privacy impacts of particular data activities and to act with the best interests of individuals and society in mind."[209]

It is not at all clear whether society suffers more from the misuse of personal data or the failure to use personal data for public-good creation. Such failures are likely to be regressive. Wealthier citizens are more likely to opt out of poorly performing public services, or seek alternative provision of public good.[210] Populations that are economically or socially marginal, in contrast, will not benefit from personal data's absent public interventions.[211] The failure to leverage the public-good potential of personal data economies, in short, will have the dynamic effect of exacerbating existing economic disparities, in ways that resonate with the critique from structural economic inequalities.

H. The Costs of Personal Data Economies Recapitulated

Platform economies, data brokers, and sensing nets have been subject to a barrage of criticism. Some of it rests on controversial metaphysical premises. Other elements have an uncertain relationship to the empirical evidence. Yet the powerful and persuasive critiques based on privacy, exploitation, and inequality are hard to wave away. Concerns about democratic backsliding, state domination, the under-supply of public goods, in my view, all have some force too. The focus of these critiques has been platform economies. Data brokers have successfully kept a lower public profile—but there is no reason to think that these normative concerns do not bite on their doings. Legal scholars are only now starting to explore the sensing net. Its capacious

identifying "particular individual[s]" be processed such that "it is not reasonably likely that the person's identity will be deduced" even when "taken together with other information").

[206] Lisa M. Austin & David Lie, *Safe Sharing Sites*, 94 N.Y.U. L. REV. 581, 600 (2019).

[207] *See, e.g.,* Robert M. George, *Data for the Public Good: Challenges and Barriers in the Context of Cities*, *in* PRIVACY, BIG DATA, AND THE PUBLIC GOOD 153, 153 (Julia Lane et al., eds. 2014) ("Comprehensive, high-quality multidimensional data has the potential to improve the services cities provide ….").

[208] Ira S. Rubinstein & Bilyana Petkova, *Governing Privacy in the Datafied City*, 47 FORDHAM URB. L.J. 755, 791 (2020)

[209] Kelsey Finch & Omer Tene, *Smart Cities: Privacy, Transparency and Community*, *in* THE CAMBRIDGE HANDBOOK OF CONSUMER PRIVACY 126-27 (Evan Selinger et al. eds., 2018).

[210] In the Covid-19 pandemic, for example, mortality rates for racial minorities has been higher than for white populations. Tiffany N. Ford, Sarah Reber, and Richard V. Reeves, *Race gaps in COVID-19 deaths are even bigger than they appear*, BROOKINGS INSTITUTE (June 16, 2020), https://www.brookings.edu/blog/up-front/2020/06/16/race-gaps-in-covid-19-deaths-are-even-bigger-than-they-appear/ ("In every age category, Black people are dying from COVID at roughly the same rate as white people more than a decade older.").

[211] *See, e.g.,* George, *supra* note 207, at 162-64 (discussing the use of spatial and personal databases to improve foster care, public housing, and poverty-eradication efforts in Chicago).

Electronic copy available at: https://ssrn.com/abstract=3794780

ability to acquire behavioral data suggests that normative objections lodged against platform economies will also resonate there.

There is no single normative value at issue across these arguments. To the contrary, they rest on different theories of the relation between the self and personal data economies. Critiques sounding in autonomy (and to some extent privacy) posit a static, authentic self undermined by the capture and commodification of personal data. In contrast, exploitation and economic inequality argument take a dynamic view of the self as subject to economic change. Similarly, the argument from democratic backsliding assumes that data-derived interventions have a causal effect polarizing the public. Platform economies, that is, make people themselves worse.

Just as they start from divergent premises, these critiques invite different, contradictory cures. The autonomy and privacy critiques could be met if platform economies did not transform personal data into an asset.[212] This would likely entail front-end prices for search and networking services. But a switch from payment in data to payment in cash is likely to have a regressive effect. Wealthier users are more likely to be able to afford cash payments. Data acquired from wealthy users might also be more 'valuable' (for example, for targeted advertising purposes) than less wealthy users' data. This would exacerbate ambient economic inequalities.

Nevertheless, it is possible to rank and organize the critiques to give them a measure of coherence. As a rough generalization, several arise from disparities of information and influence. Concerns about exploitation, inequality, and the under-supply of public goods pivot on the regressive effects of personal data economies. Retail privacy worries about improper sharing, data breaches, and unanticipated affordances also have a distributive element: They are all instances in which platform economies or sensing nets extract a larger informational surplus than consumers reasonably anticipate. Concerns about democratic backsliding and state repression also hinge on the emergence or reinforcement of political hierarchies. In contrast to these distributive concerns, objections based on efficiency do not loom large in critiques of personal data economies.[213]

This prominence of distributive concerns follows from the basic architecture of personal data economies. Platform economies, data brokers, and sensing nets all have a one-to-many logic on the consumer-facing side. Collective-action costs make it difficult for users to monitor or respond to objectionable practices. Many practices, moreover, occur in a different business-to-business market, and turn on quite technical details hard for consumers to comprehend. In contrast, personal data economies provide ample opportunities for improving efficiency. Firms have strong incentives to seek them out. Welfarist concerns can arise when platform economies such as Facebook or Google engage in monopolist practices.[214] But market concentration not only raises

[212] *See* ZUBOFF, *supra* note 112, at 67-70 (presenting a positive account of early platform economies that had)
[213] *But cf.* Bar-Gill, *supra* note 142, at 221 (identifying circumstances in which personalized advertising can generate inefficient outcomes).
[214] Dina Srinivasan, *Why Google Dominates Advertising Markets Competition Policy Should Lean on the Principles of Financial Market Regulation*, 24 STAN. TECH. L. REV. 55, 63 (2020) ("However, the majority of advertising revenue and growth has gone to large firms like Google and Facebook that both sell their own ad space and simultaneously run an electronic marketplace.").

Electronic copy available at: https://ssrn.com/abstract=3794780

concerns of "distorted growth and high trading costs."[215] It also might dissolve the constraints that competition might otherwise impose on exploitative or privacy-invading practices.[216]

III. Governance Regimes for Personal Data

In the origin fable of Harold Demsetz's famous property theory, the Montagne tribe living in what is now Quebec developed a system of individual property rights to hunt for beaver pelts in response to a spike in demand and overhunting.[217] Like the Montagne, users and regulators of personal data economies today confront a familiar resource—information rather than furs—but unfamiliar technologies of production and use. Like the Montagne, they face a question of what governance regime—including what sort of property rights—best encourages desirable resource allocations, while limiting undesirable spillover costs.

The concerns aired in Part II have inspired a broad range of policy proposals, ranging from government control of platform moderation[218] to tech worker unionization.[219] Few change the basic structure of personal data economies. Rather, they fiddle at the margin. This Part zooms in upon the boldest alternative governance proposals aimed at mitigating personal data economies' costs. The leading proposal, tracking Demsetz, involves the creation of individual ownership rights to data. A second affixes a fiduciary duty to platform economies. A third idea entails new structural antitrust remedies. This survey of extant structural proposals illuminates a gap: The array of governance tools commonly considered for personal data economies falls short of addressing all relevant policy concerns, particularly the mitigation of structural economic inequality and the supply of positive public goods. Some other governance intervention hence seem worth considering.

Proposals for a new governance regime for personal data rarely linger on the question of what legal regime applies now. But the status of personal data as property now is shot through with ambiguity. It is said that "the law does not presently recognize a property right in a particular piece of data."[220] Yet the tort of conversion is available when data is misappropriated.[221] In practice, platforms assert a "de facto if not de jure" proprietary interest in both data and algorithms.[222] Trade secrets also operate as "de facto property arrangements that affect large numbers of people."[223] Platforms or other participants in personal data economies are likely under

---

[215] *Id*. at 65.
[216] Khan, *Separations*, *supra* note 55, at 1004-05.
[217] Harold Demsetz, *Toward a Theory of Property Rights*, 57 AM. ECON. REV. PAPERS & PROC. 347, 356 (1967).
[218] Kyle Langvardt, *Regulating Online Content Moderation*, 106 GEO. L.J. 1353, 1363 (2018) (exploring this possibility).
[219] Kate Conger, *Hundreds of Google Employees Unionize, Culminating Years of Activism*, N.Y. TIMES (Jan. 5, 2021), https://www.nytimes.com/2021/01/04/technology/google-employees-union.html.
[220] Michael C. Pollack, *Taking Data*, 86 U. CHI. L. REV. 77, 106 (2019).
[221] *Thyroff v Nationwide Mutual Insurance Co*, 864 N.E.2d 1272, 1278 (NY 2007) (holding that a claim for conversion of electronic data is cognizable because "it generally is not the physical nature of a document that determines its worth, it is the information memorialized in the document that has intrinsic value").
[222] COHEN, *supra* note 47, at 44; *see also* Julie E. Cohen, *Law for the Platform Economy*, 51 U.C. DAVIS L. REV. 133, 156 (2017). For a similar view, see Birch et al., *supra* note 46, at 480 ("[L]egally speaking personal data are not assigned to individuals; instead they are treated as belonging to the entity—usually a private business—that collects and processes them … mostly without contractual consent ….").
[223] COHEN, *supra* note 47, at 45.

Electronic copy available at: https://ssrn.com/abstract=3794780

no general obligation to share data.[224] The pervasive reliance on contract arrangements is an implicit concession by firms with the most acute fiscal stake in the matter that in rem, property rights are unavailable.

A. Personal Property in Data

The idea of an alienable individual property interest in personal data is neither new nor uncontroversial. Although it has experienced a recent revival, it dates back to the emergence of the internet. Different iterations of the idea have been crafted to appeal to different constituencies. Yet it has never been broadly accepted.[225] Nor do standard justifications for a regime of individual rights squarely apply in the data context. The persisting allure of an individual rights framework, therefore, may reflect a lingering assumption that the standard form of property rights used for real and personal property can be extended to a new and different context—even though its justifications do not quite attach.

Proposals to treat personal data as a species of discrete property date back at least to the 1990s. In 1996, economist Harry Laudon proposed a highly regulated "National Information Market" to allow the sale and purchase of "personal information."[226] To enter this market, individual users and consumers would sell their data to local banks, which would bundle and sell data on national exchanges.[227] Responses within the legal academy to Laudon's idea were frosty. Some argued that the institutional infrastructure of a new personal data economy would be too expensive, that individuals would not be able to value accurate their data, and that propertization would distort "normative understandings about acceptable and unacceptable uses of personal data."[228] Others were concerned that "more, not less, trade" would entail "producing less, not more, privacy."[229] Markets were perceived as inconsistent with a normatively attractive level of privacy derived independently of users' expressed preferences. To these concerns might be added worries

---

[224] Two federal courts have held that competitors of a digital platform are entitled to interim injunctions that guaranteed continued access to consumer information based on state and federal law theories. *See hiQ Labs, Inc. v. LinkedIn Corp.*, 938 F.3d 985, 996-1004 (9th Cir. 2019) (issuing a preliminary injunction under a state tortious interference in contract claim and a claim under the federal Computer Fraud and Abuse Act); *PeopleBrowsr, Inc. v. Twitter, Inc.*, No. C-12-6120 EMC., 2013 WL 843032 (N.D. Cal. 2013). *But see* Stackla, Inc. v. Facebook Inc., No. 19-CV-05849-PJH, 2019 WL 4738288, at *6 (N.D. Cal. Sept. 27, 2019*)* (denying an injunction for plaintiff access to Facebook user data, as such a remedy "would compel Facebook to permit a suspected abuser of its platform and its users' privacy to continue to access its platform and users' data ... issuing an injunction at this stage could handicap Facebook's ability to decisively police its social-media platforms in the first instance").

[225] Beauvisage & Meller, *supra* note 45, at 3 (noting the "repeated and unsuccessful attempts to create a consumer-to-business … market for personal data").

[226] Kenneth C. Laudon, *Markets and Privacy*, COMM. ACM, Sept. 1996, at 92.

[227] *Id.* For support of this position from within the legal academy, see LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE 142-63 (1999).

[228] Pamela Samuelson, *Privacy As Intellectual Property?,* 52 STAN. L. REV. 1125, 1138–46 (2000); Jessica Litman, *Information Privacy/information Property*, 52 STAN. L. REV. 1283, 1303 (2000) ("The weaknesses of the property model are, first, that it encourages transactions in data that most of us would prefer be discouraged and, second, that its reliance on alienability and easy waiver tend to vest control over personal data in the data miner rather than the data's subject.").

[229] Cohen, *Examined Lives*, *supra* note 194, at 1391. Conversely, there was a worry about "circumstances in which society may want to make use of information that the individual does not want to release." Mark A. Lemley, *Private Property*, 52 STAN. L. REV. 1545, 1553 (2000). Lemley advocated for "government regulation of the behavior of data collectors" in lieu of a property rights regimes. *Id.* at 1554.

