The Structural Law of Data

Bridget A. Fahey† *& Raul Castro Fernandez*††

The central concern of structural constitutional law is the organization of governmental power, but power comes in many forms. This Article is about how the law of structure regulates decision-making about, and popular control over, an increasingly potent form of power: the power government obtains from data. The government has always relied on information to meet its objectives, but the digitization of information over the last half century has yielded a distinctive form of governmental power—one that is liquid, transferable, minable, dynamic, and vital to virtually all governmental activity today.

But despite the significant literature on private-sector "data governance," public law scholarship about data has focused centrally on privacy rights and far less on the structural law of data—and, in particular, the forms of data governance our constitutional democracy requires as data rises in importance as a form of governmental power.

This Article develops an original account of data's structural law—the processes, institutional arrangements, transparency rules, and control mechanisms that, we argue, create distinctive structural dynamics for data's acquisition and appropriation to public projects. Doing so requires us to reconsider how law treats the category of power to which data belongs. Data is what we call an instrument of power*—the means (money, land, arms, and the like) that the government uses to accomplish its many ends. The Constitution, we argue, facilitates popular control over material forms of power like data through specific and distinctive strategies, ranging from defaults to accounting mechanisms. Assessing data's structural ecosystem against that backdrop allows us to both map the structural law of data and provide an initial diagnosis of its deficits.*

Drawing on our respective fields—law and computer science—we conclude by suggesting legal and technical pathways to asserting greater procedural, institutional, and popular control over the government's data. Indeed, we argue that data has distinctive structural possibilities because of the capacity to both channel and constrain data through technical design.

---

† Professor of Law, University of Chicago Law School.

†† Assistant Professor of Computer Science, University of Chicago. This Article has benefited from workshops at Harvard Law School, Northwestern Pritzker School of Law, the University of Chicago Law School, the University of Virginia School of Law, and Yale Law School, in addition to helpful comments from, and conversations with, Ian Ayres, Will Baude, Curt Bradley, Danielle Citron, Alex Hemmer, Aziz Huq, Alison LaCroix, David Strauss, David Weisbach, and Taisu Zhang. We finally thank the Neubauer Collegium and the University of Chicago Data Science Institute for their generous financial support.

INTRODUCTION

The central concern of structural constitutional law is the organization of governmental power.[1] But the forms of power available to the government change over time, and each form of power operates differently. This Article develops an account of how the Constitution—and the statutes, regulations, and norms that chart its structural ecosystem—organize an increasingly essential but understudied form of governmental power: the power government derives from data.

The government has always relied on information to meet its objectives—information about economic conditions; about enemies foreign and domestic; and about who its people are, where they reside, and how they behave.[2] But the digitization of information over the last half century has yielded a distinctive form of power, one that is liquid, accumulable, and easily transferred among governmental officials; minable and dynamic in capacity; vital to virtually all governmental activity; and a means by which policy goals both admirable and contestable can be achieved. As data has become a central driver of value in private markets, so too it has become an indispensable source of capacity and power in the public sector.

But the federal government's decision-making about what data to gather and how it should be used is often opaque and forced into public view only episodically—when confidential data collections are revealed, after dramatic shifts in Supreme Court doctrine, or when new uses of data are uncovered.[3] While policymaking about public health or immigration plays out in the halls of Congress or closely followed agency rulemakings, scholars have long lamented that policymaking about data is insulated from public cognizance and popular control, as sweeping data programs arise without clear congressional authorization or in ways contrary to written law—not just for national security and law enforcement purposes but, perhaps more surprisingly, across civil programs as well.[4]

---

1 *See* Marbury v. Madison, 5 U.S. (1 Cranch) 137, 176 (1803) ("[The Constitution] organizes the government, and assigns, to different departments, their respective powers.").

2 *See generally* JAMES C. SCOTT, SEEING LIKE A STATE (1998).

3 *See infra* Part I.A.

4 Data gathering and use has long occupied a legal gray area, as the Church Committee chronicled nearly fifty years ago. *See infra* note 153. Decades on, national security agencies continued to embark on large-scale surveillance efforts without congressional authorization. *E.g.*, James Risen & Eric Lichtblau, *Bush Lets U.S. Spy on Callers Without Courts*, N.Y.

The first months of President Donald Trump's second term brought a dramatic new episode in government data policy, as the President rechristened the sleepy United States Digital Service as the Department of Government Efficiency (DOGE) and made businessman Elon Musk its de facto leader.[5] The agency quickly sought access to dozens of the most powerful and sensitive federal databases without citation to legal authority and for purposes left largely undisclosed and undefended.[6] Even before his hundredth day in office, the President had expanded his unprecedented effort to assert White House control over government data by issuing an Executive Order instructing agencies across the government to compile, aggregate, and make available all unclassified data to the President and his designees.[7]

TIMES, (Dec. 16, 2005), https://www.nytimes.com/2005/12/16/politics/bush-lets-us-spy-on-callers-without-courts.html (describing the "special collection program" initiated after the September 11, 2001 terrorist attacks and designed to exceed the limits of the Foreign Intelligence Surveillance Act); James Risen & Eric Lichtblau, *How the U.S. Uses Technology to Mine More Data More Quickly*, N.Y. TIMES, (June 8, 2013), https://www.nytimes.com/2013/06/09/us/revelations-give-look-at-spy-agencys-wider-reach.html (observing that with "little public debate, the [National Security Agency] has been undergoing rapid expansion in order to exploit the mountains of new data being created each day"). Law enforcement also draws on a reservoir of implied and inherent powers, rather than express statutory structuring, for its most frequently used and significant databases. *See* Bridget A. Fahey, *Data Federalism*, 135 HARV. L. REV. 1007, 1032 (2022) [hereinafter Fahey, *Data Federalism*]. But even for civil purposes, the government draws on ambiguous legal authorities (or cites no authority at all) to acquire new data assets, develop novel data technologies, and allocate existing data to new purposes. *See infra* notes 163–64 (election data), 185–87 (judicial data), 235–37 (data sharing across civil and criminal agencies). Of particularly pressing importance today is the government-wide use of sensitive data to train and deploy artificial intelligence without express statutory authorization. *See infra* notes 220–24 and accompanying text.

5 *See* Exec. Order No. 14,158, 90 Fed. Reg. 8,441 (Jan. 20, 2025) (locating the newly named Department of Government Efficiency inside the Executive Office of the President); Madeleine Ngo & Theodore Schleifer, *How Trump's Department of Government Efficiency Will Work*, N.Y. TIMES (Jan. 21, 2025), https://www.nytimes.com/2025/01/21/us/politics/doge-government-efficiency-trump-musk.html (describing Musk's role in DOGE).

6 *See, e.g.*, Andrew Duehren & Cecilia Kang, *Struggle over Americans' Personal Data Plays Out Across the Government*, N.Y. TIMES (Feb. 19, 2025), https://www.nytimes.com/2025/02/19/us/politics/elon-musk-doge-personal-data.html; Laurel Wamsley, *The Government Already Knows a Lot About You. DOGE is Trying to Access All of It*, NPR (Mar. 11, 2025), https://perma.cc/A8EB-7HMM; Jonathan Swan, Theodore Schleifer, Maggie Haberman, Ryan Mac, Kate Conger, Nicholas Nehamas & Madeleine Ngo, *How Elon Musk Executed His Takeover of the Federal Bureaucracy*, N.Y. TIMES (Mar. 3, 2025), https://www.nytimes.com/2025/02/28/us/politics/musk-federal-bureaucracy-takeover.html; Hannah Natanson, Joseph Menn, Lisa Rein & Rachel Siegel, *DOGE Aims to Pool Federal Data, Putting Personal Information at Risk*, WASH. POST (May 7, 2025), https://www.washingtonpost.com/business/2025/05/07/doge-government-data-immigration-social-security/.

7 Stopping Waste, Fraud, and Abuse by Eliminating Information Silos, Exec. Order No. 14,243, § 3, 90 Fed. Reg. 13,681 (Mar. 20, 2025) (ordering agencies "to ensure Federal

These actions, like the data-related controversies that preceded them, raise significant privacy concerns. But they also raise structural questions about what data power the government lawfully possesses, which institutions hold or share that power, and how it can be exercised. For compliance with privacy rights cannot alone render data policy legitimate. Data, like any other form of governmental power, must also have the democratic validation that is the central concern of structural constitutional law.[8] Scholarly literatures on data, however, have focused disproportionately on questions of rights, and far less on questions of structure.[9] That data is a form of power that must be controlled by deliberately designed structures, processes, and institutions is a proposition familiar to the private sector, where the concept of "data governance" is the object of significant scholarly attention.[10] But public law

officials designated by the President or Agency Heads (or their designees) have full and prompt access to all unclassified agency records, data, software systems, and information technology systems . . . for purposes of pursuing Administration priorities related to the identification and elimination of waste, fraud, and abuse," including through the "intra- and inter-agency sharing and consolidation of unclassified agency records").

8 Although questions of data and structure have, we think, received less than their fair share of scholarly attention, many scholars have identified data problems that are structural in nature and on which we build. Scholars writing about national security and intelligence programs have highlighted the distinct deficiencies in political process that have characterized data gathering in those areas. *See, e.g.*, Barry Friedman & Danielle Keats Citron, *Indiscriminate Data Surveillance*, 110 VA. L. REV. 1351, 1373 (2024); Christopher Slobogin, *Panvasive Surveillance, Political Process Theory, and the Nondelegation Doctrine*, 102 GEO. L.J. 1721, 1758 (2014). Literatures on transparency have cast data disclosure from the government to the public as a tool of democratic structure. *See* David E. Pozen, *Transparency's Ideological Drift*, 128 YALE L.J. 100, 102–03 (2018) (collecting literature). Professors Aziz Huq and Zachary Clopton have interrogated the judiciary's constitutional power to collect and manage its data. Zachary D. Clopton & Aziz Z. Huq, *The Necessary and Proper Stewardship of Judicial Data*, 76 STAN. L. REV. 893, 928–31 (2024). Privacy literatures have also increasingly urged the use of structural design to protect data from commercial exploitation. *See, e.g.*, Daniel J. Solove, *The Limitations of Privacy Rights*, 98 NOTRE DAME L. REV. 975, 993 (2023) ("To be effective, control can't just be placed in the hands of individuals; control must come from society."); Aziz Z. Huq, *The Public Trust in Data*, 110 GEO. L.J. 333, 377–80 (2022); Salomé Viljoen, *A Relational Theory of Data Governance*, 131 YALE L.J. 573, 645–50 (2021) (arguing for the normative value of democratic control over public data and canvassing programs designed to use data for the public good). Finally, one of us has previously written about how data power has reshaped a different axis of structural constitutional law—federalism's vertical division of powers between the levels of government. *See* Fahey, *Data Federalism*, *supra* note 4, at 1024.

9 *See infra* Part I.B.

10 Writing about private sector data governance has long focused on how to identify and monetize private data assets. *See* Rene Abraham, Johannes Schneider & Jan vom Brocke, *Data Governance: A Conceptual Framework, Structured Review, and Research Agenda*, 49 INT'L J. INFO. MGMT. 424, 424 (2019) ("The purpose of data governance is to increase the value of data and minimize data-related cost and risk."). But data governance has also expanded to encompass structures that diffuse control over data to a wider range

scholarship has thought far less about what forms of data governance our constitutional democracy has and requires.

This Article provides a way to conceptualize, and a baseline account of, the structural law of data: how the Constitution allocates the power to collect and use data among the branches of the federal government; what forms of accountability and institutional organization the Constitution (and subconstitutional statutory law) uses to organize control over government data; and what avenues the public has to control data policy and prevent the misuses of power that data can enable. By contextualizing data and the power it confers in the normative concerns and legal strategies of structural constitutional law, we argue that data has systematically escaped the tools of popular control the Constitution charts for other, similarly consequential forms of governmental power.[11]

Because data is not typically conceptualized in a structural frame, Part I develops the case for thinking structurally about data. Focusing on data collected about individuals, we briefly sketch the varied and voluminous flows of data from individuals to the government, data's increasingly significant applications, and the ways that data can be misused not just by violating individual privacy rights but also in a structural sense—by being seized by unauthorized officials or branches; by being put to new uses contrary to legal authorization, popular preferences, or mandatory procedures; by being wasted or misappropriated; or by improperly favoring one social or political constituency at the expense of others.

We then offer a set of analytical tools for understanding and assessing data's current structural ecosystem. The standard analytical frames of structural constitutional law, we argue, are ill-suited to understanding how the power data confers on government

---

of stakeholders. Energetic debates over "information fiduciaries" and "data trusts," for example, imagine ways of rearranging corporate data governance to include representation for, and control by, data producers. *See, e.g.*, Huq, *supra* note 8, at 374–77; Lina M. Khan & David E. Pozen, *A Skeptical View of Information Fiduciaries*, 133 HARV. L. REV. 497, 507 (2019). And an emerging literature in data governance has articulated the normative importance of democratic engagement in data governance. *See, e.g.*, Viljoen, *supra* note 8, at 638–39; Amy Kapczynski, *The Law of Informational Capitalism*, 129 YALE L.J. 1460, 1499 (2020).

[11] Following the institutional turn in structural constitutional law, we begin with the Constitution, but we understand the law of structure to encompass the "authoritative legal norms that guide the workings of government and the distribution of government power"—whether in constitutional text, administrative statute, or durable institutional norm. *See* Jonathan S. Gould & David E. Pozen, *Structural Biases in Structural Constitutional Law*, 97 N.Y.U. L. REV. 59, 64 (2022).

is organized and locating the deficits in those arrangements. For when scholars and jurists write about how the Constitution organizes governmental power, the focus is most often on the government's legal capabilities—the President's power to conduct diplomacy, for example, or Congress's power to declare war—and how they are distributed and balanced across institutions.[12]

But the Constitution also thinks about power in a different and less scrutinized sense: It cares not just about the formal legal ends the branches may pursue but also about the material means they use to accomplish those ends.[13] The Constitution authorizes the use of, and organizes popular control over, what we call *instruments of power*: the money, lands, troops, armaments, and (we argue) data that compose the government's material capacity and that are the means through which its legal and political projects come to fruition.[14] The normative concerns of structural constitutional law, we argue, require as much control over the means of government as over the ends.[15] The Founding generation's effort to ensure popular control over means as well as ends is evident in, among other things, the elaborate structural controls that the Constitution establishes to govern fiscal power, charted by more than a dozen procedural and institutional mechanisms that ensure popular control over government funds.[16]

Still, instruments of power pose particular challenges for structural constitutional law because of their dual legal and material character. The government must have legal authority to obtain or use them, but it must also materially amass and steward them. They can, in turn, accumulate, be transferred, and be stockpiled in ways that simple legal authorities cannot. So too must they

---

12 *See* U.S. CONST. art. II, § 2, cl. 2 (diplomacy power); *id.* art. I, § 8, cl. 11 (power to declare war); *id.* art. III (judicial power); *Marbury*, 5 U.S. (1 Cranch) at 177.

13 Data and other instruments of power are not alone in escaping scrutiny. As Professor Daryl Levinson explained, "[F]or all the attention that issues relating to power have received in U.S. constitutional law, courts and theorists seem surprisingly at sea about basic questions of where power is located in the American political system, . . . [and] what 'power' means or which kinds of power should matter for different purposes." Daryl J. Levinson, *The Supreme Court 2015 Term—Foreword: Looking for Power in Public Law*, 130 HARV. L. REV. 31, 33 (2016) [hereinafter Levinson, *Looking for Power*].

14 *See* THE FEDERALIST NO. 58, at 359 (James Madison) (Clinton Rossiter ed., 1961) (calling the purse that "powerful instrument" by which the legislature expands its capacity and resists the "overgrown prerogatives of the other branches of the government"); *see also infra* Part I.D.3.

15 *See infra* Part I.E.

16 *See infra* Part II.

be secured, stored, and maintained. And they must—in their acquisition, appropriation, transfer, and use—be tracked, measured, and accounted for.[17]

Part II develops a framework for understanding how structural law facilitates public control over instruments of power in light of the challenges posed by this dual legal-material character, a baseline that we can use to consider whether analogous controls exist for data. Data is often compared to money—it is said to be a currency, a means of payment, an asset[18]—so to set this baseline, we look primarily to the elaborate and robust controls used to steward the government's fiscal power. Drawing on a range of sources—constitutional provisions, Supreme Court doctrine, structural statutes, regulations, and norms—we identify five structural approaches that the Constitution employs to address the distinctive problems that accompany the government's use of that power.

First, to protect against the temptation to amass and stockpile money, the Constitution uses default rules—and a restrictive default, in particular—to channel the distribution of funds through Congress and an annual appropriations cycle. Second, the Constitution subjects taxation and allocation decisions to generalist policymakers, who can account for and trade off uses across government, rather than empowering specialists to raise and fund their own initiatives. Third, because money, like all instruments, must be maintained, stored, and secured, the Constitution envisions specialized technical administration by creating, by name, the Treasury Department. Fourth, to address the need to shift funds around the government as they are put to lawful ends, structural law creates a regime of ordered movement, in which transfers of money are tightly, and centrally, controlled. Finally, the little-known Accounting Clause establishes a transparency mechanism specific to fiscal power (and its instrumental character) by requiring Congress to render an ongoing accounting, not simply disclosure, of its revenues and expenditures—a requirement that has been implemented through a structured system of fiscal tracking and public communication.

With that set of potential controls in mind, Part III maps data's structural law. The Constitution, of course, does not envision the kind of liquid, accumulable, and minable data that fuels

---

17 *See infra* notes 95–102 and accompanying text.

18 *See infra* note 80 and accompanying text.

contemporary governance, so it does not organize the power it confers on the government by name. But it would be wrong to assume that in the absence of express constitutional terms, data power has been left unstructured. The Constitution does address information—and, through it, data—and by considering together this set of "informational powers" and the data-related statutes, durable norms, and regulations that supplement it, we can better understand how public law structures government data acquisition and use.