Electronic copy available at: https://ssrn.com/abstract=3794780

about the intermediaries that Laudon posited. The closest parallel to these entities now are data brokers. These, however, have been very successful in resisting regulation and oversight.[230] By baldly positing that banks will behave differently without supporting evidence, Laudon's proposed governance regime risked recreating existing dynamics.

Several efforts to realize Laudon's vision launched and then foundered in the early 2000s. Companies such as Personal, Datacoup, Handshake, and Yes Profile all stumbled because "the capture and sale of web users' traces was already a widespread practice," and none were able to press for legal changes that would have forced (say) data brokers to purchase that information from individuals.[231] Not all, though, were dissuaded. In one of the most prominent interventions of the era, Paul Schwartz proposed "use-transferability restrictions in conjunction with an opt-in default."[232] Schwartz suggested that personal data could be sold but that limits on its use and its transferability would "follow the personal information through downstream transfers and thus limit the potential third-party interest in it."[233] In effect, he proposed a regime of servitudes that "run" with data and so "pass automatically to successive owners."[234] He also argued that "given the right information and incentives," consumers would opt in to such a regime, and that technology would reduce the transaction costs of such contracting.[235] Rather than the centralized market architecture proposed by Laudon, Schwartz thus offered a decentralized array of venues, coupled to private rights of action against data breaches.[236]

Schwartz's model focused solely on "personal privacy," and did not account for the other normative critiques.[237] It placed heavy epistemic and cognitive demands on consumers. For every platform or sensing net they employ, they must provide detailed schedules of preferences. At the time Schwartz wrote, the number of such choices would have been manageable. But as the number of apps and tools that harvest personal data has risen, it has become less tractable. Today, Schwartz's proposal would likely have regressive effects given the scarcity of time and decision-making support in low-income communities. Just as common-law lawyers resisted the system of servitudes in land because of the risk of excessively complex encumbrances, moreover, so participants in personal data economies might resist Schwarz's proposal because of high transaction costs. Among the "features … especially likely to make servitudes problematic" are "the remote relationship between the burdened and benefited parties, the durability and ubiquity of the restrictions imposed, the fragmentation of rights to control use of a single resource, [and] the potential lack of salience to purchasers."[238] The "friction and disruption" from Schwartz's proposal would be greater than he allows.[239] In this light, it seems unlikely that Schwartz's

---

[230] *See supra* text accompanying notes – to --.
[231] Beauvisage and Meller, *supra* note 45, at 7.
[232] Schwartz, *Property*, *supra* note 99, at 2094.
[233] *Id.* at 2097.
[234] RESTATEMENT (THIRD) OF PROP.: SERVITUDES § 1.1 (2000).
[235] Schwartz, *Property*, *supra* note 45, at 2105.
[236] *Id.* at 2111-12.
[237] *Id.* at 2126; *see supra* Part II.
[238] Molly Shaffer Van Houweling, *The New Servitudes*, 96 GEO. L.J. 885, 890 (2008).
[239] Cameron F. Kerry and John B. Morris, Jr., *Why data ownership is the wrong approach to protecting privacy*, Brookings Institute (June 26, 2019), https://www.brookings.edu/blog/techtank/2019/06/26/why-data-ownership-is-the-wrong-approach-to-protecting-privacy/.

Electronic copy available at: https://ssrn.com/abstract=3794780

proposal would be entertained now—except by those who value privacy and autonomy to the extent that they wish to preclude most or all of the personal data economy.

A second wave of proposals for individual property rights broke in the late 2000s. In a popular book, technologist Jaron Lanier argued that users should have compensation for "information taken from them."[240] In 2009, prominent computational scientist Alex Pentland presented a paper at the World Economic Forum in Davos calling for individual property rights in personal data. Pentland argued for transposing "three basic tents of ownership" to the data context: "possession, use, and disposal."[241] He also called for "the combination of massive amounts of anonymous data to promote the Common Good."[242] Oddly, Pentland assumed a clear division between personal (identified) data and anonymous (non-identifiable) data. But by the time he wrote, research on deanonymization had demonstrated that any such crisp distinction was already implausible.[243]

More recently, the Laudon-Pentland proposal has been taken up and expanded in several high-profile treatments. In 2019, Senator John Kennedy (R-LA) introduced the "Own Your Own Data Act of 2019," which stipulated that "each individual owns and has an exclusive property right in the data that individual generates on the internet."[244] At three pages, the Act is long on sentiment but short on implementing detail. Harkening back to the 2014 World Economic Forum, the musician Will.I.Am took to the pages of *The Economist* to argue that control of data was "a central human value. The data itself should be treated like property and people should be fairly compensated for it."[245] In October 2017, the European Commission proposed a data producer's right for nonpersonal, anonymized machine-generated data.[246] In 2021, internet originator Tim Berners-Lee proposed the use of "pods," or individualized data safes in the cloud, to house personal data about "websites visited, credit card purchases, workout routines, [and] music streamed."[247] A pod system is indeed being tested with dementia patients in the United Kingdom.[248]

Most recently, Eric Posner and E. Glen Weyl identify a concern about structural economic inequality. On their account, the creation of personal form is a "form of labor" meriting

---

[240] JARON LANIER, WHO OWNS THE FUTURE? 50-51 (2013).
[241] Alex Pentland, *Reality Mining of Mobile Communications: Toward a new Deal on Data* 79 (2009), http://dml.cs.byu.edu/~cgc/docs/atdm/Readings/RealityMining2.pdf. On the context in which Pentland presented his paper, see Morovoz, *supra* note 195, at 33-34.
[242] Pentland, *supra* note 241, at 79.
[243] *See, e.g.,* Arvind Narayanan & Vitaly Shmatikov, *Robust De-Anonymization of Large Sparse Datasets*, *in* PROC. OF THE 2008 IEEE SYMP. ON SECURITY AND PRIVACY 111.
[244] S.806, Own Your Own Data Act of 2019 (Mar. 14, 2019), https://www.congress.gov/bill/116th-congress/senate-bill/806.
[245] Will.I.Am, *We need to own our data as a human right—and be compensated for it*, THE ECONOMIST (Jan. 21, 2019), https://www.economist.com/open-future/2019/01/21/we-need-to-own-our-data-as-a-human-right-and-be-compensated-for-it
[246] *Commission Communication on "Building a European Data Economy,"* at 13, COM (2017) 9 final (Oct. 1, 2017); *see also* P. Bernt Hugenholtz, *Against 'Data Property,' in* 3 KRITIKA: ESSAYS ON INTELLECTUAL PROPERTY 48, 51-52 (Harms Ullrich et al. eds., 2018) (explaining the European proposal as a response "demands of the automotive industry").
[247] Steve Lohr, *He Created the Web. Now He's Out to Remake the Digital World*, N.Y. TIMES (Jan. 10, 2021), https://www.nytimes.com/2021/01/10/technology/tim-berners-lee-privacy-internet.html.
[248] *Id.*

Electronic copy available at: https://ssrn.com/abstract=3794780

compensation.[249] They frame a personal property interest in personal data as a mechanism to alleviate structural economic inequalities exacerbated by automation.[250] Conceding that today most users would earn "only a few hundred dollars a year" from the sale of their data, they argue that once jobs have been destroyed by automation, "people will have plenty of time to supply that data."[251] Doing so, they suggest will "make them feel like more useful members of society."[252] Like Laudon, they recognize that a market in personal data would entail both new regulatory infrastructure[253] and intermediary institutions to bundle together individuals' contributions—called "data labor unions" or "data vaults" depending on their audience.[254]

The Posner-Weyl proposal invites some of the same objections as earlier iterations of individual property proposals, and raises a few new concerns.[255] Their writings are ambiguous about what kinds of data fall within their proposal. In some moments, they seem to suggest that their proposal applies to data created today as a byproduct of interactions with platform economies. At other times, they look forward to a future in which "people have time to supply" data by concerted labor rather than as a byproduct of activities aimed at other aims.[256] Moreover, unlike the earlier generation of 'data as property' legal scholars, Posner and Weyl are at best indifferent to privacy goals. Their lead example of valuable labor involves a person disclosing personal, even intimate, details about friends' relationships to a social network seeking to better understand its own data.[257]

Posner and Weyl's proposal is unlikely to mitigate structural economic inequalities. To see this, we should distinguish between the market as presently structured, and the market as it might operate in the future. At present, "personally-identified data is not scarce,"[258] and hence will rarely produce significant value. To the extent that some data has a marginally greater value than other data, it will likely be because it has been produced by a wealthier consumer, who is (by dint of their wealth) a more attractive object of targeted advertising. Implemented today, therefore, a regime of the kind that Posner and Weyl propose would likely have regressive effects. To their

---

[249] Eric A. Posner and E. Glen Weyl, *Want Our Personal Data? Pay for It*, WALL ST. J. (Apr. 20, 2018), https://www.wsj.com/articles/want-our-personal-data-pay-for-it-1524237577 [hereinafter, "Posner and Weyl, Pay for it"]. The argument for markets in personal data is developed in ERIC A. POSNER AND E. GLEN WEYL, RADICAL MARKETS: UPROOTING CAPITALISM AND DEMOCRACY FOR A JUST SOCIETY 243-49 (2018) [hereinafter "POSNER AND WEYL, RADICAL MARKETS"].

[250] Posner and Weyl, *Pay for it*, *supra* note 249.

[251] *Id.*; *see also* Eduardo Porter, *Your Data Is Crucial to a Robotic Age. Shouldn't You Be Paid for It?*, N.Y. TIMES (Mar 6, 2018), https://www.nytimes.com/2018/03/06/business/economy/user-data-pay.html (broadly endorsing this proposal, but noting "transition" costs).

[252] POSNER AND WEYL, RADICAL MARKETS, *supra* note 249, at 248.

[253] *Id.* at 245.

[254] Posner and Weyl, *Pay for it*, *supra* note 249 ("data vaults"); POSNER AND WEYL, RADICAL MARKETS, *supra* note 249, at 241-43 ("data labor union").

[255] Peter Yu, for example, has noted the risk that any individualized right may lead to an inefficient anticommons. Peter K. Yu, *Data Producer's Right and the Protection of Machine-Generated Data*, 93 TUL. L. REV. 859, 889 (2019).

[256] Posner and Weyl, *Pay for it*, *supra* note 249; *see also* POSNER AND WEYL, RADICAL MARKETS, *supra* note 249, at 223 ("[D]ata as labor may offer important supplemental earning opportunity and sense of social contribution to citizens affected by rising inequality.").

[257] POSNER AND WEYL, RADICAL MARKETS, *supra* note 249, at 205-07.

[258] Cohen, *Examined Lives*, *supra* note 194, at 1387.

Electronic copy available at: https://ssrn.com/abstract=3794780

credit, they acknowledge this.[259] But their solution of hoping for a 'broader range of niches" for data production seems implausible.[260] It does nothing to mitigate a future labor market characterized by under-employment, "wage stagnation and worsening conditions."[261] To the contrary, it is of a piece with that dystopia.

No less unpersuasive is their hypothesized future in which 'data as labor' produces "supplemental income."[262] It is difficult to see this supplement could meaningful change income distributions. Rather, the 'data as labor' proposal would increase demand in the low-wage, low-skill segment of the labor market only marginally. And it does nothing to make up for the middle-income positions actually lost to automation. Posner and Weyl imagine that demand for data will grow over time because of its growing utility for training machine-learning tools.[263] They pay little attention, to the exponential growth of data, as platforms and sensing nets grow.[264] Against their optimism, therefore, it is possible to imagine a future in which data about so many forms of behavior is so cheap to acquire and store that the marginal benefit of anything consciously produced via intentional labor is vanishingly small. Predictions of progressive effects from a 'data as labor' economy in the future, therefore, seem fragile.