We show that data's control mechanisms are impoverished relative to other instruments of power. Perhaps most importantly, the structural ecosystem for data is characterized by a permissive default. The Constitution's informational powers are unusually diffuse and redundant: Each branch has claimed a degree of inherent power to gather information, and by extension digital data, outside of express statutory authority, multiplying and easing pathways for the government to acquire and use individual data. But structuring also happens in the Constitution's negative spaces, through the statutes, regulations, and procedures that have arisen to steward the government's data assets absent meaningful constitutional constraint. We show that subconstitutional structural rules and norms in this context also permit the government greater latitude to acquire and use data power than it possesses with respect to other instruments. In effect, we argue, the branches and agencies of government have become accustomed to exercising broad authority to gather data unless specifically restricted by congressional statute or constitutional right.

Turning to the other strategies for control common to instruments of power, we likewise notice data exceptionalism. Decisions about the acquisition and allocation of data are made by specialist policymakers, including by procurement and IT officials deep within agencies. Data has decentralized technical administration, without oversight or stewardship by a focused agency that serves the function for data that the Treasury and Bureau of Land Management serve for money and land. Data generally moves between agencies and levels of government not by ordered central planning but through bilateral negotiation among individual government agents and agencies, as if it were a private asset rather than a public good. And data is only tentatively measured and accounted for, leaving the public with only intermittent knowledge of what data the government holds and the uses to which it has been put.

We conclude that the institutional and procedural management of data is diffuse rather than centralized, ad hoc rather than proceduralized, occluded rather than transparent, and—in final measure—lacking the structures designed for democratic control and used by other instruments of power. Absent structuring devices of this sort, we believe, important decisions about data will continue to be made outside public view and without popular superintendence.

Part IV charts a path forward. We begin with data's permissive default and the disproportionate role of specialists, rather than generalists, in acquiring and allocating government data. We are skeptical that data's permissive constitutional default will be amenable to dramatic doctrinal shifts, at least in the near term, so we focus instead on subconstitutional structural law. Against data's permissive constitutional backdrop, those structures will play an outsized role in improving data's popular control. To that end, we offer a series of statutory, regulatory, and norm-based recommendations for enhancing generalist oversight, in both Congress and the Executive Branch, over data policy.

But our account of data's structural apparatus also shows that law alone cannot mend its defects. To provide the public the knowledge that is a prerequisite to informed popular control, data must be rigorously and consistently accounted for—not in a general and summary form, but with precision and detail. And to be controllable by the public, data must move around government through systems that share and aggregate data with care, calibration, and transparency.

Informed by our respective fields of study—we are a constitutional law scholar and a computer scientist—we see these structural deficits as problems at the intersection of law and technical design. And we theorize that data systems, database design, and network architecture can perform some of the structural functions for data that constitutional and subconstitutional law perform for other instruments of power. First, we propose reimagining the concept of data provenance as a democratic transparency tool that can do for data what financial accounting does for money and surveying does for land: It can describe and disclose the value government derives from the data it collects about us. Second, we draw on research in data escrow systems to suggest a consistent and controllable infrastructure for sharing and regulating access to the government's data across agencies and projects. Finally, because these are new applications of emerging research and will

require future development, we also hope to chart a path for collaboration across our fields—collaboration that can draw the intersection of law and computer science beyond the already well-trodden domain of privacy rights into the new interdisciplinary space of data structure.

I. THINKING STRUCTURALLY ABOUT DATA

A. The Government's Data and Its Power

We tend to think about government data episodically—when intelligence failures prompt sweeping data policy reform,[19] when clandestine data-gathering initiatives come to light,[20] when the Supreme Court weighs in on novel data-collection techniques,[21] or when intergovernmental data disputes force voters to choose sides.[22] But the scope of data collection and the government's capabilities for extracting value from data have expanded rapidly, and it is worth pausing to think systemically about government data collection in its everyday observance, not just in its moments of breach.

There is no general accounting of the data held by the government, as we discuss in greater detail below. But there are indicators of the size and variation of the government's data stores. Over the last twenty-five years, for example, federal agencies have published over 6,200 "system of records notices," or SORNs, in the Federal Register.[23] Those notices are required when the government creates or modifies databases that hold records about

---

19 *See, e.g.*, NAT'L COMM'N ON TERRORIST ATTACKS UPON THE U.S., THE 9/11 COMMISSION REPORT, at xvi (2004).

20 *See, e.g.*, Leslie Cauley, *NSA Has Massive Database of Americans' Phone Calls*, USA TODAY (May 11, 2006), https://www.usatoday.com/story/money/2022/09/13/nsa-secretly-collecting-americans-phone-call-records/7940563001 (revealing the National Security Agency's large-scale domestic surveillance program).

21 *See, e.g.*, Carpenter v. United States, 138 S. Ct. 2206, 2223 (2018) (holding that warrantless collection of cell-site location information violates the Fourth Amendment).

22 For example, when "sanctuary cities" across the country sought to withhold immigration data and the federal government forced those claims into court. *See* Fahey, *Data Federalism*, *supra* note 4, at 1011, 1028; *see also id.* at 1011 nn.4–8 (collecting additional front-page stories about intergovernmental data disputes).

23 A search of the Federal Register for SORNs published since the year 2000 reveals over 6,250 SORNs and SORN revisions. *Document Search*, FED. REG. (updated daily), https://www.federalregister.gov/documents/search?conditions%5Bnotice_type%5D%5B%5D=sorn&conditions%5Bsearch_type_id%5D=6&conditions%5Bterm%5D=SORN&conditions%5Btype%5D%5B%5D=NOTICE.

individuals.[24] But even that number likely undercounts the government's databases by a substantial margin: Many data-rich areas, including intelligence, policing, and purely statistical efforts, are exempt from the SORN requirement.[25]

What kind of data do the government's databases contain? The government, of course, collects data to produce information about its population, its geographical distribution, its market participation, and its demographics.[26] It also requires individuals to turn over personal data about themselves to aid in program administration. Sometimes individuals are simply instructed to supply the relevant data—as most do annually on Forms W-2, 1040, and the like.[27] In other contexts, access to government services is conditioned on the transfer of information, as when individuals provide photographs to get a driver's license or a passport,[28] financial records to obtain student or small-business loans,[29] or biometric data to travel on airlines.[30]

But many public projects require data that cannot (or that the government would rather not) be provided directly by the data's subject. In these contexts, the government instead seeks data about individuals through intermediaries—from schools, banks, employers, health care providers, and more.[31] Those data transfers can be conspicuous and predictable to the data subject or, by equal measure, unexpected and obscured. As the COVID-19

---

[24] *See* 5 U.S.C. § 552a(a)(5) (defining a "system of records" as "a group of any records . . . from which information is retrieved by the name [or identifier] of the individual").

[25] *See* 5 U.S.C. § 552a(j)–(k).

[26] *See* ANTHONY GIDDENS, THE NATION-STATE AND VIOLENCE 180 (1985) ("The administrative power generated by the nation-state could not exist without the information base that is the means of its reflexive self-regulation.").

[27] *See generally Form 1040: U.S. Individual Income Tax Return*, U.S. INTERNAL REVENUE SERV. (2024), https://perma.cc/XNU9-TQNP.

[28] *See, e.g.*, ILL. ADMIN. CODE tit. 92, § 1030.90(a) (2025) ("Every driver's license issued . . . shall include, as an integral part of the license or card, a head and shoulder, full-faced color photograph of the [driver].").

[29] *See, e.g.*, System of Records Notice for the National Student Loan Data System, 88 Fed. Reg. 41,934, 41,936 (June 28, 2023) ("The information contained in this system is maintained . . . to determine the eligibility of aid applicants and recipients for Federal student financial aid programs.").

[30] *See* DEP'T OF HOMELAND SEC., PRIVACY IMPACT ASSESSMENT FOR THE TRAVEL DOCUMENT CHECKER AUTOMATION USING FACIAL IDENTIFICATION 2 (2022) (describing Credential Authentication Technology used to authenticate travel documents by conducting real-time facial scans); *see also* DEP'T OF HOMELAND SEC., PRIVACY IMPACT ASSESSMENT UPDATE FOR TSA ADVANCED IMAGING TECHNOLOGY 4 (2015).

[31] *See, e.g.*, Privacy Act of 1974, as Amended; System of Records Notice, 79 Fed. Reg. 20,969 (Apr. 14, 2014); Privacy Act of 1974; System of Records, 89 Fed. Reg. 7,381 (Feb. 2, 2024); Protection of Sensitive Information, 49 C.F.R. § 1520.5 (2024).

pandemic unfolded, for example, it would have come as no surprise to learn that health professionals shared disease and vaccination information with state and federal public health agencies.[32] But it might be more surprising to learn that employers share quarterly wage data on virtually all employees to a database administered not by taxing authorities, but by the Department of Health and Human Services, and used not for tax purposes, but to assist the administration of various programs ranging from child support to public benefits.[33]

State and local governments also act as data intermediaries, sharing information about their residents with the federal government, as one of us has shown.[34] The federal government collects significant amounts of data from states and cities for a wide range of public purposes—from immigration enforcement to the administration of public benefits programs to the validation of U.S. census numbers, to name only a few.[35]

Significant data also comes to the government not from voluntary disclosure by an individual or an intermediary but from the government's own ability to record and observe—as it does via surveillance cameras, license plate readers, and the like.[36]

Finally, we generally do not assume that government gathers the kind of behavioral and predictive data that private firms hold about our browsing habits, transit habits, viewing habits, purchase habits and the like. But even that form of data, too, has channels into government coffers. Government agencies increasingly purchase data, with little or no administrative process (a subject we explore further below), from private "data brokers"—a largely unregulated flow of digital information collected by private companies to government, which has been broadly documented in news reports but is rarely formally disclosed.[37] Moreover, local governments increasingly require or facilitate the

---

[32] *See Case Surveillance History*, CTRS. FOR DISEASE CONTROL & PREVENTION (Nov. 20, 2024), https://perma.cc/JL78-2QBU.

[33] *See* 42 U.S.C. § 653a (describing state databases); *see also id.* § 653(i).

[34] *See* Fahey, *Data Federalism*, *supra* note 4, at 1016–17.

[35] *See id.*

[36] *See* Barry Friedman, *Lawless Surveillance*, 97 N.Y.U. L. REV. 1143, 1148 (2022).

[37] *See, e.g.*, Byron Tau, *IRS Used Cellphone Location Data to Try to Find Suspects*, WALL ST. J. (June 19, 2020), https://www.wsj.com/articles/irs-used-cellphone-location-data-to-try-to-find-suspects-11592587815; NINA WANG, ALLISON MCDONALD, DANIEL BATEYKO & EMILY TUCKER, AMERICAN DRAGNET: DATA-DRIVEN DEPORTATION IN THE 21ST CENTURY 2 (2022) (describing U.S. Immigration and Customs Enforcement (ICE) purchases from private data brokers of "call records, child welfare records, credit headers, employment records, geolocation information, health care records, housing records and social media

transfer of large quantities of behavioral data—about transit habits, neighborhood activity, and even what happens inside homes—from individuals or private technology firms.[38] And as one of us has previously shown, there is a robust intergovernmental market for data, so it is reasonable to assume that data collected at the local level can and does migrate to other levels of government.[39]

B. The Inadequacy of Data Privacy Rights

It is intuitive to see individual rights, both constitutional and statutory, as the central safeguard against government misuse of these varied and extensive datastores. But despite the significant attention paid to them in academic literatures, judicial proceedings, and legislative chambers, privacy rights are an imperfectly implemented and inadequate tool for controlling the power government obtains from our data.

Constitutional rights etched in place in a predigital age have only haltingly been reimagined to confront the transformative ways data shapes government capacity. The Fourth Amendment's protection against "unreasonable searches and seizures,"[40] the Constitution's most express regulation of government acquisition of personal information, has placed guardrails on some forms of large-scale digital surveillance—including GPS monitoring and cell phone searches.[41] But its scope and doctrinal logic limit its influence over government use of personal data.

---

posts"); Orin Kerr, *Buying Data and the Fourth Amendment* 8 (Aegis Series Paper No. 2109, 2021); Chris Jay Hoofnagle, *Big Brother's Little Helpers: How ChoicePoint and Other Commercial Data Brokers Collect and Package Your Data for Law Enforcement*, 29 N.C. J. INT'L L. & COM. REG. 595, 599 (2004) (summarizing findings from a Freedom of Information Act (FOIA) lawsuit litigated by the Electronic Privacy Information Center); *see also* Matthew Tokson, *Government Purchases of Private Data*, 59 WAKE FOREST L. REV. 269, 283–88 (2024).

38 *See, e.g.*, Dave Lee, *US Police and Fire Departments Partnering with Amazon's Ring Passes 2,000*, FIN. TIMES (Jan. 29, 2021), https://www.ft.com/content/61968b3b-c093-4c4a-a7b7-29b565bc0bc0 (describing partnerships with Ring Doorbell to transfer doorbell surveillance to law enforcement); Los Angeles, Cal., Ordinance 185,785 (Sept. 13, 2018) (requiring e-scooter companies to disclose locational data on all rides as a condition of licensure); Naperville Smart Meter Awareness v. City of Naperville, 900 F.3d 521, 529 (7th Cir. 2018) (describing a public-private partnership to place smart meters in homes and share data with the city).

39 *See* Fahey, *Data Federalism*, *supra* note 4, at 1028–29.

40 U.S. CONST. amend. IV.

41 *See* Riley v. California, 573 U.S. 373, 403 (2014) (requiring a warrant for a cell phone search); *Carpenter*, 138 S. Ct. at 2218 (calling cell-site location data "near perfect surveillance" and finding it to be a search under the Fourth Amendment); United States v. Jones, 565 U.S. 400, 404 (2012). *But see* Orin S. Kerr, *The Mosaic Theory of the Fourth Amendment*, 111 MICH. L. REV. 311, 344 (2012).

Most obviously, the Fourth Amendment only regulates the government's collection of data, and it only applies to data gathered through "unreasonable searches." It does not apply to data the government obtains as a condition of receiving a benefit;[42] to much data obtained from third parties;[43] to data consensually offered to the government;[44] to data created by the government about individuals, like the Social Security number; to data supplied on government forms like tax returns;[45] to data shared across government agencies or between levels of government; or, for now, to data purchased from private data brokers.[46]

The application of the Fourth Amendment to contemporary data collection is also constrained in many contexts because it is doctrinally pegged to whether a search violates a "reasonable expectation of privacy."[47] Those expectations have been relentlessly dulled by aggressive data collection in consumer markets—including by conditioning the use of products on the surrender of personal data—and are thus, as Justice Sonia Sotomayor has argued, "ill suited" for use as a constitutional benchmark in "the digital age."[48] Most importantly, though, the Fourth Amendment does not govern what government does with data lawfully obtained.

---

[42] *See* Wyman v. James, 400 U.S. 309, 317–18 (1971) (finding a search incident to a public benefits program consistent with the Fourth Amendment because it was not "forced or compelled" and, in the alternative, because it was reasonable in light of government administrative interests).

[43] *See* United States v. Miller, 425 U.S. 435, 439 (1976); Orin S. Kerr, *The Case for the Third-Party Doctrine*, 107 MICH. L. REV. 561, 563 (2009).

[44] *See* Schneckloth v. Bustamonte, 412 U.S. 218, 219 (1973) ("[O]ne of the specifically established exceptions to the requirements of both a warrant and probable cause is a search that is conducted pursuant to consent.").

[45] *See* Flint v. Stone Tracy Co., 220 U.S. 107, 175 (1911) ("Certainly the [Fourth] Amendment was not intended to prevent the ordinary procedure . . . of requiring tax returns."); Cassano v. Carb, 436 F.3d 74, 75 (2d Cir. 2006) ("[T]he Constitution does not provide a right to privacy in one's [Social Security Number].").

[46] *See* Charlie Savage, *Intelligence Analysts Use U.S. Smartphone Location Data Without Warrants, Memo Says*, N.Y. TIMES (Jan. 25, 2021), https://www.nytimes.com/2021/01/22/us/politics/dia-surveillance-data.html ("[The Defense Intelligence Agency] does not construe the Carpenter decision to require a judicial warrant . . . [for] use of commercially available data for intelligence purposes." (quotation marks omitted)); Kerr, *supra* note 37, at 1 ("[E]xisting law leads to a clear answer: The government can buy business records without a warrant or any cause. The Fourth Amendment does not apply."). *But see* Tokson, *supra* note 37, at 288 (arguing the contrary and concluding that "Fourth Amendment law and the principles that undergird it require the government to obtain a warrant before purchasing private data").

[47] *See* Katz v. United States, 389 U.S. 347, 360 (1967) (Harlan, J., concurring).

[48] *See, e.g.*, *Jones*, 565 U.S. at 417 (Sotomayor, J., concurring).

Other constitutional rights, including general privacy rights,[49] property rights,[50] free expression rights,[51] and due process rights,[52] though theorized as tools for empowering individuals to assert control over their data, have yet to gain meaningful judicial embrace or operate as systemic constraints on government use of data. Nor has Congress enacted a statutory regime that applies comprehensively to the federal government's growing data stores—the equivalent, for instance, of the General Data Protection Regulation in Europe.[53]

One victory sometimes cited by advocates of data-related rights is the widespread adoption of the Fair Information Practices. First introduced by the Department of Health, Education, and Welfare in 1973, at the dawn of the transition into the digital age,

---

[49] *See, e.g.*, Griswold v. Connecticut, 381 U.S. 479, 484–85 (1965) ("Various guarantees [in the Constitution] create zones of privacy."); Whalen v. Roe, 429 U.S. 589, 605 (1977) (acknowledging the "threat to privacy implicit in the accumulation of vast amounts of personal information in computerized data banks"); Nixon v. Adm'r of Gen. Servs., 433 U.S. 425, 456–57 (1977); *see also* Lior Jacob Strahilevitz, *Reunifying Privacy Law*, 98 CALIF. L. REV. 2007, 2016 (2010) (cataloguing relevant law in the Courts of Appeals).