B. Data Governance through Fiduciary Duties

A second popular proposal to mitigate the costs of personal data markets uses not property law but fiduciary principles. The most prominent iteration, tendered by Jack Balkin and Jonathan Zittrain in 2016, focused on platform economies such as Google, Facebook, and Uber as "information fiduciaries."[265] Elaborating the idea in subsequent work, Balkin has argued that "certain types of online service providers [should] take on fiduciary responsibilities" towards their users because they hold themselves out as trustworthy recipients of personal data.[266] This fiduciary duty was initially characterized as a way "to encourage creativity without facilitating betrayal" of consumers' trust.[267] In practice, this would leave much of platform economies' business model intact.[268] In a more recent writing, Balkin has clarified the content of fiduciary "duties of care,

[259] POSNER AND WEYL, RADICAL MARKETS, *supra* note 249, at 247.
[260] *Id.*
[261] Aaron Benanav, *Automation and the Future of Work—2*, 120 NEW LEFT REV. 117, 123 (2019).
[262] POSNER AND WEYL, RADICAL MARKETS, *supra* note 249, at 223.
[263] Indeed, they posit increasing return as the supply of data rises. POSNER AND WEYL, RADICAL MARKETS, *supra* note 249, at 227.
[264] Consider here trend in recent data management is the creation of "Data Management Platforms" that can "merg[e] heterogeneous data sources into a single place" for a firm, and thereby "drive business actions toward consumers in the wild." Beauvisage & Meller, *supra* note 45, at 15. The ensuing profusion of "data lakes" indexes the futility of insisting on the meaningful financial value of single items of data.
[265] Jack M. Balkin & Jonathan Zittrain, *A Grand Bargain to Make Tech Companies Trustworthy*, ATLANTIC (Oct. 3, 2016), https://www.theatlantic.com/technology/archive/2016/10/information-fiduciary/502346/; *see also* Jack M. Balkin, Lecture, *Information Fiduciaries and the First Amendment*, 49 U.C. DAVIS L. REV. 1183 (2016) [hereinafter Balkin, *Information Fiduciaries*].
[266] Balkin, *Information Fiduciaries*, *supra* note 265, at 1221.
[267] *Id.* at 1224.
[268] *Id.* at 1227 ("Because personal data is a key source of wealth in the digital economy, information fiduciaries should be able to monetize some uses of personal data, and our reasonable expectations of trust must factor that expectation into account. What information fiduciaries may not do is use the data in unexpected ways to the disadvantage of people who use their services or in ways that violate some other important social norm.").

Electronic copy available at: https://ssrn.com/abstract=3794780

confidentiality, and loyalty."[269] Echoing Schwartz, he has explained that these would "run with the data."[270] He has also suggested that at least certain forms of behavioral advertising would be impermissible, so one should "not take existing business models as given."[271] The fiduciary model, that is, can be interpreted in both narrow and broad ways.

Balkin's model has been critiqued for failing to account for existing market structures. Platform economies are not merely passive recipients of data; they also leverage their monopoly status to promote the "loss of privacy and control" over personal data.[272] Lina Khan and David Pozen's argue that since Balkin's "user-centric" model fails to address structural problems of concentration and market power, it "is bound to be at best highly incomplete."[273] They see no utility in reforms without a radical transformation of market structures. Their objection to the information fiduciary model turns on the empirical question whether a revenue model based on targeted advertising can be cabined in morally acceptable ways.[274]

Even aside from Khan and Pozen's criticisms, the information fiduciary model provides limited traction for managing the panoply of harms canvassed in Part III. Its advocates frame it as an intervention against platform economies, not against data brokers and sensing nets. Indeed, the fiduciary obligation attaches to an entity "because of their relationship with another."[275] But for much personal data entering commercial circulation beyond the platform, there is no such relationship. Consider sensors that capture images and speech in public and private places, monitors that track activity in physical locations, and even cookies that track activity on third-party websites on Facebook's behalf.[276] All lack the dyadic relationship Balkin posits as necessary to a fiduciary duty. Worse, inferences about specific individuals can be drawn not only from data gathered from them, but also from third parties.[277] To encompass the larger personal data economy, Balkin's model would need to be unmoored from its starting analogy to the fiduciary obligations of doctors and lawyers. It would have to become a more free-floating, miasmatic duty of trustworthiness and constraint. It would become, in short, a coat cut from a different doctrinal cloth altogether.

Even in that reimagined form, moreover, a fiduciary principle would remain focused on the "user-centric"[278] concern of privacy. As we have seen, a governance regime that mitigates

---

[269] Jack Balkin, *The Fiduciary Model of Privacy*, 132 HARV. L. REV F. 11, 14 (2020) [hereinafter "Balkin, *Fiduciary Model*"].
[270] *Id.* (citation and quotation marks omitted).
[271] *Id.* at 29. The concession came in response to a critical account in Lina M. Khan & David E. Pozen, *A Skeptical View of Information Fiduciaries*, 133 HARV. L. REV. 497 (2019).
[272] Khan and Pozen, *supra* note 271, at 517-18. Confusingly, Khan and Pozen argue both that the information fiduciary principle largely tracks existing state law, *id.* at 521-24, and also that it would overwhelm courts' dockets within new cases, *id.* at 524. How can both of these things be true?
[273] *Id.* at 528; *id.* at 534 (criticizing the information fiduciary framework because it "characterizes Facebooks, Google, Twitter, and other online platforms as fundamentally trustworthy actors who put their users' interests first").
[274] *Compare id.* at 513 & n.74 (no), *with* Balkin, *Fiduciary Model*, *supra* note 269, at 28-29 (yes).
[275] Balkin, *Information Fiduciaries*, supra note 265, at 1209.
[276] *See supra* text accompanying notes – to --; Srinivasan, *supra* note 60, at 42–43.
[277] *See* Michele Loi, *The digital phenotype: A philosophical and ethical exploration*, 32 PHIL. & TECH. 155, 161 (2019) (noting "the possibility of discrimination harm due to generalizable knowledge, for persons who are not identified by the data, which the data protection framework is not equipped to solve").
[278] Khan and Pozen, *supra* note 271, at 528.

Electronic copy available at: https://ssrn.com/abstract=3794780

privacy concerns is not one that will equally address objections from structural economic inequalities or the underproduction of public goods.[279] At best, therefore, the fiduciary intervention proposed by Balkin and others in respect to platform economies is a partial response to the problems of personal data economies.

A variant on the idea of fiduciary obligations is worth mentioning because it is a step toward the public trust model explored in Part IV.[280] Michele Loi, Paul-Olivier Dehaye, and Ernst Hafen have proposed the creation of "personal data platform cooperatives" as vehicles through which individuals could make "collective choices" about "what data to share and with whom," and how the surplus from such exploitation should be used.[281] Sylvie Delacroix and Neil Lawrence have similarly proposed a "bottom up" trust mechanism in which data subjects are both settlers and beneficiaries of trusts that manages data in their name.[282] In a related vein, the Alphabet Subsidiary Sidewalk Labs proposed the creation of a "data trust" to govern data gathered as part of an ambitious 'smart city' initiative in Toronto.[283] The Information and Privacy Commissioner of Ontario, however, raised concerns about the proposed trust's "lack of independent oversight."[284] In effect, the Commissioner worried that the trust would exercise "exceptional regulatory powers" without itself being a public entity amenable to democratic control.[285] The Sidewalk project was canceled in March 2020, so the data trust idea was never implemented.[286] Hence, it is unclear whether a private "data trust" model that is distinct and difficult from a fiduciary model, or indeed, the intermediaries first imagined in 1996 by Harry Laudon, is in fact feasible.

C. Structural Antitrust Remedies

Finally, a wave of scholarship has challenged the concentration of market power in a small number of platforms such as Facebook, Google, and Amazon. That literature has elicited proposals aimed at fundamentally altering market structure for platform economies. Lina Khan, for example, has argued for structural remedies that "proscribe certain organizational structures," such that "platform activity and commercial activity [would] be undertaken through separate corporations with distinct ownership and management."[287] Alphabet, on this view, might be split into firms that supply search and firms that produce content.[288] Sanjukta Paul has criticized antitrust doctrine's

---

[279] *See text* accompanying notes – to --.
[280] For the linkage between fiduciary proposals and data trusts, see Anna Artyushina, *Is civic data governance the key to democratic smart cities? The role of the urban data trust in Sidewalk Toronto*, 55 TELEMATICS & INFORMATICS 101456, at *5 (2020).
[281] Michele Loi, Paul-Olivier Dehaye, and Ernst Hafen, *Towards Rawlsian 'property-owning democracy' through personal data platform cooperatives*, -- CRIT. REV. INT'L SOC. & POL. PHIL, 1, 8 (2020); *see also* Pistor *supra* note 146, at 118-22 (positing a similar idea).
[282] Sylvie Delacroix and Neil D. Lawrence, *Bottom-up data Trusts: disturbing the 'one size fits all' approach to data governance,* 9 INT'L DATA PRIVACY L. 236, 240 (2019).
[283] Sean McDonald, *Reclaiming Data Trusts*, Center for International Governance Innovation (Mar. 5, 2019), .
[284] Letter from Brian Beamish, Information and Privacy Commissioner of Ontario, to Stephen Diamond, Chairman, Waterfront Toronto (Sept. 24, 2019), https://www.ipc.on.ca/wp-content/uploads/2019/09/2019-09-24-ltr-stephen-diamond-waterfront_toronto-residewalk-proposal.pdf
[285] Artyushina, *supra* note 280, at *10. Among the wider public, Sidewalk's proposal to establish "consent through signage" (i.e., treating a pedestrian's observation of a sign detailing what data was being collected for the trust as consent) roused particular ire. *Id.*
[286] *Id.* at *1-*2.
[287] Khan, *Separations*, *supra* note 55, at 980 & 1034.
[288] *Id*. at 1034.

Electronic copy available at: https://ssrn.com/abstract=3794780

assumption that the "business firm is the central locus of economic coordination," and suggested instead making more "space for more democratic, horizontal forms of economic coordination."[289] Sabeel Rahman has argued that "Google, Facebook, and Amazon" should be treated as "foundational utilities" and regulated as such.[290]

This "neo-Brandeisian" approach, unlike proposals to install individual rights or fiduciary duties, is laser-focused on "power," and in particular private power, rather than privacy or consumer welfare.[291] Concentrated private power is perceived as antithetical to "true democracy and liberty in our political sphere."[292] Large platforms, it is suggested, wield excessive influence not just by shaping flows of information and public debate,[293] but also more directly by wielding disproportionate influence as lobbyists in legislatures.[294] Interventions aimed at fracturing platforms therefore rest on the (controversial and unsupported) normative premise that a desirable understanding of democracy requires a specific diffusion of both private and public power.[295]

These structural remedies may well have substantial effects on privacy, innovation, and digital flows of speech and information. But there are reasons for resisting the temptation to view them as panaceas. First, as this approach's leading proponents candidly admit, it is far from clear that federal agencies and courts have the political will necessary to execute a neo-Brandesian program.[296] Second, arguments for new antitrust enforcement are focused almost exclusively upon platform economies. No argument has been advanced that data-broker and sensing-net parts of the economy are overly concentrated. Indeed, available evidence suggests that neither market is presently characterized by concentration of the kind that might trigger Sherman Act liability.[297] Since many of the harms canvassed in Part II are plausibly thought to arise as a consequence of data brokers and sensing nets, an antitrust-only approach will leave them unchanged. Third, competition may be consistent with exploitation, massive privacy losses, and economic inequality. E-commerce platforms in China, for example, are characterized by "fierce competition" with no dominant firm akin to Amazon, and yet abound with "group deals, social media, gaming, instant messaging, short-form video, and live-streaming celebrities."[298] Competition may thus be consistent with many of the critiques of personal data economies adumbrated in Part II.

---

[289] Sanjukta Paul, *Antitrust As Allocator of Coordination Rights*, 67 UCLA L. REV. 378, 430 (2020).
[290] K. Sabeel Rahman, *The New Utilities: Private Power, Social Infrastructure, and the Revival of the Public Utility Concept*, 39 CARDOZO L. REV. 1621, 1669–70 (2018).
[291] Lina Khan, *The New Brandeis Movement: America's Antimonopoly Debate*, 9 J. EUR. COMPETITION L. & PRAC. 131, 131-32 (2018); *see also* TIM WU, THE CURSE OF BIGNESS: ANTITRUST IN THE NEW GILDED AGE 9-11 (2018).
[292] Khan, *Separations*, *supra* note 55, at 1061.
[293] *Id.* at 1071.
[294] See Zephyr Teachout & Lina Khan, *Market Structure and Political Law: A Taxonomy of Power*, 9 DUKE J. CONST. L. & PUB. POL'Y 37, 72 (2014).
[295] *Id.*
[296] Khan, *Separations*, *supra* note 55, at 1065 (noting the "enfeebling of antitrust"). States, however, also have the "ability to bring suit under federal antitrust law and the … ability to enact and enforce their own state antitrust laws." Note, *Antitrust Federalism, Preemption, and Judge-Made Law*, 133 HARV. L. REV. 2557, 2560 (2020). The latter are "heterogeneous themselves." *Id.*
[297] *See supra* text accompanying notes – to --.
[298] *The future of global e-commerce*, THE ECONOMIST, Jan. 2, 2021, at 7 (noting also that Alibaba's market capitalization fell from 81% to 55% in 2021).