[50] Debates about the feasibility, and desirability, of extending property rights to individual data are as energetic and contested as ever. *See, e.g.*, James Toomey, *Property's Boundaries*, 109 VA. L. REV. 131, 186 (2023) (arguing that data cannot be "owned" in the conventional property sense); James Grimmelmann & Christina Mulligan, *Data Property*, 72 AM. U. L. REV. 829, 843 (2023) (arguing that data should be treated like tangible property); Paul M. Schwartz, *Property, Privacy, and Personal Data*, 117 HARV. L. REV. 2055, 2094 (2004); Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 STAN. L. REV. 1373, 1415, 1423 (2000) (rejecting a property right in data in favor of an "individual autonomy" right); Jessica Litman, *Information Privacy/Information Property*, 52 STAN. L. REV. 1283, 1289–92 (2000) (describing earlier debates).

[51] Data privacy rights and speech rights can be mutually reinforcing or mutually detracting. *Compare* NAACP v. Alabama *ex rel.* Patterson, 357 U.S. 449, 462 (1958) (noting that "compelled disclosure of affiliation with [advocacy] groups" can violate both individual privacy and the First Amendment right to free association), *with* Sorrell v. IMS Health Inc., 564 U.S. 552, 557–58 (2011) (invalidating a state consumer privacy law because it violated the First Amendment rights of data-owning firms); *see also* Eugene Volokh, *Freedom of Speech and Information Privacy: The Troubling Implications of a Right to Stop People from Speaking About You*, 52 STAN. L. REV. 1049, 1101–03 (2000).

[52] *See* Aziz Z. Huq, *Constitutional Rights in the Machine-Learning State*, 105 CORNELL L. REV. 1875, 1906 (2020) ("[D]ue process is violated when an algorithm fails to achieve an adequate level of accuracy."); Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 125 (2014) (arguing that due process should provide "those who may suffer from predictive privacy harms an opportunity to intervene in the predictive process"); Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1300 (2008).

[53] *See* Fahey, *Data Federalism*, *supra* note 4, at 1037 (describing the general approach and its area-specific exceptions, including the Health Insurance Portability and Accountability Act, which robustly safeguards health information and is an exception that, by its comparative muscle, proves the rule).

the Fair Information Practices are styled in the language of individual rights and outline best practices for data management by administrative agencies.[54] But in the United States, they are rarely given full-throated articulation or wholehearted enforcement.[55] And they ultimately take no position on the scale of government data gathering or the substantive objectives that can justify its pursuit so long as the individual has notice of the data's use.[56]

Thinking about the rights that individuals have with respect to the dazzling growth of the government's data stores is undoubtedly important. We should have a more robust understanding of when, and in what form, individuals can shield their data from government and what wrongs in data collection and use they are entitled to individually redress. But as other scholars have explained, privacy rights can be cost- and labor-intensive for individuals to assert and difficult for them to value—a necessary act in the contemporary data landscape because government collects so much data by asking individuals to voluntarily waive their rights.[57]

More importantly, individual rights by their nature can address only some of the problems—and draw into focus only some of the harms—that stem from the large, liquid, and dynamic data stores that now power virtually every governmental program. While rights establish the outer limits of governmental power by telling government what it cannot do, structure shapes how the government decides what it does with the power it retains.

---

54 SEC'Y'S ADVISORY COMM. ON AUTOMATED PERS. DATA SYS., RECORDS, COMPUTERS, AND THE RIGHTS OF CITIZENS 41 (1973). The Fair Information Practices recommend disclosing government databases, facilitating individual knowledge about their inclusion in databases, providing them with notice about the purpose of the data collection (and obtaining consent for any new use), allowing individuals to correct inaccuracies in their data, and guaranteeing reasonable data security. *Id.* at 53–63.

55 *See* Kenneth A. Bamberger & Deirdre K. Mulligan, *Privacy on the Books and on the Ground*, 63 STAN. L. REV. 247, 254 (2011) ("Legal academics and privacy experts have labeled the U.S. approach '[Fair Information Practices]-Lite,' an unfavorable comparison to the European Union." (quoting PRIV. RTS. CLEARINGHOUSE, PRIVACY TODAY: A REVIEW OF CURRENT ISSUES (2010)).

56 *See* Woodrow Hartzog & Neil Richards, *Privacy's Constitutional Moment and the Limits of Data Protection*, 61 B.C. L. REV. 1687, 1722 (2020) ("[Fair Information Practices]-based data protection is built around the idea that as long as data processing is fair to the data subject, the law should . . . create a legal structure to enable it.").

57 *See* Solove, *supra* note 8, at 985–88 ("People are not data scientists. They have trouble understanding the implications of their personal data at face value, let alone the downstream uses."); Alicia Solow-Niederman, *Information Privacy and the Inference Economy*, 117 NW. U. L. REV. 357, 360–61 (2022) (arguing that "individual control over personal information" is inadequate because government and firms use individual data to "generate further information about both those individuals and about other people"); Viljoen, *supra* note 8, at 598–600.

To illustrate the distinction between regulating data with rights, on the one hand, and with structure, on the other, consider a stylized example. Assume that a small number of individuals consensually volunteer extravagantly detailed behavioral information about themselves for use in crafting public regulations. Because that group conveyed its data voluntarily, their individual privacy rights have been respected. But the knowledge that information provides government not just about those individuals, but about all similarly situated people, might raise contentious questions about what we think the government should know about its polity.[58] It presents, in short, a *structural* question about what power government should be able to access and how it should be able to use it, not a *rights* question about a particular individual's privacy interests. No set of individual rights, we posit, no matter how rigorously enforced, can ensure that the government uses its properly obtained data in a way that is democratically responsive.[59]

The law of structure toggles the locus of control over government power from the individual to the group. It shifts the focus from how individuals can limit governmental activity when their own interests are directly affected to how groups—including the public writ large—can control the aggregated power government possesses.[60] The idea that controlling governmental power requires both rights and structure is an old one. But the idea that controlling the government's data power requires both rights and structure is not.

C. Data's Structural Problems

Government, of course, misuses data when it violates individuals' privacy rights. But compliance with even the most robust regime of rights does not by itself render data's use by the government legitimate. Because data is a source of governmental capacity and a form of governmental power, it must also be used in

---

58 Professor Salomé Viljoen has elegantly theorized this "relational" or "horizontal" dynamic in the data economy by showing the nuanced ways that one person's surrender of her data may confer knowledge and capacity on a firm (or government) that can have profound consequences for other individuals on whom the capacity is brought to bear. *See* Viljoen, *supra* note 8, at 580; *see also* Solow-Niederman, *supra* note 57, at 385–86 (collecting related ideas in the data privacy literature).

59 Structure can also, of course, help reinforce rights by constraining the government's ability to violate them ex ante rather than burdening the individual to police them ex post.

60 *See generally* CHRISTIAN LIST & PHILIP PETTIT, GROUP AGENCY: THE POSSIBILITY, DESIGN, AND STATUS OF CORPORATE AGENTS (2011).

ways that are constitutionally and democratically legitimate. Here we briefly sketch data's structural risks.

First, and most fundamentally, the law of structure in a constitutional democracy facilitates popular control over government policy. But decision-making structures can vary across policy areas. As data increasingly enables high-profile, and often controversial, government programs—in areas ranging from abortion[61] to gun regulation,[62] election integrity,[63] policing,[64] immigration,[65] national security,[66] and the training and use of artificial intelligence (AI) to perform governmental functions[67]—data's allocation to, or withholding from, those objectives may reflect, or may be contrary to, the policy preferences of the voting public. Voters may wish to scaffold the state's capacity to track women's reproductive choices or the mental health of gun owners—or they may wish, in Professor Daryl Levinson's terms, to "incapacitate" the state from accessing those forms of knowledge.[68] It is a basic structural

---

61 *See* Aziz Z. Huq & Rebecca Wexler, *Digital Privacy for Reproductive Choice in the Post*-Roe *Era*, 98 N.Y.U. L. REV. 555, 576–77 (2023) (documenting the "digital data trail" inevitably produced by an individual's reproductive choices and noting that states may enhance their enforcement capacity by "exploit[ing]" this data).

62 Federal law, for example, both creates and constrains data capacity for gun regulation by requiring background checks using the National Instant Criminal Background Check System for most gun purchases, 34 U.S.C. § 40901, while also prohibiting the creation of a "any [national] system of registration of firearms, firearms owners, or firearms transactions or dispositions," 18 U.S.C. § 926(a)(3).

63 Michael Morse, *Democracy's Bureaucracy: The Complicated Case of Voter Registration Lists*, 103 B.U. L. REV. 2123, 2145–60 (2023) (arguing that some data-related election fraud programs have been used to neutrally identify instances of voting fraud, like the Electronic Registration Information Center, while others, like the Crosscheck program, have amplified false claims of voting fraud).

64 *See, e.g.*, U.S. GOV'T ACCOUNTABILITY OFF., GAO-21-518, FACIAL RECOGNITION TECHNOLOGY: FEDERAL LAW ENFORCEMENT AGENCIES SHOULD BETTER ASSESS PRIVACY AND OTHER RISKS 15 (2021) [hereinafter GAO, FACIAL RECOGNITION TECHNOLOGY] (describing the FBI's controversial facial recognition database); Fahey, *Data Federalism*, *supra* note 4, at 1022 (describing the National Crime Information Center, "[l]ikely the nation's largest information pooling system").

65 *See, e.g.*, WANG ET AL., *supra* note 37, at 1 ("In its efforts to arrest and deport, ICE has—without any judicial, legislative or public oversight—reached into datasets containing personal information about the vast majority of people living in the U.S.").

66 There are sections of libraries dedicated to controversaries over national security surveillance. For a thoughtful summary, see generally RICHARD A. CLARK, MICHAEL J. MORELL, GEOFFREY R. STONE, CASS R. SUNSTEIN & PETER SWIRE, LIBERTY AND SECURITY IN A CHANGING WORLD: REPORT AND RECOMMENDATIONS OF THE PRESIDENT'S REVIEW GROUP ON INTELLIGENCE AND COMMUNICATIONS TECHNOLOGIES (2013).

67 *See infra* notes 209–14 and accompanying text.

68 *See* Daryl J. Levinson, *Incapacitating the State*, 56 WM. & MARY L. REV. 181, 182 (2014) [hereinafter Levinson, *Incapacitating the State*].

shortcoming if data policy is made outside of public view, without voter input, or contrary to voter preferences.[69]

Dynamics unique to data complicate this basic story still further by opening new avenues for governmental agents to depart from public preferences—introducing opportunities for what public choice literatures call "agency slack."[70] Because data can be easily duplicated and repurposed, for example, there is a heightened risk that data collected for one voter-authorized purpose—to verify local eligibility for public benefits, for example—may be shared with a different agency and put to a different, and voter-disfavored, purpose—to enforce immigration law, for instance.[71] And because data can be aggregated and transformed with powerful analytics, voters may authorize a data program because they assume data's capabilities are limited—that the data in question can be used only for simple record retrieval, for example—but find their authorization exploited to allow that same data to be repurposed with new technologies to substantially more powerful ends. The technical constraints that informed the original authorization may, in short, quickly become anachronistic. This was the case, for example, with databases collected for basic record retrieval that were used for data mining projects in the early 2000s.[72] And substantially the same dynamic is occurring today, as data collected for administrative projects has been repurposed to train and power machine learning tools.[73]

Second, because data is a form of power, it can be distributed among officials and institutions in ways that facilitate "the proper checks and balances between the different departments," as is the titled goal of the famous *Federalist No. 51*,[74] or in ways that do the opposite and dangerously concentrate governmental power.

---

69 *See* Slobogin, *supra* note 8, at 1750–51 (arguing that data policymaking is in some contexts insulated from popular control); Danielle Keats Citron & Frank Pasquale, *Network Accountability for the Domestic Intelligence Apparatus*, 62 HASTINGS L.J. 1441, 1459 (2011) (making a similar argument).

70 *See* Jide O. Nzelibe & Matthew C. Stephenson, *Complementary Constraints: Separation of Powers, Rational Voting, and Constitutional Design*, 123 HARV. L. REV. 617, 619 (2010) ("[M]ost advanced democracies rely on a combination of both electoral discipline and some form of internal separation of powers to reduce political 'agency slack' (the deviation between the behavior of political agents and what the voter-principals would prefer).").

71 *Cf.* California v. U.S. Dep't of Health & Hum. Servs., 2025 WL 2356224, at *1–2 (N.D. Cal. Aug. 12, 2025) (granting a preliminary injunction against a data sharing program between the Department of Health and Human Services' Center for Medicaid Services and ICE).

72 *See infra* note 198 and accompanying text.

73 *See infra* note 223 and accompanying text.

74 THE FEDERALIST NO. 51, at 322 (James Madison) (Clinton Rossiter ed., 1961).

Control over the government's data may be divided among branches of government to ensure that "[a]mbition . . . [is] made to counteract ambition" and no one branch can unilaterally deploy and use its power unimpeded.[75] So too a structural strategy for data must consider how the capacity to control governmental data enhances, impedes, and reorders other forms of power. Among the most striking episodes in the first few months of the Department of Government Efficiency was the agency's effort to access the Department of the Treasury's powerful payment's database.[76] Because the database was essential to the distribution of federal funds appropriated by Congress, the President's apparent strategy was to use Executive control over government data systems to obstruct Congress's constitutional prerogative to appropriate and expend the government's fiscal power.[77]

Third, once distributed among institutions of government, data—like other governmental assets—can be used inefficiently or ineffectively, causing waste and raising the costs of accomplishing desired policy objectives if not subjected to proper procedures and accountability mechanisms. Concerns about using data inefficiently may seem counterintuitive. Data is not generally regarded as a rival asset—the kind of asset, like money, for which each potential use trades off with every other potential use. Data is, in economic terms, often thought to be *nonrival*: Multiple parties can access it without depleting it.[78] In theory, then, we might care less about using data efficiently. We can always course correct, the intuition might go, without needing to refresh our data supply.

But that intuition, which may account for some of the shortage in structural thinking about preventing data's misuse, is too crude. Data often functions in public contexts as a limited asset, despite its nonrival characteristics. Data almost always, for instance, has associated interests that make its use costly rather than costless: Each act of access, duplication, or transformation

---

75 *See id.*

76 *See* Andrew Duehren, Maggie Haberman, Theodore Schleifer & Alan Rappeport, *Elon Musk's Team Now Has Access to Treasury's Payments System*, N.Y. TIMES (Feb. 1, 2025), https://www.nytimes.com/2025/02/01/us/politics/elon-musk-doge-federal-payments-system.html ("The system could give the Trump administration another mechanism to attempt to unilaterally restrict disbursement of money approved for specific purposes by Congress.").

77 *See* Bridget A. Fahey, *Musk's Madisonian Insight—And Its Troubling Consequences*, THE ATLANTIC (Mar. 13, 2025), https://www.theatlantic.com/ideas/archive/2025/03/doge-change-constitution/682019/.

78 Charles I. Jones & Christopher Tonetti, *Nonrivalry and the Economics of Data*, 110 AM. ECON. REV. 2819, 2822 (2020).

of personal data debits our tolerance for privacy and security risks. Data's use likewise has social consequences that should reduce the sense that the government is playing with a limitless resource. The U.S. Census Bureau, for example, has famously rich data, but the Bureau has long been concerned that sharing it might diminish public trust and willingness to participate in future census surveys.[79] That is, we may be able to duplicate current census data without depleting it, but the duplication itself impacts the Bureau's capacity to obtain future data of similar quality. Data also has temporal scarcity: It goes stale because its value depends on its accuracy, and its accuracy depends on its timeliness. Although data does not materially deplete in the sense of being used up, in other words, its value can depreciate over time.[80] Data is therefore not insulated from the concerns of waste that also animate the use of government money, land, and the like.

Fourth, and relatedly, data can be misappropriated and access to data abused—like public money that is spent for personal gain, public data can be applied to personal projects, as when an official uses data to shade the truth or propound misinformation designed to support his or her reelection or job security.[81]

Fifth, government decisions about data's acquisition and use can have profound distributional consequences. The government can choose to collect data from some people and not others, unevenly distributing the burdens of governmental knowledge. Or it can transfer data—and its value—to some groups while denying it to others.[82] Similarly, there are significant distributional effects in how the government uses its data. The government may use data collected about wages to police disadvantaged populations

---

79 ALEIA CLARK FOBIA, MIKELYN MEYERS, ARYN HERNANDEZ & LUCIA LYKKE, FINAL REPORT OF THE PRIVACY ACT COGNITIVE TESTING PROJECT 4 (2022).

80 *See* Laura Veldkamp, *Valuing Data as an Asset*, 27 REV. FIN. 1545, 1552 (2023).

81 For a more extensive discussion of the risks associated with government information production, see Bridget A. Fahey, Yuping Lin & Taisu Zhang, *The Law of Information States: Evidence from China and the United States*, 65 VA. J. INT'L L. 371, 379 (2025).

82 Professor Margaret Kwoka, for example, has argued that FOIA's legal design and administrative practice favor corporate requesters and represent, in effect, a transfer of government wealth to corporate interests. *See* Margaret B. Kwoka, *FOIA, Inc.*, 65 DUKE L.J. 1361, 1415 (2016); David E. Pozen, *Freedom of Information Beyond the Freedom of Information Act*, 165 U. PA. L. REV. 1097, 1103 (2017) (building on that argument); *see also* Christopher J. Morten & Amy Kapczynski, *The Big Data Regulator, Rebooted: Why and How the FDA Can and Should Disclose Confidential Data on Prescription Drugs and Vaccines*, 109 CALIF. L. REV. 493, 522 (2021) (making a similar claim in the Food and Drug Administration context); Clopton & Huq, *supra* note 8, at 920–21 (arguing that PACER, the primary system for judicial data, transfers value to private firms by erecting technical barriers to access that only sophisticated users can surmount).

but not to affirmatively recruit them to the social welfare programs for which they are eligible. Professor Michael Morse, for example, has argued that election data used today to chase down theories of voter fraud could also be used to help enfranchise hard-to-reach voters, but it generally isn't.[83] And Professor Andrew Crespo has shown that courts have used their significant data stores to improve "administrative efficiency," but not to develop the kind of "systemic facts" about the behavior of police and prosecutors that could "illuminat[e] important constitutional issues" for criminal defendants.[84] How the benefits and burdens of data power are distributed—like how the benefits and burdens of government's other material assets are distributed—should be disclosed to, and presumptively guided by, the public.

Finally, data decision-making can be, and often is, made invisible or illegible, preventing the public from forming views about it in the first instance. This is true of any form of governmental policy—Presidents can guide agency heads through back channels, for example, or Congress can obscure major policy shifts in cumbersome statutory language—but the risks associated with obfuscation are particularly acute for data policy. What data enables government to do can be inscrutable because the value of a particular data asset and the uses to which it can be put depend on many contingent factors, including its provenance, its likelihood of being productively combined with other data, and its use in conjunction with algorithmic models. Data's capabilities, moreover, are often concealed in technical systems and languages that require translation for the general public. And data custodians may be tempted to exploit the inscrutability of those systems to avoid accountability rather than inviting scrutiny by making them publicly legible.

D. Data and the Law of Structure

How should we gain entry into data's structural ecosystem—to consider whether the public can assert adequate control over the allocation of the government's data power and the structure of the government's data policy? A constitution's basic task is to

---

[83] Morse, *supra* note 63, at 2162.

[84] Andrew Crespo, *Systemic Facts: Toward Institutional Awareness in Criminal Courts*, 129 HARV. L. REV. 2049, 2108–09 (2016). Open access and interoperability requirements can likewise operate as a form of government-mandated data redistribution. *See* Dan Awrey & Joshua C. Macey, *Open Access, Interoperability, and DTCC's Unexpected Path to Monopoly*, 132 YALE L.J. 96, 161–67 (2022).

organize governmental power.[85] But power comes in many forms, and we use different structures to bring different forms of power under popular control. In this Section, we argue that in order to understand the power government gains from data and how that power is organized, we have to first consider the less scrutinized category of power—a category we call *instruments of power*—to which data belongs. To introduce that concept, we first review the somewhat more familiar forms of power the Constitution organizes.

1. Power as legal authorization.

Much of the doctrine and scholarship about structural constitutional law focuses on power that takes the form of *legal authority*. The Constitution empowers agents and institutions of government by legally sanctioning their performance of specified tasks. This is what we mean when we talk of the Constitution's "enumerated powers." The Constitution gives Congress the "Power to" regulate commerce, establish inferior courts, raise armies, enforce the Reconstruction Amendments, and so on.[86] It grants to the President the "Power to" form treaties, appoint officials, and grant pardons.[87] It extends the "judicial Power" to specified cases and controversies.[88] And it reserves to the states the "powers" not delegated to the federal government.[89] In each case, those provisions confer power by deeming the acts of governmental officials lawful under the described circumstances. An official is acting lawfully when she uses power assigned to her office by one of these provisions; she is acting ultra vires when she goes beyond them.[90]

2. Power as popular control.

The Constitution also organizes another form of power: the popular political power of different segments of the voting (and

---

85 Marbury v. Madison, 5 U.S. (1 Cranch) 137, 176 (1803) ("[The Constitution] organizes the government, and assigns, to different departments, their respective powers.").

86 *See, e.g.*, U.S. CONST. art. I, § 8, cls. 3, 8, 11; *id.* amend. XIV, § 5.

87 *Id.* art. II, § 2.

88 *Id.* art. III, § 2.

89 U.S. CONST. amend. X. The Constitution also confers power at a broader level of generality in the Vesting Clauses, which authorize Congress to deploy "legislative Powers," the President "executive Power," and the courts "judicial Power." *Id.* art. I, § 1; *id.* art. II, § 1; *id.* art. III, § 1.

90 For a selection of high-profile cases addressing that type of question, see Seila L. LLC v. Consumer Fin. Prot. Bureau, 140 S. Ct. 2183, 2211 (2020); Zivotofsky *ex rel.* Zivotofsky v. Kerry, 576 U.S. 1, 61 (2015); Nat'l Fed'n of Indep. Bus. v. Sebelius, 567 U.S. 519, 551 (2012); Immigr. & Naturalization Serv. v. Chadha, 462 U.S. 919, 955–59 (1983).

sometimes nonvoting) public *over* governmental institutions. This form of constitutionally organized power describes not the "power of the State," but the power of the people "*over* the State."[91] It gives effect to the basic idea that "[i]n a democracy, . . . it is important for the government as a whole to be controlled *by the people*."[92] To that end, the Constitution charters procedures and institutional arrangements that (at least in theory) allow the public—in different configurations[93]—to control how government officials act.[94]

3. Instruments of power.

Least often discussed is a third way the Constitution organizes power: It authorizes the government to obtain and structures the use of what we call *instruments of power*. Instruments of power are tools—the money, land, arms, troops, and (we argue) data that materially constitute and enable the use of other state powers.[95]

What is surprising about instruments of power (and their low profile in legal scholarship[96]) is the significant billing they get in

---

91 *See* Levinson, *Looking for Power*, *supra* note 13, at 45 (emphasis in original).

92 JEREMY WALDRON, *Constitutionalism: A Skeptical View*, *in* POLITICAL POLITICAL THEORY: ESSAYS ON INSTITUTIONS 23, 31 (2016) (emphasis in original).

93 *See* ROBERT A. DAHL, WHO GOVERNS? DEMOCRACY AND POWER IN AN AMERICAN CITY 247 (1961).

94 As Levinson has argued, doctrine and scholarship often elide these two forms of power by assuming, but not showing, a relationship between the distribution of legal authority among institutions of government and the distribution of power among the people who control government. *See* Levinson, *Looking for Power*, *supra* note 13, at 80–82.

95 Instruments of power are not assimilable into the social understandings of power common in conversations about democracy and constitutional law, like the view of power influentially advanced by Professor Robert Dahl that "*A* has power over *B* to the extent that he can get *B* to do something that *B* would not otherwise do." Robert A. Dahl, *The Concept of Power*, 2 BEHAV. SCI. 201, 202–03 (1957); *see also* STEVEN LUKES, POWER: A RADICAL VIEW 19–64 (2d ed. 2004) (canvassing variations on that theme). That sense of power describes a social relationship in which one person controls another. Instruments of power are a conceptually distinct way of capturing the ability to make change in the world: They speak not to power in the sense of control but to power in the sense of capacity. To that end, our democratic claim, *infra* Part I.E., is that to control an agent in the Dahlian sense, a principal must understand and be able to control the instruments that amplify the capacity of her agents.

96 This is perhaps related to what political scientist Francis Fukuyama has called the "strange absence of the state in political science." *See* Francis Fukuyama, *The Strange Absence of the State in Political Science*, THE AM. INT. (2012), https://perma.cc/J96M-GTKH ("[M]ost people are interested in studying political institutions that limit or check power . . . but very few people pay attention to the institution that accumulates and uses power"). Echoing this idea, Professor Jeremy Waldron has lamented the disproportionate attention

the Constitution itself. To see how the Constitution organizes access to instruments of power, consider fiscal power—its most elaborately structured instrument.[97] Twelve clauses in the Constitution address money directly.[98] The Constitution permits the government to acquire a supply of money and creates an intricate structural apparatus by which the resulting fiscal assets are stewarded—with specialized institutions, processes, and transparency techniques—as we discuss in greater detail in the next Part.[99]

The Constitution likewise contemplates that the government will obtain and manage land—another instrument that has been central to its many projects. The Constitution, for example, permits the acquisition of property and territory (by agreement or by force) as well as the exchange of land between the federal government and states, and subconstitutional structural law provides for the measurement, use, transfer, and disposal of public lands.[100] And the Constitution contemplates that the federal government will develop and procure other archetypical instruments

---

paid to how constitutions limit state power relative to their role in creating and guiding that power. WALDRON, *supra* note 92, at 29–30.

97 Fiscal power is also the instrument of power to which the most academic attention has been paid. Professor Kate Stith's 1988 article remains a foundational treatment of the "power of the purse." *See generally* Kate Stith, *Congress' Power of the Purse*, 97 YALE L.J. 1343 (1988). And a recent set of thoughtful articles likewise seek to "tak[e] appropriations seriously in public law doctrine." Gillian E. Metzger, *Taking Appropriations Seriously*, 121 COLUM. L. REV. 1075, 1083 (2021); *see also* Zachary S. Price, *Funding Restrictions and the Separation of Powers*, 71 VAND. L. REV. 357, 378–82 (2018); Richard Briffault, *Foreword: The Disfavored Constitution: State Fiscal Limits and State Constitutional Law*, 34 RUTGERS L.J. 907, 941–43 (2003).

98 U.S. CONST. art. I, § 6, cl. 1; *id.* § 7, cl. 1; *id.* § 8, cls. 1–2, 5–6; *id.* § 9, cls. 5–7, *id.* § 10, cls. 1–3.

99 *See* THE FEDERALIST NO. 58, *supra* note 14, at 359 ("Th[e] power over the purse, may . . . be regarded as the most complete and effectual weapon with which any constitution can arm the immediate representatives of the people, for obtaining a redress of every grievance, and for carrying into effect every just and salutary measure.").

100 The Territories Clause empowers Congress to "dispose of and make all needful Rules and Regulations respecting the Territory or other Property" of the United States. U.S. CONST. art. IV, § 3, cl. 2. *See generally* GREGORY ABLAVSKY, FEDERAL GROUND: GOVERNING PROPERTY AND VIOLENCE IN THE FIRST U.S. TERRITORIES (2021). Although the Clause does not specifically authorize the acquisition of new territory, it is now settled that the Territories Clause in conjunction with other congressional powers, like the power to make treaties, authorize Congress not only to manage existing territory but also to acquire new territory. *See* Sarah H. Cleveland, *Powers Inherent in Sovereignty: Indians, Aliens, Territories, and the Nineteenth Century Origins of Plenary Power over Foreign Affairs*, 81 TEX. L. REV. 1, 168–70 (2002) (describing the debate over Congress's power to conclude the Louisiana Purchase). The Enclave Clause, for its part, authorizes Congress to federalize state lands after obtaining state consent. U.S. CONST. art. I, § 8, cl. 17. The Admissions Clause permits Congress to add new states to the Union. U.S. CONST. art. IV, § 3, cl. 1. And Congress, to the chagrin of some, has long drawn from a bundle of express and implied

of power, including soldiers and armaments, by allowing Congress to "raise and support Armies," "provide and maintain a Navy," and assume control over state militias.[101]

What makes instruments of power unique is that they have both legal and material dimensions. Governmental agents must, of course, be legally authorized to obtain or use an instrument of power. But the instrument must also be possessed to be used.[102] And materiality introduces a host of other complications. Once procured, for instance, instruments of power can be transferred: They can be moved or their custody reassigned.[103] That allows them, in turn, to be combined, aggregated, and separated. Instruments must also be secured, stored, and maintained. Their value and the capacity they afford government can vary with a relevant market, with their suitability to a given task, by being reconfigured, or by combination with complementary powers. And to understand the capacity they afford government, they can—and must—be measured, tracked, and accounted for.

All of this means that instruments of power have distinct structural dynamics. To understand and regulate their use, we must have mechanisms of control adapted to their allocation, movement, and measurement—so that we understand precisely what instruments the government has amassed, what their value is, how that value changes over time, where they are located, who may deploy them, and to what ends. We need to know not just, for example, that the government could levy a tax, or that it has done so, but also what funds were raised, what their value is, how they have been allocated among governmental projects, and whether they were ultimately expended. Data, like money, can be moved and aggregated, mined and reengineered. In order to understand data's scale and value, we need a system that can adequately measure and track it not merely at the point of initial collection but also as it flows through government.

To take a step back, though: Because we rarely consider instruments of power as a category, we have no ready framework for thinking about how this form of power should be stewarded

---

powers to obtain new property through eminent domain. *See* William Baude, *Rethinking the Federal Eminent Domain Power*, 122 YALE L.J. 1738, 1745–46 (2013).

101 U.S. CONST. art. I, § 8, cls. 12–13.

102 Congress has legal authority to "establish . . . post Roads," but to bring those roadways to fruition, it also of course needs the land to site them and the money to build them. *See* U.S. CONST. art. I, § 8, cl. 7.

103 Land, for example, cannot be physically relocated, but its owner, proprietor, or ultimate sovereign can be shifted.

and rendered accountable, nor a vocabulary for thinking about the structural controls used to guide the government's use of them. We understand intuitively that these tools are a form of power, but we rarely stop to think—as the next Part begins to do—about the distinctive problems they present for popular control or how public law navigates them.

E. The Normative Case for Data Control

The final premise of our argument for thinking structurally about data is normative. The normative goals that animate the Constitution's structural choices are plural and contested, but virtually all normative lenses on structural constitutional law begin with a basic agency framework: The Constitution chartered the federal government as an agent of "the people," who are its principals and ultimate sovereign.[104] The primary goal of structural public law—constitutional, statutory, and regulatory—is to organize that agency relationship. The Constitution transfers power from the people to the government. It directs how that power will be distributed among different governmental actors. And it specifies processes—from formal legislative, executive, and administrative requirements to informal interbranch negotiations—that those actors must use to exercise it. Through those processes, in turn, the people may direct, limit, or veto particular uses of the powers conferred upon their government.

There are, of course, many different normative theories of governing. Each might advocate broader or narrower delegations of power from people to government, might press different allocations of power among the branches, might embrace different processes of popular control, and might emphasize different segments of the public in distinct ways.[105] We cannot, of course, resolve those debates here. Our goal for now is simply to establish a basic predicate proposition: The power the government gains from data is one of the forms of capacity that "the people" delegate to their government and one that requires popular control for its use to be understood as legitimate.

To see that argument more clearly, consider two counterarguments. First, perhaps it is best to think of all instruments of

---

[104] *See* M'Culloch v. Maryland, 17 U.S. (4 Wheat.) 316, 404–05 (1819) ("The government of the Union . . . is, emphatically, and truly, a government of the people. . . . Its powers are granted by them, and are to be exercised directly on them, and for their benefit.").

[105] *See* Levinson, *Looking for Power*, *supra* note 13, at 44–45.

power, including data, as subordinate forms of power—as simple tools used to advance the political or policy ends that should themselves be the central object of democratic control. We do not need popular control over *means*, the argument might go, if we have adequate popular control over the *ends* they advance.

But the Constitution's structural frameworks for creating and managing other instruments of power refutes that idea. As the next Part explains, the Constitution uses intricate and bespoke structural strategies to ensure that the public plays a direct role in authorizing and allocating the fiscal power at the government's disposal. Instruments of power, in other words, are not just capacity *enabling*, in that they allow the government to fully realize the use of its other powers; they are capacity *enhancing*, altering the character, reach, and form of state power and thus allowing the state to function in ways it otherwise could not. The size of the government's bank account, or its standing army, or (we think) its data stores shapes the Overton window within which government makes choices about what policies are possible—and appropriate—to pursue.

All of this means, in turn, that we must think specifically about how to control individual instruments of power, not merely the power that takes the form of legal authority or popular control. As Levinson has explained, "the more confidence we have in our ability to control the state, the more state capacity we will be willing to countenance."[106] That applies equally to legal and material forms of state capacity.

Second, perhaps (one could argue) only some instruments—the power of the purse, the power of the sword, and the power of land and territory, for example—require specific mechanisms of control to ensure that the government acts according to popular will. More mundane instruments—vehicles, office buildings, machinery, infrastructure, and perhaps data—are important, too, in making the government run and helping to constitute its material capacity, but each cannot practically be the object of specialized treatment. There do not need to be structural controls, that is, for trucks in the same way that there are for money.

We agree that not all instruments of power require detailed treatment. But we think the power of data is more like the power of the purse and the power of the sword than the power of trucks or buildings. And the reason, perhaps counterintuitively, is money.

---

[106] Levinson, *Incapacitating the State*, *supra* note 68, at 203.

Mechanisms for controlling the government's fiscal power also, by proxy, provide opportunities to control the instruments that government obtains *with* that power—the trucks it buys, the office space it rents, and the supplies it procures.

But money is a poor proxy through which to control data, as the raft of data-rich, money-poor start-ups achieving unicorn valuations illustrates. Fiscal appropriations to data projects are not only an inadequate proxy for the power government obtains from data; they are arguably a misleading one, given that government can obtain an immense value of data with little cost (or can purchase a data set at great cost that quickly becomes obsolete). Indeed, the government does not (for the most part) purchase data in private markets; it gathers data primarily from individuals. And the government can both gather and "expend" data without any acknowledgment in the federal budget. The power that government derives from data, moreover, changes across time and context as data is transformed, shared, exchanged, and combined with other complementary data—none of which are reflected in the data-related entries in the federal budget. We therefore need to see and control that data directly in order to steer the power it confers on government.[107]

II. A FRAMEWORK FOR CONTROLLING INSTRUMENTS OF POWER

How does law organize, discipline, and facilitate popular control over instruments of governmental power? To develop an initial framework of strategies, we mine the government's elaborate structural apparatuses for stewarding its fiscal power (and supplement that account, where relevant, with similar patterns for other archetypal instruments of power, including land, arms, and troops). We distill five structural strategies tailored to problems distinctive to money's instrumental form. They include restrictive defaults, standardized measurement, ordered movements, generalist acquisition and allocation, and specialized technical administration.[108]

---

[107] The same is true of money's relationship to other instruments of power. Money, for example, is also a poor proxy for the power government gains from land—a reason that land deserves, and for the most part has, its own tailored structural architectures. Knowing the market value of the government's sovereign or proprietary land simply does not tell us what the value *to the government* is of that land—which depends on its strategic location, fit with government functions, complementary plots, and legal status (whether the government exercises sovereign or only proprietary authority over it).

[108] Important strains of structural constitutional scholarship have focused on how structure shapes and is shaped by political parties, social and economic classes, and other interest groups. *See* Levinson, *Looking for Power*, *supra* note 13, at 82–83 ("The ultimate

Our claim is not that all instruments of power use (or ought to use) each of these control strategies in the same ways or to the same degree. For example, although we think that instruments of power require standardized accounting and measurement, the measurement techniques that are appropriate for money will be different than those appropriate to land, troops, or data. We cannot air all of those nuances here. Our goal instead is to understand how each of these strategies helps the public assert control over the features distinctive to instruments of power and to establish structural benchmarks that can, in turn, help us map and evaluate the components of data's structural ecosystem.