Electronic copy available at: https://ssrn.com/abstract=3794780

Finally, the effect of antitrust remedies upon the larger labor market will vary dramatically. Breaking up platforms with large workforces (e.g., Amazon and Uber) would mitigate monopsony effects, perhaps allowing wages to rise.[299] But most of the interventions proposed—such as separating search from content—are unlikely to have labor market effects since the ratio of market value to firm size has "exploded" for firms such as Apple, Google, and Facebook.[300] The historical record also suggests reasons for caution. Looking back at the major antitrust remedies issued against AT&T and IBM, it remains "hard to know exactly how much they shaped" product markets.[301] It seems wise to maintain a certain modesty about the ramifications of complex structural interventions by constrained government actors in the dynamic context of technological and social change.

If antitrust succeeds in breaking up large platforms, the effect on national politics is also indeterminate. The empirical literature on campaign finance contributions by individuals and corporations suggests a more complex story than the neo-Brandeisian account. For the past two decades, corporate expenditures have been less ideological and more focused on incumbents than the spending of individual executives of the same company.[302] It has aimed at influencing policy outcomes, not electoral outcomes. The largest tech firms did not build up "a large lobbying presence" in Washington, D.C. until after 2010, when they perceived a rising risk of regulation.[303] They have since focused on "immigration, net neutrality, rules governing advertising, and company-specific issues."[304] The idea that this represents a failure of democracy rests on the (strong) assumption that the policy outcomes that these firms aim to foster are themselves democratically problematic.[305] Because corporate spending flows to established rather than insurgent candidates, it may buffer partisan sectarianism in ways that have positive, system-level effects even if they impede certain regulatory reforms. A world in which Amazon, Facebook, and Google are spending less on political influence is not necessarily a world in which policy outcomes are more tightly linked to popular preferences. It might be that ramping up antitrust scrutiny on firms in the personal data economies in effect increases the relative influence of other corporate actors. Where Amazon steps back, for example, this might simply leave Walmart with more influence. It is premature to assume this outcome is more 'democratic.'

More intensive antitrust enforcement in the personal data economy may well have salutary effects on privacy, innovation, and perhaps certain kinds of democratic dysfunction. Given the federal dominance in antitrust policy-making, though, it is also a hostage to political fortune. For these reasons, while it may be part of an effective regulator's toolkit, it is also unlikely to be a cure-all for the problems canvassed in Part II.

---

[299] Suresh Naidu et. al., *Antitrust Remedies for Labor Market Power*, 132 HARV. L. REV. 536, 601 (2018).
[300] BOIX, *supra* note 158, at 186 (noting that in mid-2017, Apple had 116,000 employees, Google had 61,000, and Facebook had 23,000, despite all having market capitalizations in the hundreds of billions).
[301] Randal C. Picker, *The Arc of Monopoly*, 87 U. CHI. L. REV. 523, 548 (2020).
[302] Adam Bonica, *Avenues of influence: On the political expenditures of corporations and their directors and executives*, 18 BUS. & POL. 367, 367-78 (2016); Val Burris, *The two faces of capital: Corporations and individual capitalists as political actors*, 66 AM. SOC. REV. 361, 362 (2001) (finding that individual spending is aimed at changing the outcomes of elections, whereas but corporate spending aims to influence the regulatory decisions of incumbents regardless of party).
[303] THOMAS PHILIPPON, THE GREAT REVERSAL: HOW AMERICA GAVE UP ON FREE MARKETS 260-62 (2019).
[304] *Id.* at 260.
[305] Especially in respect to immigration, this seems hard to know for sure.

Electronic copy available at: https://ssrn.com/abstract=3794780

D. The Regulatory Gap in Personal Data Economies

The leading proposals for regulation of personal data economies are not likely to address all of the pressing normative concerns raised by their operation. Proposals to create individual entitlements to data, now more than two decades old, have never successfully addressed their considerable logistical and practical impediments. They lean on implausible assumptions about individuals' capacity for knowing and controlling their own data use. They are also unlikely to mitigate the regressive effects of commodifying personal data, and are prone to exacerbating privacy losses. The imposition of a fiduciary duty upon platform economies, in contrast, would reach only a portion of the firms trading personal data as an asset. Its impact would depend on the uncertain extent to which platforms' business models would have to change. While the scope of those obligations in familiar contexts is tolerably clear, how they would apply in new digital environments remains up in the air. Finally, structural antitrust remedies would accomplish important goals, including perhaps better privacy arrangements. But they too would be partial in scope, uncertain in effect, and largely targeted at welfarist ends orthogonal to the critiques lodged in Part II.

All these interventions, in short, rest on powerful justifications. Yet none takes up all structural economic effects of personal data economies. Nor does any mitigate the absence of positive public good production. There is therefore still room in the regulatory toolkit for something more when it comes to new personal data economies.

IV. The Public Trust in Data

The repertoire of structural responses to personal data's pathologies can be enriched, surprisingly, by reaching back to a common-law doctrine of property crafted in a nineteenth century society only passingly familiar with the perils and pleasures of commodifying information. The doctrine in question is called the "public trust." This Part explores the possibility of a "public trust in data" as an instrument that states and localities can deploy to address some of the harms arising from personal data economies.

To develop the case for a public trust in personal data, I begin by setting out the doctrine's common-law origins and American applications. I then explain why there is a close fit along several margins between earlier uses of the public trust and its proposed deployment in the digital age. Parallels exist between the kinds of resources subject to management under public trust doctrine in the past and personal data. They also run between past and present justifications for the creation of public trust. Finally, I suggest that the public trust form can be adapted to address the specific distributional and public-good related problems of personal data economies. That is, it is a way of durably bundling together solutions to several problems created by data economies. Through user fees, limits on permissible data deployments, and mandates to create public goods, a public trust can mitigate some of the power asymmetries and regressive effects of present data economies. At the same time, the trust can employ safeguards to foreclose governmental misuses of data, much as the rich personal data disclosed to social security and tax authorities is shielded from misuse. And through a trustee or public enforcement mechanism, all these constraints can simultaneously be given durable effect.

Electronic copy available at: https://ssrn.com/abstract=3794780

To be very clear, what follows is not intended as a comprehensive account of public data trusts. There are too many kinds of data, and too many local specifities, to allow for that. Rather, the aim of this Article is to provide a 'proof of concept' for a generally applicable legal idea. I hence close by offering general suggestions about how a public trust for data might be implemented by state or local governments in respect to sensing net data and extensions into platform economies.

A. The Public Trust Doctrine as a Resource for Governance

The core ambition of public trust doctrine is to facilitate long-term management of assets to benefit a broad cross-section of the public.[306] An asset in "public trust" is owed and managed by the state. Yet the public trust doctrine differs from the idea of "public land" owned by the state free of any supervening obligation.[307] Instead, the state has obligations of trusteeship "to protect the people's common heritage."[308] These constrain its ability to authorize wholesale private exploitation of a public-trust asset.[309] At the same time, certain controlled forms of commercial exploitation, including the alienation of some 'sticks' of the property bundle, may be allowed. Importantly, the balance struck between public and private uses of a trust asset must account for both the risk that a resource enjoyed by a broad public may be spoiled or exhausted through commercial exploitation, and also the possibility that state actors fail to meet their obligations to ensure public resources are properly husbanded and avail the public as a whole.[310] A further advantage of the public trust established through either legislation or state constitutional text is that it creates a platform for democratic deliberation and decision about an asset's mix of uses. Post hoc judicial review locks in democratic choices and guards against later defection.

Nevertheless, like any other doctrinal tool used to further important policy goals, the public trust doctrine is no panacea. On the one hand, its promise in the digital context derives from its combination of rules meant to preserve an asset for common enjoyment, with permissions for controlled commercial exploitation. Even as it would allow the continued commercial use of

---

[306] Ryan, *supra* note 4, at 161 ("The public trust's doctrinal infrastructure shows that it doesn't just protect the public nature of these common resources--it also assigns responsibility for their protection--specifically, to the government.").

[307] *See, e.g.,* 43 U.S.C. § 1702(e) (defining "public lands" as "any land and interest in land owned by the United States within the several States and administered by the Secretary of the Interior through the Bureau of Land Management"). *But see* 43 U.S.C. § 1702(e)(2) (excluding from definition of "public lands" those "lands held for the benefit of Indians, Aleuts, and Eskimos"). Under the Property Clause of Article IV of the Constitution, "Congress exercises the powers both of a proprietor and of a legislature over the public domain" without caveats or limitations. *Kleppe v. New Mexico*, 426 U.S. 529, 540 (1976).

[308] *Nat'l Audubon Soc'y v. Superior Court*, 33 Cal. 3d 419, 441, 658 P.2d 709, 724 (1983).

[309] *See* Sax, *supra* note 3, at 477 (enumerating specific limits in the use of a public trust asset).

[310] *Cf.* Sax, supra note 3, at 521 ("The 'public trust' doctrine has no life of its own and no intrinsic content. It is no more—and no less—than a name given by courts to their concerns about the democratic process."). Concerns have been raised about the "democratic deficit" created by judicial enforcement of the public trust doctrine. Thomas W. Merrill, *The Public Trust Doctrine: Some Jurisprudential Variations and Their Implications*, 38 U. HAW. L. REV. 261, 284 (2016) [hereinafter "Merrill, *Public Trust*"]. This argument rests on the fallacy of composition: It assumes a polity cannot be democratic unless all its consistent elements are democratically responsive. But this is false. No one thinks that elected leaders should have plenary power over police forces, election management, or the regulation of speech to ensure democratic responsiveness. The public trust doctrine is simply another way of promoting democracy by assuring some minimum level of security (here of assets) as against public misuse of private capture.

Electronic copy available at: https://ssrn.com/abstract=3794780

personal information, therefore, a public trust in data could be used for promoting privacy, dampening regressive distributional effects, enabling democracy, and eliciting the production of public goods. On the other hand, while a public trust in data can be easily established at a state or local level, a national-level trust would face practical and legal impediments. The public trust model, furthermore, might only fit certain kinds of data. Still, it would be premature to allow these barriers to preempt experimentation. Only by pursuing its possibilities through trial and error that the doctrine's potential might be realized.

The idea of a "public trust" in a common asset, under public ownership and control but subject to controlled public usage and limited private exploitation, has a long history. It can be traced back to Roman law.[311] Folded into English common-law, it has been part of American jurisprudence since the Republic began.[312] Its active use, however, traces back to the Progressive Era, when it was deployed as a prophylactic against legislative defalcation of resources enjoyed by the people in common. The ensuing history of public trust doctrine, summarized here, testifies to its adaptability and its capacity for handling shifting mixes of public and private usages.

Early American cases identified a public trust in resources such as oysters and fish. In the 1821 case of *Arnold v. Mundy*, for example, the New Jersey Supreme Court held that navigable waters were "common to all the people, and that each has a right to use them according to his pleasure," and so the public could not be excluded from oyster picking in the tidal Raritan River.[313] Two decades later, Chief Justice Roger Taney wrote for the U.S. Supreme Court—again in a case about the Raritan River—that "a public trust [existed] for the benefit of the whole community, to be freely used by all for navigation and fishery."[314] This public trust doctrine, the Court ruled four years later in a case about navigable tributaries, applied as background law to states other than the thirteen original colonies.[315] In these early cases, the Court rejected private claims to exclude the public from a resource, such as a navigable way or an oyster bed, while also underscoring a positive obligation on the state to maintain the resource's availability.

The leading American case on the public trust doctrine emphasizes the judiciary's obligation to protect a resource from the corrupt deployment of state power.[316] At issue in *Illinois*

---

[311] *See* INSTITUTES OF JUSTINIAN 78 (John T. Abdy & Bryan Walker trans., Cambridge Univ. Press 1876) (530 C.E.) ("By the law of nature these things are common to all men; air, running water, the sea, and consequently the shores of the sea.").

[312] *Free Fishers of Whitstable v. Gann,* (1865), 11 Eng. Rep. 1305 (HL) [1, 13-14] (holding, as a matter of the common law, that the sovereign held title to the bed of all tidal rivers, estuaries, and territorial seas for the benefit of the subjects); *see also* H. Bracton, On the Law and Customs of England 39–40 (S. Thorne trans. 1968) (describing the shores of the sea "common to all" and inalienable).

[313] *Arnold v. Mundy*, 6 N.J. 1, 42 (1821); *see also Carson v. Blazer*, 2 Binn. 475, 478 (Pa. 1810) (recognizing a "right to fisheries" in tidal waters that is "vested in the state and open to all").

[314] *Martin v. Waddell's Lessee*, 41 U.S. (16 Pet.) 367, 413 (1842).

[315] *Pollard v. Hagen*, 44 U.S. (3 How) 212, 215-26 (1845) ("A right to the shore between high and low water-mark is a sovereign right, not a proprietary one…. Why? Because rivers do not pass by grant, but as an attribute of sovereignty. The right passes in a peculiar manner; it is held in trust for every individual proprietor in the state or the United States, and requires a trustee of great dignity."); *id.* at 228-29 (explaining that common law doctrines of land applied to newly admitted states, such as Alabama); *see also* Ryan, *supra* note 4, at 153-55 (providing background to *Pollard*).

[316] *See* Sax, *supra* note 3, at 489 (describing the *Illinois Central* case as "[t]he most celebrated public trust case in American law"); *see also* Joseph D. Kearney & Thomas W. Merrill, *The Origins of the American Public Trust Doctrine: What Really Happened in* Illinois Central, 71 U. CHI. L. REV. 799, 802 (2004) (discussing the prominence

Electronic copy available at: https://ssrn.com/abstract=3794780

Central Railroad v. Illinois was the Lake Front Act, an 1869 state legislative measure granting the eponymous railroad a portion of the Chicago lakeshore and over one thousand acres of submerged land for a new depot.[317] Four years later, the state legislature revoked the grant. The railroad, of course, sued. It alleged (among other things) that the Act violated its "vested rights" in lakebed property. The resulting law-suit ended in a split judgment from the U.S. Supreme Court. Key, though, to the Court's ultimate holding was its conclusion that the 1869 transfer of submerged land had never been valid—and thus the railroad had not been deprived of any "vested right"—because of the public trust doctrine.[318]

Writing for the Court's majority, Justice Stephen Field held that the state might hold title in the land, but such title was "held in trust for the people of the State that they may enjoy the navigation of the waters, carry on commerce over them, and have liberty of fishing therein freed from the obstruction or interference of private parties"[319] He went on to explain that the state could neither "abdicate its trust over property in which the whole people are interested, like navigable waters and soils under them," nor "leave them entirely under the use and control of private parties."[320] Hence, the land could only be alienated if doing so promoted "the interests of the public" and had no "substantial impairment of the public interest in the lands and waters remaining."[321]

Illinois Central's holding reflects mistrust of concentrated private power, whether manifesting as an interest-group lobby or as the monopolistic owner of an asset that would otherwise avail a broad swathe of the public. How did it come to pass? To constitutional law scholars today, Justice Field is notorious as a "pioneer and prophet" of the laissez-faire interpretation of the Fourteenth Amendment's Due Process Clause.[322] But Field was also a Jacksonian Democrat willing to "summarily … divest a major American company of an exceedingly valuable property" to forestall "corruption and special privilege."[323] Whether or not the Lake Front Act in fact was induced through corrupt means,[324] the Court's ruling hinged on its perception of interest-group capture. *Illinois Central* thus embodies bilateral constraints, arising out of the public trust doctrine, upon the state as owner and manager, and also upon private firms and individuals as potential owners and users. Its inalienability rule reflects a commitment to preserving public ownership and hence democratic control. That is, it affirms democracy, just as

---

of the *Illinois Central* decision); Carol Rose, *The Comedy of the Commons: Custom, Commerce, and Inherently Public Property*, 53 U. Chi. L. Rev. 711, 737 (1986) [hereinafter "Rose, *Comedy of the Commons*"] (describing Illinois Central as the "most famous assertion of the public trust theory"); *see also Protect Our Parks, Inc. v. Chicago Park Dist.*, 971 F.3d 722, 729 (7th Cir. 2020) (Barrett, J.) (underscoring *Illinois Central* as central to public trust doctrine).

[317] 146 U.S. 387, 440-48 (1892).

[318] *Id.* at 453.

[319] *Id.* at 452.

[320] *Id.* at 453.

[321] *Id.* Now Justice Amy Coney Barrett's recent discussion of the public trust doctrine emphasized these constraints on alienation. *Protect Our Parks, Inc. v. Chicago Park Dist.*, 971 F.3d 722, 729 (7th Cir. 2020)

[322] Charles W. McCurdy, *Justice Field and the Jurisprudence of Government-Business Relations: Some Parameters of Laissez-Faire Constitutionalism, 1863-1897*, 61 J. AM. HIST. 970, 971 (1975)

[323] *Id.* at 994.

[324] Kearney & Merrill, *supra* note 316, at 893 ("[A]lthough the documentary record from 1869 cannot be said definitely to establish that the Illinois Central used corrupt means to facilitate the enactment of the Lake Front Act, it probably leans in that direction.").

Electronic copy available at: https://ssrn.com/abstract=3794780

it keeps a beady eye on its derailment. Finally, it is—much like the public utilities and broad reading of the Sherman Act other contemporary scholars are recovering—a Progressive Era effort to manage concentrated private power by endowing the state with power and still shackling the manner in which such power is exercised.[325] Hence, it is a rule concerned with power, and oriented toward democratic ends through a mix of public control and ex post judicial safeguards.[326]

Before the end of that century, the Supreme Court went on to endorse applications of the public trust doctrine to riverine resources[327] and wildlife.[328] The *Illinois Central* decision also prodded state courts to till independently the same jurisprudential field.[329] The Minnesota Supreme Court, for instance, extended the public trust to recreational uses of lakes, such as "sailing, rowing, fishing, skating, [and] taking water."[330] In a famous series of twentieth-century cases, the New Jersey Supreme Court identified a public trust in Atlantic beach access,[331] while Pennsylvania's high court found the "ambient air" to be subject to trust duties.[332] In the wake of Joseph Sax's influential scholarship recovering the doctrine in the 1970s,[333] lawyers in the nascent environmental movement deployed it aggressively across a range of new contexts.[334]

States differed in how they implemented the public trust doctrine. In some, a public trust meant simply that "the state's title to certain resources is impressed by a trust in favor of particular public uses" or that "that certain resources are subject to a presumption that they will be devoted to particular public uses unless the state legislature specifically legislates to the contrary."[335] In yet other states, the doctrine has been constitutionalized.[336] More recently, lower federal courts have divided over whether the public trust doctrine could extend to federal government assets.[337]

---

[325] Ryan, *supra* note 4, at 161-62 (noting that the public trust doctrine operates both as a constraint upon and a grant of sovereign authority).

[326] Concerns have been raised about the "democratic deficit" created by judicial enforcement of the public trust doctrine. Merrill, *Public Trust*, *supra* note 310, at 284. This argument rests on the fallacy of composition: It assumes a polity cannot be democratic unless all its consistent elements are democratically responsive. But this is plainly false. No one thinks that elected leaders should have plenary power over police forces, election management, or the regulation of speech to ensure democratic responsiveness. The public trust doctrine is simply another way of promoting democracy by assuring some minimum level of security (here of assets) as against public misuse of private capture.

[327] *Shively v. Bowlby*, 152 U.S. 1, 14-25 (1894)

[328] *Geer v. Connecticut*, 161 U.S. 519, 527, 529 (1896).

[329] Rose, *Comedy of the Commons*, *supra* note 316, at 738 ("*Illinois Central* sparked a new line of state 'public trust' jurisprudence.").

[330] *Lamprey v. Metcalf*, 53 N.W. 1139, 1143 (Minn. 1893).

[331] *See, e.g., Matthews v. Bay Head Imp. Ass'n.*, 471 A.2d 355, 363 (N.J. 1984) (public easement to access beach); *Borough of Neptune v. Borough of Avon-by-the-Sea*, 294 A.2d 47, 47 (N.J. 1972) (public use of beach).

[332] *Robinson Twp., Washington Cty. v. Com.*, 83 A.3d 901, 955 (2013).

[333] Rose, *Idea of the Public Trust*, *supra* note 3, at 352.

[334] For a summary, see Richard J. Lazarus, *Changing Conceptions of Property and Sovereignty in Natural Resources: Questioning the Public Trust Doctrine*, 71 IOWA L REV 631, 643-56 (1986).

[335] Merrill, *Public Trust*, *supra* note 310, at 261-62.

[336] *See, e.g., Pa. Envt'l Def. Found. v. Commonwealth of Pa.*, 161 A.3d 911, 933 (Pa. 2017) (finding state constitutional obligations to "prohibit the degradation, diminution, and depletion of our public natural resources" and "act affirmatively via legislative action to protect the environment"); *Chelan Basin Conservancy v. GBI Holding Co.*, 413 P.3d 549, 558 (Wash. 2018) (noting the public trust doctrine's "constitutional underpinning").

[337] *Compare Alec L,. v. Jackson*, 863 F. Supp. 2d 11, 15 (D.D.C. 2012) (no), *with Juliana v. United States*, 863 F. Supp. 3d 1224, 1259 (D. Or. 2016), overruled on other grounds 947 F.3d 1159 (9th Cir. 2020) (yes).

Electronic copy available at: https://ssrn.com/abstract=3794780

The public trust doctrine can be put into play by ex post review of how an asset is used, or analysis of the interest-group dynamics around how the asset is used. Taking the second tack, *Illinois Central* limited the alienation of public trust assets by asking whether a sale furthered "the interests of the public" and had no "substantial impairment of the public interest in the lands and waters remaining."[338]

But this has not locked reviewing courts into one modality of review. As a recent federal district court about the use of Chicago public trust land for the Obama presidential library explains, there are several ways of implementing a public trust.[339] The district court noted that where a public trust is statutorily designed over land that has never been submerged, a reviewing court using Illinois law applies a minimal form of review. It ask only whether the law creating the trust "is sufficiently broad, comprehensive and definite to allow the diversion" at issue.[340] Where submerged land is at issue, though, the reviewing court engages in more intensive review. It asks whether "the 'primary purpose' of a legislative grant is 'to benefit a private interest.'"[341] In a similar vein, the Hawaii Supreme Court has held that courts should take a "close look" at decisions taken by public authorities respecting a public trust asset. This approach is akin to the "'hard look' that federal courts have said is required in reviewing consequential decisions by environmental and consumer safety regulatory decisions."[342] Courts in Idaho, North Dakota, and California take the same tack.[343] In Wisconsin, the "hard look" approach to public-trust assets has congealed into a more substantive form, with courts considering five factors, including the extent of public control, the existence of a public purpose, and the disappointment of those previously using the public asset.[344] Like *Illinois Central*, these decisions reflect a substantive commitment to democratic control coupled to an awareness of democracy's frailties.

A virtue of the public trust doctrine, in sum, is that it is very ductile and so capable of flexing to fit over many different kinds of assets—from oysters and fish to navigable passage to fresh-water to park land. Its overriding touchstone is democratic control of common resources tempered toward the preservation of that asset. It is also durable: It provides a way to entrench a persisting governance framework for an asset. The question today is whether it can be adapted to the personal data context.

---

[338] *Illinois Cent. R. Co. v. State of Illinois*, 146 U.S. 387, 453 (1892).

[339] The district court in that case denied relief on public trust ground, a holding that was reversed on appeal by the Seventh Circuit Court of Appeals, with then-Judge Barrett writing, on Article III standing grounds. *Protect Our Parks, Inc. v. Chicago Park Dist.*, 385 F. Supp. 3d 662 (N.D. Ill. 2019), *aff'd in part, vacated in part, remanded,* 971 F.3d 722 (7th Cir. 2020)

[340] *Id.* at 678.

[341] *Id.* at 682.

[342] Merrill, *Public Trust*, *supra* note 310, at 281.

[343] *See, e.g.*, *Kootenai Envtl. Alliance v. Panhandle Yacht Club*, 105 Idaho 622, 628-31, 671 P.2d 1085, 1091-94 (1983) (requiring a 'close look' at conveyance of trust property); *In re Stone Creek Channel Improvements*, 424 N.W.2d 894, 902-03 (N.D. 1988) (closely examining administrative record of public trust's disposal); *National Audubon Soc'y v. Superior Court*, 33 Cal. 3d 419, 658 P.2d 709, 189 Cal. Rptr. 346, *cert. denied*, 464 U.S. 977 (1983) (same).

[344] *Paepcke v. Pub. Bldg. Comm'n of Chicago*, 263 N.E.2d 11, 19 (1970) (summarizing and adopting Wisconsin law).

Electronic copy available at: https://ssrn.com/abstract=3794780

B. Fitting Data within a Public Trust Framework

The public trust form has potential in the personal data context because of congruities of form and function. First, there is a fit between the formal qualities of data as an asset and the formal qualities of other assets subject to the public trust doctrine. Second, there is a close match between the jurisprudential ambitions baked into the public trust and the desirable mix of public and commercial uses of personal data.

1. Data is an Archetypal Public Trust Asset

Let's start with a negative: There is a profound doctrinal and intellectual mismatch between the standard individualized form into which property is usually sliced, and the way in which personal data is circulated and exploited. This incongruity emerges most clearly in the sharp mismatch between the leading justification for creating discrete, fungible property interests and the manner in which value is in fact extracted from personal data.