A. Restrictive Default

We begin with defaults—an aspect of structural design that is easily overlooked but that is essential to how the Constitution ensures control over instruments of power.

To understand the relevance of defaults to questions of government structure, consider the institutions and processes that are the traditional focus of academic attention in structural constitutional law. The goal of such work is to understand how institutional choices shape the decision-making of public officials. But background conditions also shape official behavior. It is often easier, for example, to sustain the status quo than to alter it.[109] And when officials must take action to advance a desired end, the ease or difficulty of that action also shapes the outcome. To study structural arrangements, then, we need to peer into the background to see the "choice architecture" within which government officials make decisions and exercise power—and to notice, drawing from contract law, the Constitution's legal "defaults" and their "altering rules."[110] As we show, defaults and altering rules are a central

---

holders of power in American democracy are not government institutions but democratic interests: the coalitions of policy-seeking political actors—voters, parties, officials, interest groups—that compete for control of these institutions and direct their decisionmaking."); Daryl J. Levinson & Richard H. Pildes, *Separation of Parties, Not Powers*, 119 HARV. L. REV. 2312, 2329–30 (2006); Gould & Pozen, *supra* note 11, at 126–27. Our analysis in this initial effort to understand the basic structural scaffolds of data's governance does not delineate among these competing power centers, but future work can, and should, interrogate how data's structural arrangements distribute political power among different democratic interest groups.

[109] *Cf.* Cass R. Sunstein, *Deciding by Default*, 162 U. PA. L. REV. 1, 17 (2013); Richard H. Thaler & Cass R. Sunstein, *Libertarian Paternalism*, 93 AM. ECON. REV. 175, 176 (2003).

[110] We draw here principally from the work of contract scholars, particularly Professors Ian Ayres and Robert Gertner. *See* Ian Ayres & Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 99 YALE L.J. 87, 87 (1989) ("Default rules

component of the Constitution's structural approach to controlling its fiscal power, and a central driver of the challenges in controlling the government's data power.

Consider the default rules applicable in the fiscal context. The Constitution's Appropriations Clause, which directs that "[n]o Money shall be drawn from the Treasury, but in Consequence of Appropriations made by Law," is commonly cited as the textual commitment of the power of the purse to Congress.[111] But seen through the lens of defaults, it does something more foundational than simply allocate power to Congress: It establishes a punishing default, and a burdensome altering rule, for use of the government's fiscal power.[112] Absent affirmative appropriations by Congress, the government is—by default—disabled from expending funds.[113] That applies to ordinary government programs and also to operations of the White House and federal courts, who cannot access the treasury even in service of their inherent Article II and Article III powers without appropriations.[114]

Because no part of government can draw upon the fisc without appropriations made by Congress, the Clause is perhaps best understood as what Professors Ian Ayres and Robert Gertner call a "penalty default"—a deliberately disagreeable status quo used to induce decision-makers to bear the costs necessary to modify

. . . govern unless the parties contract around them."); Ian Ayres, *Regulating Opt-Out: An Economic Theory of Altering Rules*, 121 YALE L.J. 2032, 2036 (2012) ("Altering rules are the necessary and sufficient conditions for displacing a legal default." (emphasis omitted)). Public law scholars have both identified the relevance of defaults to constitutional law, if sometimes by a different name, and urged their further study. *See* John Ferejohn & Barry Friedman, *Toward a Political Theory of Constitutional Default Rules*, 33 FLA. ST. U. L. REV. 825, 838 (2006) (arguing for the "nonoptional nature of constitutional default thinking"); Aziz Z. Huq, *The Negotiated Structural Constitution*, 114 COLUM. L. REV. 1595, 1647 (2014). *See generally* Gould & Pozen, *supra* note 11 (typologizing the many ways that "structural biases" shape the institutional backdrops against which American politics occurs).

111 U.S. CONST. art. I, § 9, cl. 7.

112 *See* Stith, *supra* note 97, at 1348.

113 U.S. House of Representatives v. Mnuchin, 976 F.3d 1, 11 (D.C. Cir. 2020), *vacated sub nom.* Yellen v. U.S. House of Representatives, 142 S. Ct. 332 (2021) ("Because the clause is phrased as a limitation, it means that 'the expenditure on public funds is proper only when authorized by Congress, not that public funds may be expended unless prohibited by Congress.'" (quoting United States v. MacCollom, 426 U.S. 317, 321 (1976) (plurality opinion))).

114 *See* Stith, *supra* note 97, at 1362 n.89 (explaining that "the President [lacks] constitutional authority to spend in the absence of appropriation," even if Congress "fail[ed] to provide funds for presidential activities"); *see also* Knote v. United States, 95 U.S. 149, 154 (1877) (holding that the President cannot use her pardon power to return funds paid into the Treasury absent appropriation). *But see* J. Gregory Sidak, *The President's Power of the Purse*, 1989 DUKE L.J. 1162, 1222 (arguing against Stith's conception of presidential powers).

it.[115] And it is a steep penalty: A default that can only be altered by law is costly indeed, given the demands of bicameralism and presentment and the supermajority requirement imposed by Senate rules.[116]

And the altering rule, too, is an imposing one: Money can be drawn from the fisc not by any law but only by an "[a]ppropriations" law.[117] Legislation generally need not take a specific substantive form—it can be express or implied, broad or narrow, specific or general. But appropriations are not so flexible. They cannot be implied by, or embedded in, broad conferrals of authority; the Treasury cannot release funds to an agency on the theory that because the agency has been charged with a statutory task, it must by implication have the resources to support it.[118] Instead, appropriations are operative only when an act of Congress "specifically states that an appropriation is made," a standard heavily policed by the Comptroller General.[119] Congress must thus reach agreement on precise appropriations language, a requirement that it has the option to avoid through ambiguity when enacting other forms of legislation. Moreover, although there is no requirement that appropriations expire each year, enough are made by convention on an annual basis to make the annual appropriations process both functionally necessary for the government and politically salient to voters.[120] Text, to be sure, does not always map onto practice, but the Appropriations Clause is notable for its longevity and continued relevance, as visible in the many statutes,

---

115 Ayres & Gertner, *supra* note 110, at 97.

116 *See* U.S. CONST. art I, § 7, cl. 2; STANDING RULES OF THE SENATE, S. DOC. NO. 113-18, at r. XXII (2013).

117 *See* U.S. CONST. art. I, § 9, cl. 7.

118 *See, e.g.*, Appropriations—Payments to Certain Counties in Oregon and Washington in Lieu of Taxes, 13 Comp. Gen. 77, 80 (1933) (explaining that appropriations can be identified by a "*specific* direction to pay and a designation of the funds to be used" (emphasis added)).

119 Act of Sept. 13, 1982, Pub. L. No. 97-258, § 1301(d), 96 Stat. 877, 917 (codified at 31 U.S.C. § 1301(d)); *see also* 31 U.S.C. § 701(2). As evidence of the fastidiousness of appropriations practice, the Comptroller General—the government's chief auditor, who also issues opinions on the interpretation of disputed appropriations provisions—chose not to correct even an obvious topographical error in an appropriations statute out of concern about interfering with Congress' appropriations authority. U.S. GOV'T ACCOUNTABILITY OFF., GAO-16-463-SP, PRINCIPLES OF FEDERAL APPROPRIATIONS LAW 1-31 (4th ed. 2016).

120 *See* Consumer Fin. Prot. Bureau v. Cmty. Fin. Servs. Ass'n of Am., 144 S. Ct. 1474, 1478 (2024) ("For most federal agencies, Congress provides funding on an annual basis. This annual process forces them to regularly implore Congress to fund their operations for the next year.").

regulations, and common law rules that reinforce its default and police its circumvention.[121]

How do this restrictive default and this burdensome altering rule shape the government's stewardship of fiscal power and opportunities for public control over it? These pressure Congress to annually negotiate, debate, publicize, and record votes on fiscal policy—practices that, given the power money confers on governmental actors, many representatives would avoid if they could by embedding funding in less salient legislative formats.

The restrictive default also serves an information-forcing function. Because the President and Congress have to make affirmative choices on an ongoing basis—and because those choices must achieve institutional concurrence—both branches must be able to develop detailed information about, and publicly digestible justifications for, their proposed allocations of federal funds. The Congressional Budget and Impoundment Control Act of 1974,[122] for example, created the Congressional Budget Office, a powerful nonpartisan office in Congress responsible for analyzing and "scoring" the fiscal impact of legislative proposals.[123] The House and Senate Budget Committees likewise reflect policymaking capacity forged through annual taxing and appropriations tussles.[124] Similarly, in the twentieth century, the Executive Branch

---

[121] The Antideficiency Act, for example, prohibits agencies from spending beyond their appropriations on credit in order to pressure Congress to later appropriate funds. *See* Pub. L. No. 97-258, § 1341, 96 Stat. 877, 923 (1982) (codified as amended at 31 U.S.C. § 1341(a)(1)(A)–(B)); Stith, *supra* note 97, at 1371–72. The common law rule disfavoring apparent authority for government agents likewise safeguards Congress's appropriation power by ensuring that government funds are not committed by agents not directly empowered by appropriations. *See* Fed. Crop Ins. Corp. v. Merrill, 332 U.S. 380, 383 (1947). The same goes for the rule disfavoring claims of equitable estoppel against government officials who make representations about government payments that prove legally incorrect. Off. of Pers. Mgmt. v. Richmond, 496 U.S. 414, 433 (1990) (arguing that recovery in estoppel, absent appropriation, could constitute "improper executive attempts to frustrate legislative policy").

[122] Pub. L. No. 93-344, 88 Stat. 297 (codified as amended in scattered sections of 2 and 31 U.S.C.).

[123] *See* Jesse M. Cross & Abbe R. Gluck, *The Congressional Bureaucracy*, 168 U. PA. L. REV. 1541, 1575 (2020) (describing the Congressional Budget Office as "a nonpartisan legislative office" that has "accreted authority" over time (citing Congressional Budget and Impoundment Control Act of 1974 § 201, 88 Stat. at 302–03)); Tim Westmoreland, *Standard Errors: How Budget Rules Distort Lawmaking*, 95 GEO. L.J. 1555, 1561 (2007); BARBARA L. SINCLAIR, UNORTHODOX LAWMAKING: NEW LEGISLATIVE PROCESSES IN THE U.S. CONGRESS 125–26 (1997).

[124] *See* Elizabeth Garrett, *The Congressional Budget Process: Strengthening the Party-in-Government*, 100 COLUM. L. REV. 702, 714 (2000) ("[A] major objective of the 1974 Act . . . was to provide an overarching structure that would coordinate a process overseen by dozens of appropriating, tax-writing, and some substantive committee.").

dramatically expanded its budget expertise. Until 1921, agencies submitted budget requests directly to Congress.[125] The Budget and Accounting Act of 1921[126] established the Bureau of Budget, the predecessor to today's Office of Management and Budget (OMB), to coordinate and collate the President's comprehensive budget proposal.[127]

Finally, the restrictive default also helps control the risks associated with the movement of money. Because an agency may not spend without appropriation, a sister agency with unspent funds cannot unilaterally transfer what it has left.[128] Those funds must be appropriated by generalists in Congress, as we discuss below, who can consider their many potential uses. The default serves a security function by creating a standardized script through which authorizations to access funds are expressed. And it has prompted an elaborate tracking and measurement system, as we explore further below: In order to pass appropriations bills, Congress must have a standardized way of comparing and accounting for its fiscal resources, past expenditures, and future uses across government.

In sum, the Constitution's restrictive default and onerous altering rule for fiscal power anchor its strategy for facilitating popular control over the government's money by hardwiring public-facing process into even basic fiscal decisions. The Constitution's Framers were acutely aware of the possibility of taxation and expenditure without representation, so it is no surprise that the Appropriations Clause is a particularly pointed example of how to use a constitutional default to force publicity and accountability in the management of instruments of power. But there are also significant examples of default-style thinking in structuring other instruments of power.[129]

---

125 OFF. OF MGMT. & BUDGET, OMB CIRCULAR NO. A-11, PREPARATION, SUBMISSION, AND EXECUTION OF THE BUDGET 15-1 (2024) ("[Before the 1921 Act] there was no annual centralized budgeting in the Executive Branch. Federal Government agencies usually sent budget requests independently to congressional committees.").

126 Pub. L. No. 67-13, 42 Stat. 20 (codified as amended in scattered sections of 31 U.S.C.).

127 *See* 31 U.S.C. §§ 1104–05 (directing the President to provide a unified budget); Reorganization Plan No. 2 of 1970, 35 Fed. Reg. 7,959, *reprinted in* 84 Stat. 2085 (1970) (relocating the Bureau of the Budget from the Treasury to the Executive office of the President).

128 *See infra* notes 143–44 and accompanying text (elaborating on how the government constrains the movement of fiscal assets).

129 Public lands are not subject to a constitutional Appropriations Clause, but the Property Clause allocates to Congress the "Power to dispose of and make all needful Rules and Regulations respecting the Territory and other Property belonging to the United States." U.S. CONST. art. IV, § 3, cl. 2. That Clause establishes a less restrictive default,

As we show in the next Part, however, data gathering and use are not subject to a restrictive default. Instead, the regime established to govern the acquisition and use of data is characterized by a *permissive* default, under which various branches of government can obtain and use data freely and of their own accord—without express statutory authorization or specific process, and even at times contrary to statutory restrictions.

B. Generalist Acquisition and Allocation

Because they have a material character, instruments of power must be acquired by the government in the first instance. And once held, they must be allocated—placed into the custody of specific officials and agencies, and allotted to specific projects. Decisions about acquisition and allocation could be vested in the outer perimeter of government—in the line-level administrators who interact with the public on the front lines. Managers of national parks could decide what fees are necessary to maintain those parks, then collect and expend them. Or decisions of this sort could be given to the heads of administrative agencies capable

---

but one that still requires significant legislative—and procedurally robust—engagement in the land management process. It was common in the nineteenth century, for example, for Congress to directly exercise its power under the Clause to dispose of federal land through "private bills," which named specific plots, counterparties, and prices, and at times themselves constituted contracts of sale. For examples, see 6 Stat. iii–xcix (collecting "The Private Acts of Congress" from 1789–1845). Congress has also exercised its authority to establish "Rules and Regulations" for federal land by delegating land management decisions to the four federal land management agencies. *See generally* CAROL HARDY VINCENT, LAURA B. COMAY, ERIC P. NARDI & ANNE A. RIDDLE, CONG. RSCH. SERV., RL34273, FEDERAL LAND OWNERSHIP: ACQUISITION AND DISPOSAL AUTHORITIES (2025). And before *Immigration and Naturalization Service v. Chadha*, 462 U.S. 919 (1983), invalidated the legislative veto, it was common for Congress to directly superintend those land decisions through the legislative veto. *See* Eugene R. Gaetke, *Separation of Powers, Legislative Vetoes, and the Public Lands*, 56 U. COLO. L. REV. 559, 560 (1985) (noting that the legislative veto had "become a prominent feature of public lands legislation" as Congress "frequently reserved the power to review and reject decisions made by agencies delegated the authority to manage the public lands"). So too the famous Posse Comitatus Act of 1878, ch. 263, § 15, 20 Stat. 145, 152 (codified as amended at 18 U.S.C. § 1385), legislatively restricted the default regarding another instrument of power: the troops that compose the standing army. The concern it addressed was the ease with which the President could allocate military forces to the project of domestic policing. The Act provided that military policing was prohibited unless "*expressly authorized* by the Constitution or Act of Congress." 18 U.S.C. § 1385 (emphasis added); *see also* JENNIFER K. ELSA, CONG. RSCH. SERV., R42659, THE POSSE COMITATUS ACT AND RELATED MATTERS: THE USE OF THE MILITARY TO EXECUTE CIVILIAN LAW 5 (2018) (noting that "[n]otwithstanding the founders' aversion to the use of a standing army to control the civilian populace, the Constitution nowhere explicitly prohibits it" and troops were used for that purpose on many occasions before the Posse Comitatus Act was enacted).

of assessing the agency's needs and pricing their fees accordingly. Decisions in those cases would be vested in *specialists*—agents selected for their subject-matter expertise. Acquisition and allocation decisions could, by contrast, be entrusted to *generalists*—actors like Congress or the President, who are accountable to voters through elections rather than appointed for their subject-matter expertise. That structural choice matters particularly for instruments of power, which have many potential uses across the government. Generalists are expected to weigh broad societal interests against one other and to make tradeoffs across a wide range of priorities. Specialists, by contrast, are usually expected to optimize within the area committed to their expertise.

The Constitution generally assigns the acquisition and allocation of instruments of power to generalists, not specialists. That decision reflects two basic judgments: (1) that the instruments at the government's disposal are public assets and should be allocated by representatives who can balance priorities across government (rather than specialists seeking to aggrandize their own role or project) and (2) that, as we have explained, because instruments expand governmental capacity, such capacity should be the people's to allocate.[130] Congress, for instance, is empowered to "dispose of and make all needful Rules and Regulations respecting the Territory or other Property belonging to the United States"; to negotiate land transfers with the states under the Enclave Clause; and to admit new states under the Admissions Clause.[131] And the Constitution assigns the power to "raise and support Armies," "provide and maintain a Navy," regulate the militia, and "declare War" to Congress and the President acting through bicameralism and presentment, and the power to command those resources to the President.[132]

The Constitution likewise assigns fiscal acquisition and allocation decisions to multiple generalist bodies by vesting those choices in Congress and subjecting them to the requirements of bicameralism and presentment.[133] Revenue-raising laws, moreover, must also comply with an added constraint: They must originate in the House of Representatives, the body that, in the view of

---

[130] *See supra* notes 103–05 and accompanying text.

[131] *See supra* note 100. In an interesting exception to that rule, the Constitution has been understood to permit Congress to delegate its eminent domain power within government and to private parties. *See* PennEast Pipeline Co. v. New Jersey, 141 S. Ct. 2244, 2247 (2021).

[132] U.S. CONST. art. I, § 8, cls. 11–16; art. II, § 2.

[133] *See* U.S. CONST. art. I, § 8, cl. 1; art. I, § 9, cl. 7.

James Madison, served as the most "immediate representatives of the people."[134]

Straightforward though that point may be in the fiscal, land, and military contexts, decisions about the acquisition and allocation of data—as we develop in the next Part—cut a striking contrast. An alternative system for fiscal power might have permitted officials to entrepreneurially collect their own revenue and fund their operations through devices like fees and asset forfeiture, as do many local agencies.[135] The Constitution rejects that structure, instead placing the power over the fisc in the body that most closely represents the people in order to empower the principal over the agent. But that path is, in broad strokes, exactly how the government satisfies its data needs. Data-gathering decisions are largely made by specialists—and, in particular, by the administrative agents who stand most to benefit from expanded data collection, without routine involvement from the representatives of those who stand most to lose.

C. Specialized Technical Administration

The material dimension of the government's instruments of power is most salient in judgments about how to administratively manage them—about their storage and measurement, how they will be secured, and who may access them. The Treasury Department gets prominent billing in the text of the Constitution itself. The Treasury oversees the processes of disbursing and accounting for the government's monetary assets,[136] and, with few exceptions, agencies that have been appropriated funds must access those

---

134 *See* THE FEDERALIST NO. 58, *supra* note 14, at 359 ("The house of representatives can not only refuse, but they alone can propose the supplies requisite for the support of government."); *see also* U.S. CONST. art. I, § 7, cl. 1 ("All Bills for raising Revenue shall originate in the House of Representatives."). The constraining effect of the Origination Clause is debatable. *See* Sissel v. U.S. Dep't of Health & Hum. Servs., 799 F.3d 1035, 1062 (D.C. Cir. 2015) (Kavanaugh, J., dissenting) (describing the Senate's workaround of "gut[ting] and replac[ing]" the contents of prior, unrelated House-originated bills).

135 By contrast, even fees collected by federal agencies are generally subject to the Appropriations Clause, though Congress can appropriate the value of those fees back to the agency either at the time it authorizes the fee collection or at a later time. *See* 1 U.S. GEN. ACCT. OFF., GAO-04-261SP, PRINCIPLES OF FEDERAL APPROPRIATION LAW 2-5 to 2-6 (3d ed. 2004).

136 *See* U.S. CONST. art. I, § 6 ("The Senators and Representatives shall receive a Compensation for their Services, to be ascertained by Law, and paid out of the Treasury."); *id.* § 9, cl. 7 ("No Money shall be drawn from the Treasury, but in Consequence of Appropriations made by Law."); *id.* § 10 (providing that state-laid duties "shall be for the Use of the Treasury of the United States"). The Federal Reserve Bank of New York hosts the Treasury General Account, the main bank account of the U.S. government.

funds through accounts stewarded by the Treasury.[137] Even programs like Social Security that earmark specific taxes for designated accounts hold the resulting "trust funds" in Treasury accounts.[138]

This custodial centralization serves several structural functions for the stewardship of federal funds. The Treasury creates salience. It is a single body, known to the rest of government and to the public as the entity responsible for ensuring that fund dispersals are authorized and responsible for any defects in the way funds are handled. This also allows the Treasury to perform the role of mediator among agencies. It is responsible for handling elaborate accounting and disclosure systems, discussed below. And it centralizes expertise: As the fiscal body responsible for stewarding all government funds, the Treasury has access to the staff and expertise to accomplish its legal obligations with regularity.[139]

By contrast, the administration of government data, as we explain below, is decentralized: Each agency manages its own data stores with little salience, virtually no accounting, disbursed expertise, and minimal efficiency.

D. Centrally Ordered Movements

Because of their material form, instruments of power can be accumulated, transferred, moved, and stored. Those activities, in turn, can shape their value and their use. Some instruments increase in value in a linear manner: Adding an extra dollar to a piggy bank increases the value contained in that bank by one dollar. The value of others can increase nonlinearly in combination.

---

[137] For a list of accounts, see *Federal Account Profiles*, USASPENDING.GOV, https://perma.cc/3VYT-U237. The Federal Reserve is a notable exception. Federal Reserve Act of 1913, Pub. L. No. 63-43, 38 Stat. 251 (codified as amended in scattered sections of 12 U.S.C.); *see also* 12 U.S.C. §§ 243–44 (funding the Federal Reserve through fees paid by Federal Reserve Banks). So is the Consumer Finance Protection Bureau (CFPB), whose funding mechanism was at the heart of a high-profile 2024 Supreme Court case. *See Cmty. Fin. Servs. Ass'n*, 144 S. Ct. at 1475. The CFPB is funded by Congressional appropriations, but the funds themselves come not from a Treasury account, but from "the combined earnings of the Federal Reserve System." *See* 12 U.S.C. § 5497(a)(1).

[138] *See What are the Trust Funds?*, SOC. SEC. ADMIN., https://perma.cc/54VP-ZXF5.

[139] The Bureau of Land Management serves many of the functions for public lands that the Treasury serves for public funds, managing 245 million acres of public land, including maintaining the General Land Office Records and Mineral and Land Records System. Three other agencies—the U.S. Fish and Wildlife Service, the National Park Service, and the U.S. Forest Service—also perform technical functions for particular types of federal land. *See* VINCENT ET AL., *supra* note 129, at 1.

Land and data are vivid examples of complementary goods: The value of two pieces of well-matched land or two bits of related data can be far greater than the independent value of each piece of land or data. The value of a complementary instrument of power, therefore, depends on its transferability—its capacity to be combined physically or custodially. Structure can facilitate ordered movement of assets in centrally managed ways, ensuring careful consideration of how to gather together complementary assets and what risks and capacity their combination brings to the institutions and agents who gain control over them. Or it can permit ad hoc movement in organic ways, letting discrete custodians decide whether and how to combine those assets in bilateral ways.

Fiscal resources generally flow only from the Treasury to the specific federal agency for which appropriations were made, and then only for the named purpose or project. The President cannot, for instance, decide that one program or agency is more important than another and transfer funds between those agencies unless Congress has so authorized. A high-profile dispute over the scope of one such authorization arose in 2020, when President Trump invoked a provision of the National Emergencies Act of 1976,[140] which allows the President to repurpose unobligated or leftover funds previously appropriated for military construction projects for unforeseen construction related to "war or . . . a national emergency," in order to divert money for constructing a wall along the United States' border with Mexico.[141] His gambit failed when the Ninth Circuit found the alleged "emergency" beyond the scope of the Act.[142]

Even at the mundane level, when funds move between agencies—including for ministerial purposes—it is because of express congressional authorization. When, for example, the Postal Service handles a package for the Environmental Protection Agency, the Federal Bureau of Investigation provides a background check for the State Department, or one agency compensates another for an employee on detail, such actions are pursuant to a transfer of fiscal assets that has been congressionally authorized.[143]

---

[140] Pub. L. No. 94-412, 90 Stat. 1255 (codified as amended at 50 U.S.C. § 1601 et seq.).

[141] *See* Sierra Club v. Trump, 977 F.3d 853, 862, 864 (9th Cir. 2020).

[142] *See id.* at 878 (holding that the transfer did not satisfy the "statute's criteria," and therefore "violates the explicit prohibition of the Appropriations Clause"); *see also Mnuchin*, 976 F.3d at 6, 11.

[143] *See* 31 U.S.C. §§ 1535–1536 (authorizing interagency service provision and reimbursement).

Fiscal transfers between the federal government and the states, for their part, must first pass through the appropriations process described above, then are also subject to an additional federalism-related structuring principle of state consent.[144]

By contrast, the movement of data throughout the government is fluid, negotiated, difficult to track, and disordered, as we set out below. Yet what fiscal power teaches is that aggregation and movement restraints are central to a structural strategy for controlling governmental power.

E. Standardized Measurement

One of the most significant—and perhaps most underappreciated—tools for organizing the government's instruments of power is measurement. Knowing how much money the government has, how many acres of land it controls, what armaments are in its arsenal, the number of troops in its uniforms, and the data in its databases is a basic building block of any regime of democratic accountability. Knowing only that the government has passed a law to tax earnings, gather data, raise an army, and buy (or take) land tells us only about *potential* governmental power. The law must be traced forward to the material instruments that it in fact produces. Only then can the public gain a measure of the actual power that tax, data collection, or draft has conferred on government. Those material instruments, in turn, must be tracked as they are stored, maintained, altered, transferred, and ultimately "appropriated." Only by understanding how an instrument of power is used can the public form a complete view of the value of that instrument to the government and, ultimately, the power the government has at its disposal.

Measuring the instruments in the government's possession is essential to allowing the public to evaluate its comfort with the capacity conferred by those instruments, determine the projects it wishes to enable through them, and decide which institutions it wants to control them. Most instruments of governmental power originate with the people—the people earn the money, generate the data, and are the troops in the government's cache. Only

---

144 *See* Printz v. United States, 521 U.S. 898, 936 (1997) (O'Connor, J., concurring) (indicating that although Congress may not commandeer the states, it can obtain their participation in joint programs through voluntary consent); *see also* New York v. United States, 505 U.S. 144, 166–68 (1992); Bridget A. Fahey, *Federalism by Contract*, 129 YALE L.J. 2326, 2339 (2020). The same is true for land. The Constitution requires state "[c]onsent" for the transfer of state land to federal sovereign control. U.S. CONST. art. I, § 8, cl. 17.

by knowing what assets the government has obtained from them and what public purposes it has pursued with those assets can the people weigh the costs they bear against the benefits they obtain.

The Constitution recognizes that fact with respect to fiscal resources in the Accounting Clause, a rarely discussed provision that requires Congress to make "a regular Statement and Account of the Receipts and Expenditures of all public Money . . . from time to time."[145] Although many structural provisions of the Constitution have fallen into desuetude, the Accounting Clause has not: Congress, the Treasury, the Government Accountability Office, and the Office of Management and Budget together maintain and disclose essentially continuous statements of account for the government's fiscal assets and their uses.[146] The public can access that information in a central and standardized form—the Government Standard General Ledger, a government-wide balance sheet—rather than in variable and ad hoc expressions that require the observer to conduct her own normalization and synthesis.[147]

In a time of notable governmental dysfunction, the transparency regime this clause has established, and the window it has created into the government's fiscal power, is notable.[148] Of course, our accounting and disclosure approaches, like any aspect of government, could be improved, as scholars like Howell Jackson have suggested.[149] But we have nevertheless reached a remarkable degree of fiscal transparency and regularity, anchored by the Accounting Clause and reinforced by the information-forcing functions of each aspect of the structural ecosystem described above. Likewise, our robust tradition of land measurement and surveying illustrates a near obsession with understanding what assets are at the federal government's disposal and how they are being used.[150]

---

145 U.S. CONST. art. I, § 9, cl. 7.

146 *See, e.g.*, *Daily Treasury Statement*, FISCALDATA (updated daily), https://fiscaldata.treasury.gov/datasets/daily-treasury-statement/operating-cash-balance.

147 *Governmentwide Accounting*, BUREAU OF THE FISCAL SERV. (last updated Apr. 15, 2024), https://perma.cc/XN2U-FBQS; *see also 1 Treasury Financial Manual, Part 2: Central Accounting and Reporting*, TREASURY FIN. EXPERIENCE, https://perma.cc/2AW5-MMDH.

148 There are exceptions, and they are significant—the foreign intelligence agencies get lump sum appropriations, for example—but for the most part, the public knows what the government spends.

149 *See generally* Howell E. Jackson, *Counting the Ways: The Structure of Federal Spending*, *in* FISCAL CHALLENGES: AN INTERDISCIPLINARY APPROACH TO BUDGET POLICY 185 (Elizabeth Garrett et al. eds., 2008).

150 Indeed, our intricate public system of land surveying predates the Constitution itself, tracing its lineage to the Land Ordinance of 1785 enacted by the Confederation

III. DATA'S STRUCTURAL LAW

It is common, when sweeping data-gathering initiatives or novel data uses are publicly revealed—often in splashy news articles—for scholars to lament that data policy is so frequently made through confidential channels outside the ordinary political process.[151] What we lack is a detailed and systemic account that explains why. Using the framework developed above, this Part provides an original account of data's structural ecosystem: the defaults, transparency rules, movement controls, policymaking process, and institutions of technical administration that are used to set data policy and steward the government's data assets.

This structural ecosystem is largely *not* the product of deliberate constitutional or statutory design but of implied and adapted powers, claimed institutional prerogatives, departmental practice, and incidental statutory effects. It is accreted, not calculated. But it *is* distinctive to data. And it has yielded a structure better calibrated to motivated actors exploiting data's capacities than to a deliberative and considered form of popular control and accountability over the power data confers on government.

We begin by excavating data's structural default rules. Whereas the Constitution restricts access to the government's fiscal power except when pursuant to the burdensome altering rule of appropriations statute, constitutional law and practice have established a far more *permissive* default for government data. Because the power government gains from its large and growing stores of digitized data was not contemplated by the Constitution's framers—and because the constitutional register in which we think about data is so often that of privacy rights not structure—identifying how the Constitution organizes data power requires something of an investigative effort.

We cannot expect the Constitution's text to explicitly configure control over the government's data power. But neither is data's power guided in an entirely subconstitutional register by ordinary statutes and administrative law. Data is a form of information, and the Constitution can be understood to authorize and allocate among the branches of government a set of what we call *information powers*. By drawing them together, we describe a

---

Congress. *See* 27 JOURNALS OF THE CONTINENTAL CONGRESS, 1774–1789, at 446 (Gaillard Hunt ed., 1928). The Land Act of 1796, in turn, laid the foundation for a sprawling land measurement and management bureaucracy under the auspices of the "Surveyor General," as historian Gregory Ablavsky has detailed. ABLAVSKY, *supra* note 100, at 77–78.

151 *See generally, e.g.*, Friedman & Citron, *supra* note 8; WANG ET AL., *supra* note 37.

fluid set of constitutional claims made by each branch to obtain and expend public data unless Congress has explicitly restricted them, leaving data power to expand without the clear channeling of a provision like the Appropriations Clause.

The remaining aspects of data's structural ecosystem also depart from the structural controls that accompany the use of fiscal power. Policy around data acquisition and allocation is controlled by *specialists*, not generalists; data's technical administration is *decentralized*, not subject to the oversight of centralized decisionmakers; it moves in crude and *disordered* ways among branches and levels of government; and it is measured only *tentatively*, using crude and muddled strategies that fail to account for its scale and value to the government or even chart its basic uses over time. Thinking structurally about data, in short, reveals that the structural controls that attend one of the government's central sources of power do little to constrain the use of that power or to ensure popular control over it.

A. Permissive Default

In contrast to the restrictive default and burdensome altering rule embedded in the Appropriations Clause, the power to gather and allocate the government's data does not rest solely with Congress and does not always require enactment of laws via bicameralism and presentment (much less laws with specific statutory language).

Rather, as we show, constitutional text, departmental practice, and judicial doctrine have together multiplied the power to gather information—and by extension digital data—by lodging it in all three branches of the federal government (and, for that matter, in each house of Congress acting independently). Existing constitutional settlement does not resolve the precise contours of those powers, but each body has asserted at least some *inherent* data-gathering power (which can be operationalized without statutory authorization), and some have asserted *exclusive* data-gathering power (which is insulated entirely from statutory regulation). Data acquisition and use, in short, has an exceptional political process that is more permissive than ordinary lawmaking and far more permissive than that of fiscal appropriations.[152]

---

[152] To be clear, our argument here is not that this constitutional regime is necessarily inappropriate; there are justifications for, as well as drawbacks to, the Constitution's permissive informational powers that have significance beyond conversations about data.

Data's exceptionalism extends to administrative agencies as well. These agencies have come to draw on broad delegations of authority—many of which predate the digital age—to justify data gathering, data use, and data programs big and small. We show, moreover, that an outdated privacy law has the perverse effect of exempting administrative data acquisition and allocation from the form of public control standard in federal administrative agencies: notice-and-comment rulemaking, and its burdens of reasoned justification and judicial review. Taken together, in other words, although the federal government is broadly prohibited from raising and spending money unless *specifically authorized* by Congress, its many component parts have asserted significant flexibility to acquire and use data unless *prohibited* by Congress (and sometimes contrary to Congressional guidance).

1. The President.

The President has a long history of gathering data without congressional authorization for policing and national security purposes—asserting an inherent and, at times, congressionally unregulable authority to acquire information and data at a large scale.[153] Just after the September 11, 2001, terrorist attacks, for example, President George W. Bush confidentially ordered the warrantless surveillance of foreign targets and their domestic interlocutors, contrary to the Foreign Intelligence Surveillance Act. As Deputy Assistant Attorney General John Yoo's well-known

What is significant for our purposes is the hardwired deficit in mechanisms for popular control over data that it creates—and against which the administrative structures that Congress has devised for data must be conceptualized.