Supreme Court precedent on property in information strongly suggests that there is no individualized property interest in personal data. At least as a matter of the black-letter law, therefore, the aggregations of data that comprise the most important asset in the personal data economy are simply not within the private property system. The leading decision on information aggregations is Justice O'Connor's 1991 opinion *Feist Publications v. Rural Telephone Service*.[345] *Feist* concerned a copyright claim to the compilation of names, addresses, and telephone numbers in a white-pages directory. Taking originality as a constitutional floor, the Court held that the "selection, coordination, and arrangement of … white pages do not satisfy the minimum constitutional standards for copyright protection."[346] Mere "facts" are "uncopyrightable."[347] After *Feist*, lower courts have found compilations to be copyrightable only when they evince some "judgment" about divisions within data or summary statistics.[348] Whether a particular aggregation created through personal data economies reflects sufficiently creativity depends, of course, on its particular facts. But in at least one decision, locational data generated through a sensing net has been characterized as beyond copyright's constitutional domain.[349] So even when a given database architecture can be copyrighted, the actual data within it will not thereby become property.[350] Although contract, trade secret, antitrust, privacy, and other bodies of law may inflect how

[345] 499 U.S. 340 (1991).
[346] *Id.* at 362.
[347] *Id*; *see also Int'l News Serv. v. Associated Press*, 248 U.S. 215, 235 (1918) (rejecting the idea of "property in news").
[348] *CCC Info. Servs., Inc. v. Maclean Hunter Mkt. Reports, Inc.*, 44 F.3d 61, 67 (2d Cir. 1994).
[349] Cyrus Farivar, *Judge, Siding with Google, Refuses to Shut Down Waze in Wake of Alleged Theft*, at http://arstechnica.com/tech-policy/2015/12/judge-siding-with-googlere-fuses-to-shut-down-waze-in-wake-of-alleged-theft.
[350] *Assessment Techs. of WI, LLC v. WIREdata, Inc.*, 350 F.3d 640, 641 (7th Cir. 2003) (Posner, J.) (rejecting the idea that "a copyright owner to use copyright law to block access to data that not only are neither copyrightable nor copyrighted, but were not created or obtained by the copyright owner"); *Hutchins v. Zoll Med. Corp.*, 492 F.3d 1377, 1385 (Fed. Cir. 2007) ("Although the compilation of public information may be subject to copyright in the form in which it is presented, the copyright does not bar use by others of the information in the compilation.").

Electronic copy available at: https://ssrn.com/abstract=3794780

information can be alienated or used, constitutional basics dictate that data is "largely free from properly rights"[351] defined in terms of private ownership.

This absence of an individualized property interest in information means that there can be no objection from prior owners to the recognition of a common, aggregate form of property in personal data. In particular, it vitiates objections on Fifth Amendment grounds pursuant to the Takings Clause. It does not supply, though, a positive reason for *adopting* a public trust for data.

Yet from another perspective, the economic logic of property rights does conduce well to an aggregative, common governance regime. An individualized, granular, and standardized mode of property is appropriate when social value is realized through the *partition* of assets. In the personal data economy, however, value is created through *aggregation*.[352] A single data point is rarely of much value on its own, at least unless it concerns a celebrity or public figure. This means that the commercial value of personal data emerges only when it has been lumped together. It also means that while a few harms associated with personal data economies concern individualized data, many emerge only after aggregation. As a rough first cut, privacy, dignity, and exploitation worries attach to discrete items of data without regard to aggregation. In contrast, economic inequality, democratic backsliding, state dominance, and the underproduction of public goods are associated with data aggregates. To the extent the law seeks to mitigate the latter as well as the former, it should intervene in respect to data aggregates, not target the flow of discrete bits of information.

The public trust is commonly deployed for assets that are hard to slice up into discrete, individualized assets. These include clean ambient air,[353] navigable waters,[354] ground water,[355] the recreational use of a lake,[356] and beach access.[357] Divisible resources, such as oysters and fish, might be parceled out by quota systems, but their component items are fungible and better considered as aggregates. As such, the public trust has developed for assets with the same relation to aggregation as personal data.

Further, while the harms of personal data economies cannot be captured without an accounting of data in its aggregate form, the individuation of data as property does not yield the payoffs associated with other discrete and parceled forms of data. In the leading economic account of individual property rights schemes, Thomas Merrill and Henry Smith argue that the transaction costs of dealing in property are a function of information costs imposed on third-parties.[358] "As a

---

[351] Lothar Determann, *No One Owns Data*, 70 HASTINGS L.J. 1, 5 (2018). For example, of particular relevance to the sensing net, "manufacturers … generally cannot claim trade secret ownership rights in the data and information generated by the devices they sell to customers." *Id.* at 16.

[352] KELLEHER & TIERNEY, *supra* note 29, at 56-58 (describing a widely used process of "data capture and generation through data preprocessing and aggregation" called "CRISP-DM").

[353] *Robinson Twp., Washington Cty. v. Com.*, 623 Pa. 564, 652, 83 A.3d 901, 955 (2013).

[354] *Martin v. Waddell's Lessee*, 41 U.S. (16 Pet.) 367, 413 (1842).

[355] *Envtl. Law Found. v. State Water Res. Control Bd.*, 237 Cal. Rptr. 3d 393, 399-403 (Cal. Ct. App. 2018) (public trust doctrine protected groundwater tributaries of navigable waters).

[356] *Nat'l Audubon Soc'y v. Superior Court*, 33 Cal. 3d 419, 441, 658 P.2d 709, 724 (1983).

[357] *See, e.g., Matthews v. Bay Head Imp. Ass'n.*, 471 A.2d 355, 363 (N.J. 1984).

[358] *See* Thomas W. Merrill & Henry E. Smith, *What Happened to Property in Law and Economics?*, 111 YALE L.J. 357, 359 (2001) ("[P]roperty imposes an informational burden on large numbers of people, a burden that goes far beyond the need for nonparties to a contract to understand the rights and duties of contractual partners.").

Electronic copy available at: https://ssrn.com/abstract=3794780

consequence, property is required to come in standardized packages that the layperson can understand at low cost."[359] Because these information costs "impinge upon a very large and open-ended class of third persons"[360] in market contexts, standardization is necessary to trade's viability. Merrill and Smith point out that even though items of personal property can vary across in multitudinous ways, legal standardization is most useful "in connection with the dimensions of property rights that are least visible, and hence the most difficult for ordinary observers to measure."[361]

This logic does not translate well into the personal data context. The standardization of data does not have the same payoffs as the standardization of land and chattels. Rather, it presents different and sharper challenges. Personal data is much more difficult to standardize than goods. Data from the varied digital tributaries feeding the larger personal data economy will be as varied as personal property, but will lack the manifest and observable qualities of "size, shape, color, or texture" that obviate certain forms of standardization.[362] Data will vary in nature and content depending on whether it comes from a cellphone, a vacuum cleaner, a dating app, an artificial pancreas, or a public surveillance camera.[363] It will not reliably have "complementary attributes," while the "information-hiding and limited interfaces" used to standardize land and chattels may be available only by losing precisely that which creates value in the first instance—the informational content of the data.[364] Standardizing will often both require large investments in computation, and would likely come with heavy informational losses.

This is not to say that data is on all fours with assets historically subject to a public trust. The latter commonly preexist man-made action or commercial investments, and can easily be seen to require protection from such investments. Yet this distinction, while real, is easy to overdo. Assets such as lakebed property close to Chicago or fresh water near Los Angeles merits protection not because it is valuable in isolation. To the contrary, it has value—and needs legal shelter—because of commercial investments in proximate real property. The noncommercial interactions swept into social media networks can, similarly, be thought of as a 'natural' phenomenon that accrued value because of a shift in locus.[365]

---

[359] *Id.*; *see also* Thomas W. Merrill & Henry E. Smith, *Optimal Standardization in the Law of Property: The Numerus Clausus Principle*, 110 YALE L.J. 1, 38 (2000) [hereinafter "Merrill & Smith, *Optimal Standardization*"] (arguing that "the objective [in designing property rights] should be to minimize the sum of measurement (and error) costs, frustration costs, and administrative costs" though "optimal standardization").

[360] Thomas W. Merrill & Henry E. Smith, *The Property/Contract Interface*, 101 COLUM. L. REV. 773, 802 (2001)

[361] Merrill & Smith, *Optimal Standardization*, *supra* note 359, at 34.

[362] *Id.*

[363] An exception is locational data, which will be possible to standardize.

[364] Henry E. Smith, *Property as the Law of Things*, 125 HARV. L. REV. 1691, 1703 (2012); *see also id.* at 1705 ("Because delineation costs are not greater than zero, which strategy one uses and when one uses it will be dictated in part by the costs of delineation – not just the benefits that correspond to the use-based purposes of the property").

[365] A possible distinction between assets traditionally subject to a public trust and data is the former's rivalrous quality. That is a public trust is established when an asset is capable of exhaustion. Data, however, cannot be exhausted: It is nonrivalrous. There are traces of this idea in some cases. *See, e.g., Nat'l Audubon Soc'y v. Superior Court*, 33 Cal. 3d 419, 432, 658 P.2d 709 (1983). But reported decisions do not reflect a purely instrumental account of what is and what is not a public trust. To the contrary, they reflect a normative understanding reflecting a sense of what ought to be in the public as opposed to the private domain. I am grateful to Lee Fennell for discussion of this point.

Electronic copy available at: https://ssrn.com/abstract=3794780

The public trust form, all said and done, is well-fitted in theory to the governance and management of personal data. Information is not personal property. It comes in aggregates that are poor fits for the day-to-day system of sliced-up, discrete property entitlements for chattels and land. And the principal justification for cleanly individuated and sharply distinguished property is largely inapposite in modern data economies.

2. The Justifications for the Public Trust Doctrine Apply to Personal Data

At its core, the public trust is a governance regime designed "to protect the people's common heritage" from public and private misuse.[366] An asset fit for public trusteeship, accordingly, is a "common" one in the sense that it can be enjoyed by an economically and sociological varied public. Fishing for trout or oysters, larking about on a sandy Atlantic beach, or enjoying fresh potable water—all these are goods enjoyed by the public at large. A rule of common access is markedly progressive in its distributional effect. Moreover, in each case, the asset in question is durable: it is a resource that has historically been enjoyed from one generation to the next—and is therefore a legitimate object of people's expectations.[367]

At a high level of generality, there are five parallels between these justifications and the regulatory gaps to be found in public data economies. To begin with, the pools of information created through engagement with platform economies and sensing net are the product of common labor. Their value exists thanks to the mutual expression of our "natural compulsion to reciprocate" and "existing solidaristic bond[s]."[368] On familiar Lockean grounds, that endows their collective creator—not one single person, but a networked assemblage of all—with a collectively held title.[369] The public trust hence puts ownership in the hands of those who deserve it, , and allows them to reap a fair return via user fees.

Second, personal data is not only created by common, albeit uncoordinated, action. It could also be designed for the enjoyment and benefit of all, rather than for the benefit of a narrow coterie of monopsonistic purchasers and brokers. That is, personal data creates a choice: Should it be exploited for the good of the few, or titrated for the benefit of the many? The public trust in data is a way to create democratic control over a resource's use—barring undesirable effects and eliciting public goods.

---

[366] *Id.* at 441.

[367] *Paepcke v. Pub. Bldg. Comm'n of Chicago*, 263 N.E.2d 11, 19 (1970) (noting the role of public expectations in justifying a public trust).

[368] Fourcade and Kluttz, *supra* note 135, at 10.

[369] *See* JOHN LOCKE, SECOND TREATISE OF GOVERNMENT 19 (C.B. Macpherson ed., Hackett Publ'g Co. 1980) (1690) ("Whatsoever then he removes out of the state that nature hath provided, and left it in, he hath mixed his *labour* with, and joined to it something that is his own, and thereby makes it his *property*."). Locke justified the individual's ownership right by the tendency of individual ownership to conduce to more productive uses of land. Which "does not lessen, but increase the common stock of mankind." *Id.* at 23. By crude analogy, the assetization of aggregate data serve the same net welfarist end. Conscripting Locke for the cause of common property in information is not as odd as it might first seem. In the informational domain, Locke opposed the Licensing Act of 1662 because of the chill it cast on "authors' abilities to create derivative works, inhibiting communal knowledge and progress. Alexander D. North over, *"Enough and As Good" in the Intellectual Commons: A Lockean Theory of Copyright and the Merger Doctrine*, 65 EMORY L.J. 1363, 1374 (2016). He urged a "limited copyright term that promote[d] a robust public domain."

Electronic copy available at: https://ssrn.com/abstract=3794780

Third, personal data as an asset is durable. It cannot be exhausted (although its misuse can yield spoilage of the public square). It endures for generations.[370] The reservoirs of personal data being filled now are thus as much a kind of common heritage as the air we breathe. What is spoiled is less the resource, but the ambient social conditions of equality and adequate resources for all that make personal data economies useful in the first instance.