153 United States v. U.S. Dist. Ct. for the E. Dist. of Mich., 407 U.S. 297, 299 (1972) ("Successive Presidents for more than one-quarter of a century have authorized [electronic] surveillance [in internal security matters] . . . without guidance from the Congress."). The report from the Senate's celebrated Church Committee—which was charged with reviewing decades of confidential information-collection initiatives after several were publicly revealed in early 1970s—catalogues the claims of inherent presidential authority used to justify those and other mass data-gathering programs. *See* 2 FINAL REPORT OF THE SELECT COMMITTEE TO STUDY GOVERNMENTAL OPERATIONS WITH RESPECT TO INTELLIGENCE ACTIVITIES, S. REP. NO. 94-755, 36–37 (1976) (describing President Franklin D. Roosevelt's reliance on inherent powers to allow FBI wiretapping contrary to the Federal Communications Act of 1934); *id.* at 70 (describing the FBI's effort to "collect domestic intelligence under sweeping authorizations issued by the Justice Department in 1974 for investigations of 'subversives,' potential civil disturbances, and 'potential crimes,'" which were "explicitly based on broad theories of inherent executive power"); *id.* at 129–30 (describing efforts by FBI Director Clarence M. Kelly, successor to J. Edgar Hoover, to "urge[ ]" the President to use his inherent information gathering authority to permit certain FBI surveillance activities).

memorandum defending that action argued: "[T]he executive branch possesses the inherent constitutional power to conduct warrantless searches for national security purposes," a power "consistently asserted[ ] and exercised" by presidents and endorsed by the Office of Legal Counsel "across different administrations."[154]

But assertions of the President's inherent data-gathering power also extend beyond the national security context. President Trump's Department of Government Efficiency sought access to many different federal databases—structured and regulated in distinct ways by congressional statutes.[155] The legal claim appeared to be not that those statutes authorized DOGE's access but that the President had an inherent authority to learn what was in the databases held by federal agencies—however Congress substantively or procedurally organized their use.

So too Executive Order 14,243—which instructs agencies to provide Presidential designees with "full and prompt access to all unclassified agency records, data, software systems, and information technology systems" to advance "Administration priorities related to the identification and elimination of waste, fraud, and abuse"—appears to be implicitly constitutionally grounded in a theory that the President holds inherent power to view and use any data held by administrative agencies, though the legal theory on which the order is grounded has not been publicly described or defended.[156]

During his first term, President Trump likewise acted as if he had inherent power to gather and hold data outside the national security sector when he established the Presidential Advisory Commission on Election Integrity, which sought hundreds of millions of voter files from the states in order to assist the White

---

[154] John C. Yoo, *Memorandum for the Attorney General, Re: Constitutionality of Expanded Electronic Surveillance Techniques Against Terrorists*, U.S. DEP'T OF JUST. 9 (Nov. 2, 2001), https://www.justice.gov/olc/page/file/1154156/dl; *see also* John C. Yoo, *Memorandum for the Attorney General, Re: Review of the Legality of the STELLAR WIND Program*, U.S. DEP'T OF JUST. (May 6, 2004), https://www.justice.gov/sites/default/files/pages/attachments/2014/09/19/may_6_2004_goldsmith_opinion.pdf (revising the original memorandum).

[155] *See, e.g.*, Duehren et al., *supra* note 76; Nicholas Nehamas, *DOGE Seeks Access to Social Security Data*, N.Y. TIMES (Aug. 26, 2025), https://www.nytimes.com/2025/08/26/us/politics/doge-social-security-data.html.

[156] Exec. Order No. 14,243, § 3, 90 Fed. Reg. 13,681 (Mar. 20, 2025). Combatting "waste, fraud, and abuse" is not (or at least not obviously) an inherent Article II power, suggesting that the President's power to use the vast quantities of data held by federal agencies is grounded in an informational power over the data itself, rather than an ends-based power over the President's specific proposed uses of that data.

House in evaluating claims of voter fraud.[157] The announcement prompted a flurry of lawsuits[158] challenging the Commission's compliance with the Administrative Procedure Act,[159] the Freedom of Information Act[160], the Privacy Act of 1974,[161] the E-Government Act,[162] and the Federal Advisory Committee Act.[163] What is notable for our purposes is that so few of those lawsuits, and timidly at that, contested the President's *constitutional* authority to solicit data without congressional authorization, and that claim was never addressed by a court.[164]

When the President acts without congressional delegation, her actions must generally be drawn from inherent executive powers.[165] And if she acts contrary to congressional instruction, she must show that her constitutional authority in that area is exclusive—that Congress has no right to constrain her.[166] What sources of constitutional authority might empower the President to unilaterally (or exclusively) collect or control data? The answer has not been settled by constitutional practice or doctrine.[167]

---

157 *See* Exec. Order No. 13,799, 82 Fed. Reg. 22,389 (May 11, 2017).

158 *See, e.g.*, Laws.' Comm. for C.R. Under L. v. Presidential Advisory Comm'n on Election Integrity, 316 F. Supp. 3d 230, 231 (D.D.C. 2018); Elec. Priv. Info. Ctr. v. Presidential Advisory Comm'n on Election Integrity, 266 F. Supp. 3d 297, 315 (D.D.C. 2017), *aff'd on other grounds*, 878 F.3d 371 (D.C. Cir. 2017); Dunlap v. Presidential Advisory Comm'n on Election Integrity, 464 F. Supp. 3d 247, 252 (D.D.C. 2020); Joyner v. Presidential Advisory Comm'n on Election Integrity, 2018 WL 4776089, at *1 (S.D. Fla. May 30, 2018).

159 Pub. L. No. 79-404, 60 Stat. 237 (1946) (codified as amended in scattered sections of 5 U.S.C.).

160 Pub. L. No. 89-487, 80 Stat. 250 (1966) (codified as amended at 5 U.S.C. § 552).

161 Pub. L. No. 93-579, 88 Stat. 1896 (codified as amended at 5 U.S.C. § 552a).

162 Pub. L. No. 107-347, 116 Stat. 2899 (2002) (codified as amended in scattered sections of 44 U.S.C.).

163 Pub. L. No. 92-463, 86 Stat. 770 (1972).

164 *See Joyner*, 2018 WL 4776089, at *2.

165 Youngstown Sheet & Tube Co. v. Sawyer, 343 U.S. 579, 637–38 (1952) (Jackson, J., concurring).

166 *See id.* at 638.

167 *See* Zweibon v. Mitchell, 516 F.2d 594, 616 (D.C. Cir. 1975) (explaining that the Supreme Court regards "long-standing Executive practice of conducting surveillance 'in cases vitally affecting the domestic security' as indicative of the unchallenged Executive power to obtain intelligence information" but "not as determinative of the proper procedures to be followed in so doing" (quoting *U.S. Dist. Ct.*, 407 U.S at 310)); *id.* at 614 (declining to "address the substantive scope of that power"); United States v. Truong Dinh Hung, 629 F.2d 908, 914 (4th Cir. 1980) (acknowledging "the principal responsibility of the President for foreign affairs and concomitantly for foreign intelligence surveillance"); *In re* Sealed Case, 310 F.3d 717, 740–42 (FISA Ct. Rev. 2002) (cataloging case law). Standing doctrine makes precise resolution of these structural claims challenging. *See generally* Clapper v. Amnesty Int'l USA, 568 U.S. 398 (2013).

The President has generally claimed data gathering and use powers connected to specific Article II authorities.[168] And it would be difficult to dispute that the President has the power to gather information about the deployment of foreign forces as an adjunct to her powers as Commander in Chief, for example.[169] But each such power needs to be established through a context-specific inquiry focused on the nature of the presidential power in question and its relationship to Congress's coordinate powers—especially so if the claimed executive power is insulated from congressional regulation.[170] Only exclusive executive powers can have exclusive attendant data powers.[171]

2. Congress.

Pursuant to the Necessary and Proper Clause, Congress has wide authority to authorize or direct data gathering and use by federal administrative agencies.[172] But Congress's subpoena power operates as an additional, expansive information power through which Congress can itself gather and use data without preceding through ordinary lawmaking procedures. Although "Congress has no enumerated constitutional power to conduct investigations or issue subpoenas," the Supreme Court has long

---

[168] *See* Cass R. Sunstein, *Clear Statement Principles and National Security:* Hamdan *and Beyond*, 2006 SUP. CT. REV. 1, 38:

> [I]t might be urged that as Commander-in-Chief, the President has the inherent power to engage in foreign surveillance . . . . On the most extreme version of this view, Congress cannot limit that power even if it chooses to do so. Foreign surveillance is a presidential prerogative, akin to dictation of the movement of troops.

[169] *See Zweibon*, 516 F.2d at 620–24; *Truong Dinh Hung*, 629 F.2d at 913.

[170] Much of the litigation in this area has focused on whether the President has the inherent power to conduct *warrantless* surveillance, a power so potent that it supersedes the Fourth Amendment, *see supra* note 41. Given the many forms of data gathering that are exempt from the Fourth Amendment, as Part I develops, disentangling the inherent powers question from the warrant-requirement question will only become increasingly important going forward.

[171] The State of the Union Clause requires the President to "give to the Congress Information of the State of the Union, and recommend to their Consideration such Measures as he shall judge necessary and expedient"—an interbranch information-sharing requirement that could in theory bear on Presidential power over data. U.S. CONST. art. II, § 3. The President also has the power to "require the Opinion, in writing, of the principal Officer in each of the executive Departments," suggesting a particular form of vertical information flow—at least in the form of written "opinions"—from agency heads to the President. *Id.* art. II, § 2. But those thin provisions could not plausibly ground a presidential power to supersede the rules Congress sets for data gathering and use by administrative agencies.

[172] *See id.* art I, § 8, cl. 18.

held that "each House has power 'to secure needed information' in order to legislate" by request or subpoena.[173] The Court has emphasized that "[t]his power of inquiry—with process to enforce it—is an essential and appropriate auxiliary to the legislative function. Without information, Congress would be shooting in the dark, unable to legislate 'wisely or effectively.' The congressional power to obtain information is," therefore, "'broad' and 'indispensable.'"[174]

But Congress's power to obtain information is not just broad, it is also exempt from the ordinary lawmaking requirements of bicameralism and presentment. It need not be scaffolded by formal statute nor exercised by the institution as a whole (as opposed to by each house of Congress acting independently). Each house of Congress routinely issues subpoenas independently and holds noncompliant targets in contempt. Indeed, the power is more diffuse still: "[T]he subpoena power may be exercised by a committee acting . . . on behalf of one of the Houses."[175]

Today, refusal to comply with a congressional subpoena is a federal criminal offense.[176] But even before 1857, when the relevant criminal statute was enacted, the House and Senate contended—and the Supreme Court agreed—that their inherent legislative powers supported their ability to issue subpoenas and punish noncompliant parties even absent statutory authorization.[177]

This diffusion of legally binding authority among actors within the institution of Congress is unusual. Ordinarily, Congress must act through bicameralism and presentment.[178] The canonical case, *Immigration and Naturalization Service v. Chadha*,[179] holds that "lawmaking"—defined as legislative acts that "alter[ ] the legal rights, duties and relations of persons"—is "a power to be shared by both Houses and the President."[180] Powers conferred on a single house—including impeachment, treaty ratification, confirmation of presidential appointees, and the prescription of

---

[173] *See* Trump v. Mazars USA, LLP, 140 S. Ct. 2019, 2031 (2020) (quoting McGrain v. Daugherty, 273 U.S. 135, 161 (1927)).

[174] *Id.* (internal alterations omitted) (first quoting *McGrain*, 273 U.S at 175; then quoting Watkins v. United States, 354 U.S. 178, 187, 215 (1957)); *accord McGrain*, 273 U.S. at 175 ("[W]here the legislative body does not itself possess the requisite information . . . recourse must be had to others who do possess it. . . . [and] some means of compulsion are essential to obtain what is needed.").

[175] Eastland v. U.S. Servicemen's Fund, 421 U.S. 491, 505 (1975).

[176] *See* 2 U.S.C. § 192.

[177] *See* Anderson v. Dunn, 19 U.S. (6 Wheat.) 204, 225 (1821).

[178] *See* U.S. CONST. art. I, § 7, cl. 2.

[179] 462 U.S. 919 (1983).

[180] *Id.* at 947, 952.

internal rules—are specifically and textually enumerated constitutional exceptions, not the rule.[181]

It is not obvious that Congress's subpoena power complies with the principle articulated in *Chadha*. Gathering information by subpoena alters "the legal rights, duties and relations of persons" by providing the target a choice between compliance and punishment.[182] The House and Senate's information-gathering privilege is thus one of the few nontextual exceptions to bicameralism and presentment, as the Supreme Court itself recognized in an early subpoena case.[183]

Our goal here is not to critique the legislative subpoena power but to illustrate its expansive character and its deviation from ordinary structural constraints.[184] To enact a tax—to obtain money from individuals—requires an act of law that proceeds through the ordinary processes of bicameralism and presentment. But to obtain information from individuals requires only the action of a single committee in one house of Congress.

3. The judiciary.

The federal judiciary also gathers vast amounts of data pursuant to inherent powers conferred by the Article III Vesting Clause without specific congressional authorization. Professors Zachary Clopton and Aziz Huq have recently examined the scope of the judiciary's data-gathering authority in detail.[185] They argued that the judiciary has long asserted capacious authority to gather and manage its own data absent congressional structuring.[186] And they contended that much of that power could be regulated by Congress, if Congress so chose, leaving a small core of inherent informational powers to the judiciary exclusively.[187]

---

[181] Indeed, each chamber's constitutional authority to create "[r]ules of its [p]roceedings," U.S. CONST. art. I, § 5, cl. 2, does not confer a subpoena power precisely because, as the Court has held, the subpoena alters the legal obligations of external parties and the "authority to determine the rules of its proceedings . . . cannot be construed to operate beyond the walls of the House, except on its own members, and its officers." *Anderson*, 19 U.S. (6 Wheat.) at 213–14.

[182] *See Chadha*, 462 U.S. at 952.

[183] *See Anderson*, 19 U.S. (6 Wheat.) at 213–14.

[184] *Cf.* William J. Stuntz, *O.J. Simpson, Bill Clinton, and the Transsubstantive Fourth Amendment*, 114 HARV. L. REV. 842, 864 (2001) (explaining why the broader "federal subpoena power [is] something akin to a blank check").

[185] *See generally* Clopton & Huq, *supra* note 8.

[186] *See id.* at 937.

[187] *See id.* at 928–31.

Clopton and Huq also showed that, as with Congress, judicial data policy is not generally set in a centralized manner by the judicial branch but instead by individual courts—it is doubly diffuse.[188] The courts, in other words, follow the same template as the other branches: They generally gather data on their own initiative without specific authorization, awaiting only acts of legal constraint from Congress. What matters for this high-level account of data's permissive default is the judiciary's capacity to gather data on its own initiative and the different diffuse actors within the judiciary's administrative apparatus who are empowered to make those decisions.

4. Administrative agencies.

Much of the government's data is acquired and used not by the President, Congress, or courts, but by federal administrative agencies. In contrast to the other branches, the "fourth branch" cannot claim inherent data powers: Administrative agencies have no independent constitutional power and can act only pursuant to delegations from Congress and the Executive.[189] But the anemic statutory framework that structures administrative decisionmaking when gathering new data has produced a data permissiveness in agencies as well.

There is, of course, significant variation in the way that agencies with different programmatic mandates, statutory structures, data needs, and, in some cases, area-specific data rights steward data. We cannot describe every variation here; our more modest goal is instead to describe a general pattern, one in which agencies exercise enormous discretion to gather and use individual data unless specifically directed not to.

a) Statutory minimalism. As one of us has previously observed, agencies administering major data programs and overseeing significant public "data pools" frequently draw authority from thin and outdated congressional authorizations that do not specifically contemplate the data gathering in question, its scale, its data sources, the uses to which the data can be put, the policy choices they embed, or the risks they create.[190] Some statutes

---

188 *See id.* at 917.

189 *See* City of Arlington v. FCC, 569 U.S. 290, 297 (2013) (contending that federal agencies' "power to act and how they are to act is authoritatively prescribed by Congress").

190 *See, e.g.*, Fahey, *Data Federalism*, *supra* note 4, at 1012, 1033 (describing the FBI's reliance on the Driver's Privacy Protection Act—a statute designed to *restrict* data sharing—as authorization to assemble a database of hundreds of millions of photographs); *id.*

authorizing significant data programs predate the data age entirely and envision information gathering of dramatically different size and effect.[191] Congress's approach to authorizing data-related policymaking, in other words, can often be characterized as a kind of "statutory minimalism," one in which agencies understand themselves to be authorized to collect and use data through generic grants of power that do not directly address—or, as a consequence, disclose publicly or invite popular engagement about—the data gathering and data uses that will result from those authorizations.[192]

The effects of that statutory minimalism are particularly pronounced when agencies use old statutory authority to gather and use data in new or transformative ways.[193] Consider the expanding agency practice of purchasing data from private data brokers.[194] Data brokers can offer access to forms of data not traditionally thought to be within the government's reach, including the kinds of behavioral and tracking data that consumers produce when using consumer technologies and which can provide detailed profiles of individual activity, tastes, tendencies, and motivations.[195] Congress has not affirmatively authorized the practice of data purchasing—and, indeed, many members of Congress have introduced bills to restrict it—yet it appears to have become common practice.[196] So too agencies seem to be drawing on undisclosed

---

at 1035 (describing the statute typically cited as authority for the National Crime Information Center, an enormous cross-governmental data pool used by virtually every police officer in the nation: a law passed in 1924 that generically authorizes the Attorney General to "acquire, collect, classify, and preserve identification, criminal identification, crime, and other records" (citing Act of May 28, 1924, Pub. L. No. 68-153, tit. II, 43 Stat. 205, 217)).

191 *See id.*; *see also* Friedman & Citron, *supra* note 8, at 1382 (describing the "[m]ission creep" noted by the Church Committee in which "programs initiated with limited goals, such as preventing criminal violence or identifying foreign spies, were expanded to what witnesses characterized as 'vacuum cleaners,' sweeping in information about lawful activities of American citizens" (quoting S. REP. NO. 94-755, at 3–4)).

192 *See* Fahey, *Data Federalism*, *supra* note 4, at 1032.

193 The argument is not, to be clear, that agencies are acting ultra vires or that there is a nondelegation or major questions defect in Congress's practice of using broad and dated language to empower contemporary data collections. It is instead a normative claim about a missed opportunity to control data power: Congress exerts less control than it could over how agencies gather and use data not just—as many have previously addressed—by failing to enact a broadscale data privacy law like Europe's General Data Protection Regulation and also by failing to flex its levers of structural control to more tightly oversee the agencies that use data on its behalf.

194 *See supra* notes 37–38.

195 *See supra* notes 37–38 and accompanying text.