Fourth, several of the most penetrating normative challenges of personal data economies arise from disparities of information and influence between firms and the public. Across varied fronts, the concentration of profits and knowledge in a small number of firms is a fulcrum of normative concern. Both platform economies and sensing nets extract data that firms value in ways users cannot. This many-to-one character of many platform economies, which is baked into both design and technological detail, spills over into another asymmetry: Even if Facebook yields substantial gains for individual users, the sheer gap between their numerosity and Facebook's unity has distributive effects. Small per-person profits captured by a single firm from millions daily generates a large, lopsided concentration of both wealth and influence. Technical and legal complexity allowing firms to exploit workers' and users' cognitive weak spots only exacerbates this tilt.[371] The public trust doctrine changes this many-to-one dynamic into a one-to-one contest. It hence levels the playing field.

This leveling means the public trust can be a direct response to many of the critiques lodged against data economies. Concerns about exploitation, inequality, and the under-supply of public goods are all thus best understood as objections to the regressive dynamics layered into personal data economies. Retail privacy worries about improper sharing, data breaches, and unanticipated affordances also have a distributive character: In addition to the first-order objection to privacy losses, they are all instances in which platform economies or sensing nets extract a greater informational surplus than consumers reasonably anticipate. Concerns about democratic backsliding and state repression are also objections to certain kinds of asymmetrical arrangements; they focus, however, on political rather than economic hierarchies. The America public trust doctrine as revived and rearticulated by *Illinois Central* provides a well-tailored vehicle for addressing those redistributive concerns. From its inception, it was understood as a means of curbing the influence of powerful interest groups over important common assets.[372] The *Illinois Central* Court conceived of the problem presented by the Lake Front Act in terms of legislative corruption, resulting in the improper transfer of assets to the company.[373] The state today may not act corruptly. It rather fails, either by negligence or undue influence, to prevent immediate harms or larger structural imbalances from materializing. As with the Lake Front Act, the effect is to allow an undue part of the value created by a public resource to flow to small number of firms. The data public trust corrects for that.

---

[370] Anya E.R. Prince and Daniel Schwarcz, *Proxy Discrimination in the Age of Artificial Intelligence and Big Data*, 105 IOWA L. REV. 1257, 1274 (2020) (discussing this problematic in the insurance context).
[371] *Cf.* ROSENBLATT, *supra* note 151, at 199 (explaining how Uber would not supply a handbook to drivers, leaving them in the dark and having to figure out work-related rules by networking with fellow drivers).
[372] *See* text accompanying notes – to --.
[373] *Illinois Cent. R. Co. v. State of Illinois*, 146 U.S. 387, 451-52 (1892) (noting concerns about the Lake Front Act); Kearney and Merrill, *supra* note 316, at 806 (arguing that Justice Field's opinion offered "a narrative of monopoly privilege subverting the public interest").

Electronic copy available at: https://ssrn.com/abstract=3794780

Fifth and finally, at the remedial end, the public trust harnesses "checks and balances of government" to prevent an asset's misuse, but at the same time reposes no "blind" trust in the state.[374] It accounts for both market and government failures. Hence, from Justice Field's opinion in *Illinois Central* onward, the public trust doctrine has been organized around the creation of judicial mechanisms to ward off various ways in which government might connive with interest groups to spoil or alienate an asset to the detriment of the public at large.[375] It is a means to regulate "the collective ownership [of] public property" through a mix of "inalienability" rules and other restraints.[376] Although the *Illinois Central* Court enforced an inalienability rule to void the transfer of Chicago's lakefront, the doctrinal entailments of a public trust can be more subtle and varied, extending from a light review of the formal qualities of an asset's use to a hard look at the motives and justifications for a particular arrangement. A public trust might also be a semicommons-arrangements in which common usages are mixed with extractive private uses.[377] Similarly, a public trust in data can be hedged around with rules to prevent the government's misuse of its contents, such as the kind of limits on disclosure and sharing that apply to social security and taxing authorities.

C. Imagining the Public Trust in Data

What would this mean in practice? It is possible to imagine the implementation of various public trust regimes that specifically accounted for and mitigated harms detailed in Part II. Without being exhaustive, I sketch here one way in which a governance arrangement of this sort might be deployed. I first explain why it would be wise to focus upon cities and states as the font of such regulation. Next, I offer a sketch of how a public trust in data might work on the ground.

1. Jurisdictional Choice for a Public Trust in Data

Subnational jurisdictions, and in particular cities, are the most promising starting point for a public trust in data. A "majority of the world's population lives in cities," a situation that "marks a major and unprecedented transformation of the organization of society, both spatially and geopolitically."[378] Cities are hence directly accountable to the vast agglomerations of individuals now generating most locational, behavioral, and social data. Moreover, they tend to be geographically compact. "Cities develop because they … provide residents with the advantages of big, diverse, and productive markets and creative ferment."[379] Cities are also responsible for addressing many, if not all, of the social harms spilling over from personal data economies.[380] Hence it is no surprise that we have already seen that cities such as Barcelona, Amsterdam, New York, and Washington experimenting with proto-trust forms and kindred regulatory strategies for

[374] *Robinson Twp., Washington Cty. v. Com.*, 623 Pa. 564, 655, 83 A.3d 901, 956 (2013).
[375] *See supra text* accompanying notes – to --,
[376] Richard A. Epstein, *The Public Trust Doctrine*, 7 CATO J. 411, 418-21 (1987).
[377] Henry E. Smith, *Semicommon Property Rights and Scattering in Open Fields*, 29 J. LEGAL STUD. 131, 131 (2000).
[378] Ran Hirschl, *Constitutions and the Metropolis*, 16 ANN. REV. L. & SOC. SCI. 59, 60 (2020)
[379] David Schleicher, *The City as a Law and Economic Subject*, 2010 U. ILL. L. REV. 1507, 1509 (discussing literature in agglomeration economics).
[380] This may be especially true of redistributive policies. *See* Richard C. Schragger, *Federalism, Metropolitanism, and the Problem of States*, 105 VA. L. REV. 1537, 1540 (2019) "As economic activity becomes concentrated, those cities and regions have more capacity to redistribute than the standard model predicts ….").

Electronic copy available at: https://ssrn.com/abstract=3794780

platform economies and sensing nets.[381] The extraterritorial reach of such regulation poses no barrier. In 2018, the U.S. Supreme Court affirmed the constitutional power of subnational jurisdictions to impose sales taxes upon out-of-state retailors.[382] The application of the public trust doctrine by a locality to an extraterritorial platform would, equally, be permissible and present no distinct constitutional difficulty related to extraterritorially.

As a practical matter, data subject to a public trust is likely to be stored in the cloud. This might entail the use of a single storage location or a "shard" structure whereby a "single file can be broken into components and stored in different countries, and intelligence embedded in the network decides where to send and store the data."[383] A municipal public trust regime would not necessarily require that data be held locally, as the proposed Indian law would.[384] It would, though, demand that where data falling within a public trust was stored elsewhere, it would continue to be subject to that municipality's regulation. A data storage regime that located data in a jurisdiction with conflicting or inconsistent regulation would therefore be a violation of the public trust.

The choice of a subnational unit also ensures that a public trust can more precisely correct distributive pathologies of the data economy. Such a trust can impose user fees on firms that wish to exploit data, and then direct their proceeds to the populations producing the latter. At the national level, there is a greater chance that such funds might be repurposed to other ends.

2. Creating a Public Trust in Data

What would a public trust in data look like? Its establishment would have three basic steps.

To begin with, a state or local government would by legislation or ordinance establish a public trust in the data created by its citizens within its geographical ambit. Unlike older public trusts established through case-law, this one would be created and grounded in democratic (legislative) deliberation and choice.[385] Indeed, one of the advantages of a public trust structure is the possibility of subjecting personal data aggregations to greater degrees of democratic control. The legislation would begin with a declaration that title to the data resided in the public trust, without regard to where the data was physically housed.

A jurisdiction would next have to decide on what data to include. A logical place to begin, and the starting point for Barcelona's Decidem platform,[386] is the locational data created by public and private sensing nets within the jurisdiction. Pursuant to the Decidem platform, for example, the winner of a contract to supply a city-wide bike share system would "have to give the city back all the information it collects about how citizens are using the service."[387]

---

[381] *See supra* text accompanying notes – to --.
[382] *South Dakota v. Wayfair Inc.*, 138 S. Ct. 2080, 2092 (2018) ("'It has long been settled' that the sale of goods or services 'has a sufficient nexus to the State in which the sale is consummated to be treated as a local transaction taxable by that State.'" (citation omitted)).
[383] Paul M. Schwartz, *Legal Access to the Global Cloud*, 118 COLUM. L. REV. 1681, 1695 (2018) [hereinafter "Schwartz, *Global Cloud*"].
[384] Andrew Keane Woods, *Against Data Exceptionalism*, 68 STAN. L. REV. 729, 752 (2016) (criticizing data localization rules on efficiency grounds).
[385] Araiza, *supra* note 14, at 696 (noting criticism of judicial discretion in public trust doctrine).
[386] Lewin, *supra* note 5.
[387] *Id.*

Electronic copy available at: https://ssrn.com/abstract=3794780

Decidem is not the only initiative to focus on locational data. In 2018, Washington, DC partnered with a not-for-profit called SharedStreets to give the municipality pickup and drop-off data from Uber.[388] This data is then used "to understand whether … drivers are too often blocking traffic to pick up passengers" and even to "reconsider … street designs or traffic patterns to accommodate the new ways of getting around."[389] A step toward the public trust form, SharedStreets is a nonprofit rather than a legislative creation. Uber's participation is, though, voluntary rather than mandatory. Yet this is easy to change. A year later, indeed, New York City mandated the disclosure of the same data by ride-sharing companies "to learn more about what's happening on the streets," "to plan … how to beat traffic and improve road safety," to "monitor the number of wheelchair-accessible vehicles picking up passengers" and to "enforce its new minimum wage rule for app-based drivers."[390] In other words, New York is using a public quasi-ownership strategy for data both as a means of creating public goods that would otherwise go unrealized and also to prevent drivers' exploitation. Nevertheless, the New York initiative is limited to the extent that it does not give the municipality actual title to the data. That step would not only allow for the creation of public goods, but the conditioning of firm access upon the avoidance of harms documented in Part II.

There is no reason to limit the public trust to the obvious and intuitive example of locational data. Rather, these applications might be stepping stones for more aggressive applications—to other sensing net data and to locally acquired platform economy data. Indeed, the public trust for data would not reach its full potential without these latter applications. Hence, a public trust could be extended to all sensing-net devices within a jurisdiction. This might include, as a threshold matter, all such devices operating in public spaces. It could also be extended to data from devices, such as Alexa, that operate within domestic spaces. It could be applied to the data generated through commercial transactions (such as Amazon) and through social network. Platforms would continue to generate and store all this data—to be clear, there is no thought here of cities building their own data-storage centers—but the way in which such data could be disseminated, exploited, and monetized would be constrained by the public trust. The latter would, further, have a claim to a portion of profits generated by the exploitation of user data that could be imposed in the form of user fees.

There is a further question of whose data would fall within a trust. A city has a plausible claim to data locally produced by a local resident through interactions on a platform economy. When a resident of the municipality logs on from a local IP address, accessing a social network, they are creating valuable data for aggregation and circulation. They are also risking the harms listed in Part II. A city that asserts an interest in the data thereby created, and seeks to subject that data to regulation through the vehicle of a public trust, is properly acting in its citizens' interests. It is a closer case whether the same is true for visitors' data—but the difficulty of distinguishing between residents and visitors might counsel for treating both as subject of the public trust. Moreover, recall that platforms such as Facebook install cookies on users' cellphones and computers that capture both call data and also traffic to other web sites.[391] This data, which does

[388] Aarian Marshall, *Uber Makes Peace With Cities By Spilling its Secrets*, WIRED (Apr. 16, 2018), https://www.wired.com/story/uber-nacto-data-sharing/; see also SharedStreets, https://sharedstreets.io/.
[389] *Id*.
[390] Marshall, *NYC*, *supra* note 8.
[391] Srinivasan, *supra* note 60, at 71-72.

Electronic copy available at: https://ssrn.com/abstract=3794780

not directly serve users, is also generated by local activity and has commercial value. It too presents sharp normative concerns about privacy and exploitation. Accordingly, a municipality has an interest in this data too. Data subject to this public trust need not be stored locally, but (as discussed) should be amenable to local control. The municipality should then subject the data to a schedule of permitted and impermissible uses. These should be a matter for democratic determination, albeit within the board limits imposed by the public trust's fiduciary framing. Overall, the goal would be to continue to allow commercial exploitation while constraining externalities and structural harms.

Assume, then, that a trust over some class of data has been determined. The third and final step in fashioning a public trust is the fashioning of bilateral obligations to constrain state and private handling of data to promote broad public benefits rather than narrowly channelized private profits. To this end, legislation would describe the terms and conditions for private exploitation of the data, either directly or in combination with other databases. A schedule of permitted and impermissible uses should be determined by democratic means, albeit within the board limits imposed by the public trust's fiduciary framing. Overall, the goal would be to continue to allow commercial exploitation while constraining externalities and structural harms. Further, the trust should establish rules to prevent misuse of the data by the state itself.

As a threshold matter, a municipal jurisdiction could require that information held in the global cloud—whether sharded or localized—have a local "data trustee" with "the exclusive ability to access the data" regardless of where physical storage occurred.[392] This would mitigate conflicts-of-law problem that might arise from the globalization of storage capacities. It would also provide the city with a focal point for regulation and oversight. By fortunate coincidence, both Facebook and Google have announced that they are moving legal responsibility for their data from Dublin to California "as a consequence of Brexit."[393] This change lowers one barrier to the creation of a public trust in data. The data trustee would be charged with the technical implementation of trust rules: It would, to that effect, monitor *both* private and public uses of the data to guard against misuse on either side. In the absence of a trustee, a government office (such as a state's attorney general could play this role).

The trusteeship element of the public trust distinguishes it from other feasible regulatory interventions in private data economies. It thus brings into focus the ways in which the public trust differs from other potential regulatory forms. A state or municipality in theory has the power to impose limits on how data is collected or used already. But it lacks instruments of ongoing supervision and management to ensure this a way to at downstream uses do not violate its rules. A trustee fulfills that role using the enforcement related authorities discussed below. Further, most existing regulatory strategies are aimed at controlling private malfeasance. In the context of data economies, both public and private action presents a risk of harm. A trusteeship mechanism is a way to achieve durable regulatory control over data as an asset. It is also a way to combine limits on private with a constraint on improper state action. Finally, the public trust mechanism is a means of bundling together—legislating in one fell swoop—an array of constraints on both private and public action. It is far more likely that a defensible set of measures will be adopted if this bundled approach is taken than if regulation is pursued piecemeal.

---

[392] Schwartz, *Global Cloud*, *supra* note 383, at 1697.
[393] *Chlorinated Facebook*, THE ECONOMIST (Jan. 9, 2021), at 48.

Electronic copy available at: https://ssrn.com/abstract=3794780

More substantive obligations could then take either positive or negative forms. Consider four ways in which the public trust doctrine could be crafted to mitigate harms. Again, I emphasize that what follows is very much a sketch—with more details turning on the specifics of the kinds of data subject to trust control, and the particulars of the jurisdiction.

First, the fact of state ownership of data constrains the private exploitation of informational and market inequalities. States can condition access and use of personal data on rules that minimize discrete privacy losses and acts of individual exploitation.[394] A company such as Uber that gathers and exploits individual locational data, for example, might be required to follow labor policies and pricing strategies that did not merely maximize private profits.[395] A company such as Facebook, which "does not disclose information about its uses of data … at all" would at a minimum be required to account for its commercial strategies before operating in a jurisdiction.[396] A sensing net that produced visual data that included faces might be restricted from allowing these to be used as training data for controversial facial recognition instruments.

Second, private uses of personal data could be conditioned to agreement to share a fraction of profits with the trust. In effect, use taxes would ensure that those create data benefit from its transformation into an asset. Rather than paying individuals for data, a public trust is a way to recoup some return from the emotional, intellectual, and even physical labor that allowed its creation. The trust, in turn, would be legally obliged to apply those funds to the general benefit of a city or state's residents. Local labor hence becomes a fiscal foundation for local public goods.

Third, access to data for commercial use could be conditioned on an agreement to forego certain harmful transformations. For example, a social media platform subject to the public trust regiment might be required to demonstrate that its network architecture did not facilitate the dissemination of false political information or deliberately polarizing propaganda.[397] It might be compelled to show that it did not, even inadvertently, present different interfaces to men and women.[398] It would have to demonstrate that its algorithm was designed not merely to maximize engagement as such, but to elicit forms of social-media activity that are consistent with democratic norms.

On this last score, it is worth underscoring that a public trust need not engage only in constraint. Like the Silicon Valley Data Trust,[399] it might also aim at the positive production of needful public goods. Hence, a condition of access to personal data might be the generation of

---

[394] *Cf.* COHEN, *supra* note 47, at 65 (describing how platforms are designed to "maximize opportunities for behavioral data extraction").
[395] Marshall, *NYC*, *supra* note 8, at id. (noting enforcement strategies for city's minimum wage rule).
[396] COHEN, *supra* note 47, at 61.
[397] Leading proposals focus on deepening transparency by requiring "transparency, education, and 'nudges.'" Abby K. Wood & Ann M. Ravel, *Fool Me Once: Regulating "Fake News" and Other Online Advertising*, 91 S. CAL. L. REV. 1223, 1253 (2018). The public trust form would allow more aggressive regulatory interventions going to network architecture.
[398] For an example of how this happen thanks to algorithm design and not designer bias, see Anja Lambrecht and Catherine Tucker, *Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads*, 67 MANAGEMENT SCI. 2966, 2967 (2019).
[399] *See supra* text accompanying notes – to --.

Electronic copy available at: https://ssrn.com/abstract=3794780

public goods that would otherwise be difficult to create. For example, a ride-sharing company or a traffic app that generated locational data for vehicles might be allowed to operate only if could certify that its recommendation apps minimized traffic and air pollution. An individual locational service such as FourSquare might be allowed to operate only if it also committed to sharing data with public health authorities to identify 'hotspots' during pandemics (or, indeed, flu season).[400] A two-sided platform for consumers and merchants such as Amazon might be required to place data on usage patterns into a trust, where it could be tapped by individuals looking to build new platforms and products.[401] A social network might commit to providing timely information on the diffusion of anti-democratic messaging, such as the speech and mobilization that led to the January 6, 2021, Capitol siege. At present, there is "no financial or political incentive to look for the evidence [of misinformation and conspiracies being spread]."[402] By conditioning access to personal data on a network's willingness to diligently root out misinformation, the public trust doctrine yields the beginning of a solution. The production of public goods may be part of the quid-pro-quo reached with firms in allowing them access to personal data.

Finally, a public trust needs enforcement mechanisms to head off the risk of both private abuse and interest-group capture of a state agency. This could be done through the creation or a trustee, or by giving an official such as a state's attorney general a durable oversight mandate. A trustee, for example, should have the power to seek judicial intervention to enjoin impermissible dissemination or use of public-trust data. It should be able to seek fines for past conduct, even though the sheer scale of certain platforms makes this approach somewhat less than effective. In extreme cases, a trustee might seek to permanently enjoin a firm from using or benefiting from data under the trust. Under a regime of plural trusts established by different cities, these remedies might be amplified by linking together penalties in serious cases. For example, if a breach of trust rules is serious enough, this might be treated by law in different jurisdiction as a ground for excluding a firm from data usage. By calibrating the ensuing 'cascade' of interjurisdictional penalties in different ways, the trustee could dial up and down the severity of the ensuing enforcement regime.

A trustee should also have power to intervene against improper state action. An obvious, but overblown, objection to a public data trust is likely to flow from privacy concerns. The state, of course, already has access to much personal data because it operates social security and tax systems. It is far from clear that a public data trust, where the state itself held no information, presents new or insurmountable privacy worries. In any event, a public trust can be designed to stymie improper state action. From the beginning of the Republic, courts have enforced public trust-limits against the state at the behest of individuals. Unconstrained by the rigors of Article III standing doctrine, state courts have done so even when a plaintiff could not show some distinctive

---

[400] For an example of this application, see Thomas Varsavsky et al., *Detecting COVID-19 infection hotspots in England using large-scale self-reported data from a mobile application: a prospective, observational study*. THE LANCET PUB. HEALTH (2020).

[401] This may also yield changes to market structure. *Cf.* Morozov, *supra* note 195, at 65 ("Democratizing access to [an] information infrastructure, so that all producers can build on … emerging product insights, would surely result in a system that is far less centralized than today's ….").

[402] Zack Stanton, '*The Internet is a Crime Scene*,' POLITICO (Jan. 14, 2021), https://www.politico.com/news/magazine/2021/01/14/us-capitol-disinformation-online-qanon-trump-insurrection-459505.

Electronic copy available at: https://ssrn.com/abstract=3794780

stake in an asset's deployment.[403] A basic enforcement mechanism would involve ex post review of a license to use data, or a permission to acquire data within a jurisdiction, as consistent with the public trust. Additionally, several state courts have developed a form of "hard look" of licenses and alienations of a public-trust asset to guard against the risk of interest-group capture leading to improper spoilage of a public-trust asset.[404] For example, drawing on state environmental regulation for a basic template, the California Supreme Court has required the state's Water Resources Control Board to look closely at how water diversions to Los Angeles impact nearby Mono Lake, and to protect its public-trust uses "whenever feasible."[405] Like the trans-substantive "hard look" doctrine of federal administrative law, this approach can be extended beyond its original sphere of application. The ex post examination of the justifications for how public-trust data is employed provides perhaps the most fine-grained instrument for evaluating the integrity of public decision-making about these uses.

Finally, the effect of municipal public trust regulation of this sort, moreover, holds the promise of catalyzing more extensive reforms. Imagine a city such as New York, Chicago, or Los Angeles imposing constraints on the use of its residents' personal data. These metropolises are so large, and so globally important in the digital economy, that firms would have little choice but to comply. In effect, a version of the "California effect" might take hold.[406] To be sure, the creation of municipal-level data-use regimes creates a possibility of regulatory conflict. But this already exists given the variance in European and American regimes, and the growing possibility of state-level interventions.[407] To mitigate the risk of conflicting rules, cities could coordinate policy approaches, as they have done recently in respect to global migration policies, to minimize disruptive disuniformity.[408] It seems likely that American cities will have similar democratic preferences over many issues, and so would be able to coordinate in ways that conduced to a generally harmonious regulatory environment.

* * *

In summary, the history of the public trust doctrine provides a deep repository of legal duties and remedies for the management of common-pool assets. On the one hand, this means that mere invocation of the term "public trust" does little analytic work on its own.[409] But for a jurisdiction wishing to exercise a richer measure of democratic control over personal data economies, the doctrine can be a rich storeroom of ideas. I have outlined here one way of

---

[403] *See, e.g., Paepcke v. Pub. Bldg. Comm'n of Chi.*, 46 Ill.2d 330, 263 N.E.2d 11, 18 (1970).
[404] *See* cases cited in *supra* note --.
[405] *National Audubon Soc'y v. Superior Court*, 33 Cal. 3d 419, 446, 658 P.2d 709, 728, (1983).
[406] DAVID VOGEL, TRADING UP 248 (1995).
[407] *See, e.g.,* Cal. Civ. Code §§ 1798.100-80 (2020) (extensive state privacy law); Carol Li, *A Repeated Call for Omnibus Federal Cybersecurity Law*, 94 NOTRE DAME L. REV. 2211, 2227 (2019) ("With California's recent data privacy legislation, companies who deal with the personal data of California residents will be forced to decide whether to overhaul all their data collecting operations or build in-certain operations solely for their California clients." (quotation marks and citation omitted)).
[408] Metropolis World Association of Major Metropolises, *Position Paper Submitted as a Contribution to the United Nations Global Compact for Safe, Orderly and Regular Migration, and to the Global Compact on Refugees* (dated Dec. 12, 2017), https://www.metropolis.org/sites/default/files/20171201_metropolis_decl_eng_final_declaration.pdf.
[409] Sax, *supra* note 3, at 521 ("The 'public trust' doctrine has no life of its own and no intrinsic content. It is no more—and no less—than a name given by courts to their concerns about the democratic process.").

Electronic copy available at: https://ssrn.com/abstract=3794780

appropriating these doctrinal resources for the new information economy. My aim, however, has not been to apply a definitive blueprint. It has rather been to demonstrate the plausibility, and value, of the project.

Conclusion

New economies through which personal data is extracted, aggregated, and exchanged have created great commercial gains and large windfalls in personal well-being. After the pandemic, no one should need a reminder of the value of communicating by Facetime, WhatsUp, or similar apps. Yet at the same time, these same economies have generated significant new challenges for individuals and for societies at large.

The public trust in data provides another tool for addressing those concerns. It does so by harnessing a form of public, collective property as old as the republic. That property form has done yeoman's service already in advancing environmental goals. My central ambition here has been to demonstrate its utility in the new data economy context. Such eversion of doctrinal forms should not be a complete surprise: The common law, as a shared legal heritage, is itself a kind of public good—capable of being deployed to new and unexpected ends. A public trust in data is simply one such possibility being realized.

Electronic copy available at: https://ssrn.com/abstract=3794780