196 *See, e.g.*, *Wyden, Paul and Bipartisan Members of Congress Introduce The Fourth Amendment Is Not For Sale Act*, RON WYDEN, U.S. SENATOR FOR OR. (Apr. 21, 2021),

existing statutory authorities to train machine learning technologies on governmental data, a course of conduct transparently without specific statutory authorization.[197] Those practices reproduce the pattern visible in the government's early "data mining"—which dramatically expanded analytical value government could gain from its data, and which swept through federal agencies two decades ago without express authorization from Congress.[198]

The permissive default created by Congress's statutory minimalism is only further illustrated by the exceptions that prove the rule. In pointed cases, Congress has by specific statute overridden the default—and forbidden particular collections and uses of data absent express and specific congressional authorization.[199] One important example is § 702 of the Foreign Intelligence Surveillance Act[200] (FISA), which legislatively authorized—but cabined and proceduralized—the National Security Agency's surveillance of foreign intelligence targets.[201] In an effort to constrain an executive that had previously expressed the view that the President could not constitutionally be so constrained,[202] Congress took the unusual step of subjecting its data-gathering authorization to a five-year sunset, after which the program would have to be

---

https://perma.cc/HR5Y-FQL5; *Durbin, Lee Introduce Bipartisan SAFE Act to Reform FISA Section 702*, U.S. SENATE COMM. ON THE JUD. (Mar. 14, 2024), https://perma.cc/TS7Z-TW3X; Purchased Data Inventory Act, S. 2292, 118th Cong. (2023).

197 *See infra* notes 226, 236 and accompanying text.

198 *See generally, e.g.*, U.S. GEN. ACCT. OFF., GAO-04-548, DATA MINING: FEDERAL EFFORTS COVER A WIDE RANGE OF USES (2004); JEFFREY W. SEIFERT, CONG. RSCH. SERV., RL31798, DATA MINING: AN OVERVIEW (2004); GINA MARIE STEVENS, CONG. RSCH. SERV., RL31730, PRIVACY: TOTAL INFORMATION AWARENESS PROGRAMS AND RELATED INFORMATION ACCESS, COLLECTION, AND PROTECTION LAWS (2003); DEF. ADVANCED RSCH. PROJECTS AGENCY, REPORT TO CONGRESS REGARDING THE TERRORISM INFORMATION AWARENESS PROGRAM: IN RESPONSE TO CONSOLIDATED APPROPRIATIONS RESOLUTION (2003).

199 The United States lacks general privacy protections for individual data, but some targeted privacy statutes in effect override data's permissive default. *See* Tax Reform Act of 1976, Pub. L. No. 94-455, § 1202, 90 Stat. 1520, 1667–85 (codified as amended at I.R.C. § 6103) (restricting disclosure of tax returns except for authorized purposes); Family Educational Rights and Privacy Act of 1974, Pub. L. No. 93-380, 88 Stat. 484, 571–74 (codified as amended at 20 U.S.C. § 1232g) (prohibiting disclosure of student educational data except to specifically authorized agencies and institutions); Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (codified as amended in scattered sections of 18, 26, 29, and 42 U.S.C.); 45 C.F.R. § 164.512 (2025) (describing restrictions, including on some governmental uses, of health information).

200 Pub. L. No. 95-511, 92 Stat. 1783 (1978) (codified as amended at 50 U.S.C. § 1801 et seq.).

201 *See* FISA of 1978 Amendments Act of 2008, Pub. L. No. 110-261, § 101, 122 Stat. 2436, 2438–48 (codified as amended at 50 U.S.C § 1881a).

202 *See supra* note 154.

authorized anew.[203] In § 702, Congress effectively recognized the value of subjecting at least some data collections to the kind of restrictive default that governs fiscal appropriations.[204]

b) Procedural minimalism. Agencies also make decisions about data against a permissive backdrop in a second important way. Congress channels and constrains agency action—and facilitates direct public control over it—through the procedural requirements of the Administrative Procedure Act. But here too data is exceptional: Data-related decisions are subject to less demanding process than comparable agency decisions, furthering the generally permissive environment in which data policy is made.

Most basically, agencies do not in the typical case proceed through the Administrative Procedure Act's notice-and-comment process when making decisions about what data to gather or how to use it. That process, which governs unless Congress displaces it, has been replaced for data-related decisions by a separate statute designed not to structure policymaking generally but to protect privacy rights—the Privacy Act of 1974. The Fair Information Practice Principles require government to provide notice about what data they collect from individuals and how it will be used. The Privacy Act instructs agencies to provide that notice by producing and publishing a technical announcement called a System of Records Notice before creating new databases or substantially changing existing ones.[205] They describe the data an agency intends to collect, the individuals or entities from whom they intend to collect it, the "routine uses" they intend to put it toward, and the legal authority on which the agency's data related activities rely.[206] Data may be disclosed only to advance those routine uses, to "officers and employees of the agency which maintains the record

---

[203] *See* 50 U.S.C. § 1881a note (Effective Date of Repeal); 50 U.S.C. § 1801 note (Transition Procedures for FISA Amendments Act of 2008 Provisions); *see also* Friedman & Citron, *supra* note 8, at 1435 (describing the default-like effect of § 702's sunset provision).

[204] Congress's 1974 regulation of the uses of the Social Security number likewise illustrates both the permissive default that generally attends data policy and Congress's targeted efforts to flip it. *See* Privacy Act of 1974 § 7, 88 Stat. at 1909–10. The Social Security number was created not by statute but by the Social Security Board in 1936, as historian Sarah Igo has chronicled. SARAH IGO, THE KNOWN CITIZEN 63–71 (2018). President Franklin D. Roosevelt subsequently amplified the number's power by instructing federal agencies to use it as a preferred identifier across programs "in the interest of economy and orderly administration." Exec. Order No. 9,397, 8 Fed. Reg. 16,095 (Nov. 22, 1943). In 1974, however, Congress overrode the default, reigned in the executive's administrative discretion, and banned the Social Security number's use except pursuant to express legislative permission. Privacy Act of 1974 § 7, 88 Stat. at 1909–10.

[205] *See* 5 U.S.C. § 552a(e)(4).

[206] *Id.* § 552a(e)(3).

who have a need for the record in the performance of their duties," and for a range of auditing, statistics, recordkeeping, and law enforcement objectives.[207]

As a practical matter, the requirements of the Privacy Act are relatively modest and easily circumvented. "Routine use" is a capacious term that can include any use "compatible with the purpose for which [the data is] collected."[208] And agencies are not obligated to defend the collection or use of data in SORNs using justificatory frameworks like cost-benefit analysis; SORNs disclose, but do not defend, data policymaking. Although the Act invites public comment on published SORNs, agencies are under no burden to respond to those comments, nor does the Act allow arbitrary and capricious review of the policy choices reflected in SORNs, a standard rubric through which an agency's consideration of public comments is evaluated.[209] Judicial review, in turn, is available only when an agency has handled an individual's data contrary to the SORN in a way that adversely affects her.[210] In practice, then, agencies have a strong incentive to make SORNs as general as possible—to characterize data sources generically and identify routine uses broadly—in order to insulate themselves from challenge.[211]

The deficits of the Privacy Act as a structural tool were on sharp display in the first efforts by the Department of Government Efficiency to access sensitive databases and the wave of litigation that followed.[212] First, although the Privacy Act purports to restrict data access to employees at the agency that "maintains

---

[207] *Id.* § 552a(b).

[208] *Id.* § 552a(a)(7).

[209] Agencies are instructed to receive comments on SORNs but can begin collecting data as soon as they publish the relevant SORN and can begin using data for a newly disclosed routine use just thirty days after that. *See* 5 U.S.C. § 552a(e)(4); *id.* § 552a(e)(11). A SORN, therefore, is not like a Notice of Proposed Rulemaking intended to be a draft that will be revised after comments are received; the statutory regime envisions the SORN as the end product.

[210] *Id.* § 552a(d)(3).

[211] Agencies often use templates to help draft SORNs, and those templates suggest that they yield to the incentive to insulate themselves from challenge by adopting sweeping routine uses ex ante. A common off-the-shelf routine use, for example, allows data to be used to implement or enforce virtually any law enacted by any level of government. *See, e.g.*, U.S. DEPT. OF HOMELAND SEC., SORN TEMPLATE 20170207 (available at https://www.dhs.gov/publication/system-records-notice-template) (suggesting "routine use" of: "implementing," "investigating," "prosecuting" or "enforcing" any "law, rule, regulation, or order" whether "criminal, civil, or regulatory" by an "appropriate federal, state, local, international or foreign" government).

[212] *See supra* note 6.

the record who have a need for the record,"[213] news coverage suggests that DOGE quickly circumvented that constraint. It devised a practice of detailing its staffers to several agencies at once so that they could be de jure agency employees for Privacy Act purposes, while still working de facto for DOGE itself.[214] Then, in the first Privacy Act case to reach the Supreme Court, the Court sided with the government in an unsigned order, without an accompanying opinion, which stayed a temporary injunction the district court had entered against DOGE. That district court injunction was, as Justice Ketanji Brown Jackson explained in her opinion dissenting from the Supreme Court's stay, narrow: It was "minimally burdensome" and a "short-term pause on giving DOGE unfettered" access to Social Security Administration data.[215] Yet even a serious Privacy Act challenge, securing only the limited relief the Act affords, could not earn serious consideration from the Supreme Court.[216]

The Privacy Act, in sum, does little to encourage agencies to make rational and justifiable decisions about data collection and use, much less decisions that are democratically accountable in a thick sense. Although it is in practice the central *structural* framework for agency policymaking about data, the Act is emphatically not designed for that role—it is a privacy regime (and a narrow one by modern privacy standards), not a structural statute designed to organize the effective, justifiable, and accountable deployment of government power.[217]

---

213 *See* 5 U.S.C. § 552a(b)(1).

214 *See* Faiz Siddiqui & Jacob Bogage, *Some DOGE Staffers Hold High-Powered Jobs at Multiple Federal Agencies*, WASH. POST (Apr. 14, 2025), https://www.washingtonpost.com/business/2025/04/14/elon-musk-doge-staffers/.

215 *See* Soc. Sec. Admin. v. Am. Fed'n of State, Cnty., and Mun. Emp., 145 S. Ct. 1626, 1629–30 (2025) (Jackson, J., dissenting from the grant of application for stay).

216 *Id.*

217 Since 1974, the Privacy Act has been supplemented with two other procedural requirements, but neither adequately addresses its policymaking gaps. The E-Government Act of 2002 requires agencies to conduct "Privacy Impact Assessments"—assessments that discuss a data collection's relevant privacy risks and how they will be mitigated. E-Government Act of 2002 § 208(b), 116 Stat. at 2921–22 (codified at 44 U.S.C. § 3501 note (Federal Management and Promotion of Electronic Government Services)). But like SORNs there is no cause of action to seek judicial review over their content or quality, and they focus on privacy to the exclusion of general policy justifications. *See generally id.* Another statute, the Paperwork Reduction Act, does impose a justificatory hurdle to data collections, but it focuses on just one policy consideration: how burdensome the data collection is for individuals and businesses. *See* Pub. L. No. 96-511, § 2, 94 Stat. 2812, 2812 (1980) (codified at 44 U.S.C. § 3501(1)). The Office of Management and Budget (not federal courts) is the central assessor of those burden estimates, but it is unusual for that office to disagree with the relevant analysis and to decline to approve a data collection. *See*

c) Judicial minimalism. Finally, as the previous discussion suggests, judicial review of data policymaking is in short supply. Judicial review under the Privacy Act is miserly. Only individuals "adversely affected" by information handling in violation of the Act can sue, and even then, the only remedies available are tied to the privacy-related harms suffered by those individuals.[218] There is no standard arbitrary and capricious review of the *policies* contained in SORNs—no hard look at the goal of the data collection, its burdens and benefits, or whether the agency justified its decision in light of public comments.[219]

The effects of this permissive regime are well illustrated by perhaps the most dramatic shift in the government's use of its data in the twenty-first century: the use of high-quality government data to train and deploy artificial intelligence tools for public purposes.[220] Data is the raw material that powers AI, and AI is "data-hungry."[221] The most sophisticated AI models require so much training data that they will soon exhaust the common crawl—a repository of 250 billion public web pages.[222] Curated data, of the type government frequently collects, is particularly scarce. Data policy, therefore, shapes government AI capabilities. And controlling the government's data supply is a direct and immediate mechanism through which to control government use of AI.

MAEVE P. CAREY & NATALIE R. ORTIZ, CONG. RSCH. SERV., IF11837, THE PAPERWORK REDUCTION ACT AND FEDERAL COLLECTIONS OF INFORMATION: A BRIEF OVERVIEW (2024). *See generally* Stuart Shapiro, *The Paperwork Reduction Act: Benefits, Costs and Directions for Reform*, 30 GOV'T INFO. Q. 204 (2013); 5 C.F.R. § 1320.8 (2025).

[218] *See* 5 U.S.C. § 552a(p)(3).

[219] *See* 5 U.S.C. § 552a(g)(1)(A)–(D). The same is true for the E-Government Act, 44 U.S.C. § 3501 note (Federal Management and Promotion of Electronic Government Services), and the Paperwork Reduction Act, 44 U.S.C. § 3507(d)(6) ("The decision by the [OMB] Director to approve or not act upon a collection of information . . . shall not be subject to judicial review.").

[220] We recognize that there are many ways to define and describe AI, U.S. GOV'T ACCOUNTABILITY OFF., GAO-24-105980, ARTIFICIAL INTELLIGENCE: AGENCIES HAVE BEGUN IMPLEMENTATION BUT NEED TO COMPLETE KEY REQUIREMENTS 6 (2023) [hereinafter GAO, ARTIFICIAL INTELLIGENCE] (collecting six different definitions of AI reflected just in federal statutes), and we focus here on the technologies that the federal government has itself identified as AI tools. *See generally 2024 AI Use Case Inventories Reference*, U.S. CHIEF INFO. OFFICERS COUNCIL (Dec. 16, 2024), https://perma.cc/48CS-RYCH.

[221] *See* David Freeman Engstrom, Daniel E. Ho, Catherine M. Sharkey & Mariano-Florentino Cuéllar, *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies,* 1, 72 (N.Y.U. Sch. of L., Pub. L. Rsch. Paper No. 20-54, 2020); *see also* Solow-Niederman, *supra* note 57, at 38.

[222] *See* Pablo Villalobos, Anson Ho, Jaime Sevilla, Tamay Besiroglu, Lennart Heim & Marius Hobbhahn, *Will We Run Out of Data? Limits of LLM Scaling Based on Human-Generated Data*, 41 PROC. INT'L CONF. ON MACH. LEARNING, 2024, at 2.

Yet the government is proceeding to do just that—to develop public use AI, or AI tools designed to perform public functions—largely without public engagement or oversight. Data's permissive default is one important reason why. Although a recent Government Accountability Office report found 1,200 AI projects in planning or operation, we find mentions of "AI," "artificial intelligence," or machine learning" in just four SORNs, and they are passing at that.[223] Nearly every government AI project lacks a standardized record of the data used to train or use it. Meanwhile, the Federal Trade Commission has warned private firms that "surreptitiously" amending privacy policies to permit the use of consumer data to train AI may be an "unfair or deceptive" practice.[224] The Privacy Act's permissive regime has thus enabled the government to reroute its data to AI projects in the same surreptitious ways it has condemned in private firms. That the Act has not prompted disclosure or notice about the data that agencies are using to train AI reflects its failure to facilitate that control by permitting agencies to use overly general descriptions of data's use that cannot inform the public of genuine changes in the stewardship of data assets or force agency justifications that must withstand judicial review. It reflects, in short, the Act's structural limitations as a tool for organizing public engagement with data policy.

B. Specialist Acquisition and Allocation

The Constitution's permissive environment for gathering data is not the only way in which data's structural ecosystem differs from the analogous controls that accompany the government's use of money and other assets. In the context of data, decisions around acquisition and allocation are made, at least in significant measure, by specialists rather than generalists—that is, by the very experts who benefit from expanding the government's access to data. Recall that decision-making around money is made

---

223 *See* GAO, ARTIFICIAL INTELLIGENCE, *supra* note 220, at 3 (assessing AI uses across a study group of just 20 agencies). Searching in the Federal Register, which contains every SORN published, we found one reference to "AI" as an abbreviation for artificial intelligence (in a citation to the AI in Government Act of 2020) and four to "artificial intelligence." We found three matches for "machine learning," all of which were within the three SORNs that also mentioned artificial intelligence." We caveat that the database only includes SORNs going back to 2000, but given the recent advent of AI technologies, we think it unlikely that the exclusion of earlier SORNs skewed our results. *See Document Search*, *supra* note 23.

224 *AI (and other) Companies: Quietly Changing Your Terms of Service Could Be Unfair or Deceptive*, FTC (Feb. 13, 2024), https://perma.cc/E78S-9LXH.

in key respects by three powerful generalist bodies—the House of Representatives, the Senate, and the President, who balance the branches' requests for fiscal power against other competing considerations. By contrast, policymaking around data is generally made within the branches or agencies themselves—the entities that stand most to benefit from data collection and whose staffers are least attuned to the policy costs of expansive data power.[225]

The *internal* structure of many agencies further magnifies that dynamic. Data gathering decisions are frequently made within agencies by IT officials and procurement departments—the kind of back office staffers who we generally do not imagine making significant policy choices, but who embed consequential judgments about how data will be managed and shared in the technical systems that host and make databases accessible. We hypothesize that the procurement process represents the central forum for resolving policy choices about data—a hypothesis that warrants greater exploration in future work, whether by us or others.[226]

Procurement has serious flaws as a policymaking device. Although it is highly legalized, federal acquisition law generally focuses on government efficiency and the rights of bidders and purchasers, not on public input or the assessment of broader policy objectives.[227] A procurement officer is generally not expected

---

[225] Consider, for example, the elaborate bureaucracy of federal, state, and local police who oversee the National Crime Information Center—the nation's most sweeping data pool—largely without generalist supervision. *See* Fahey, *Data Federalism*, *supra* note 4, at 1047–49.

[226] Among the striking examples in recent years is the government's decision to procure data directly. *See supra* notes 37–39. Likewise, as the government has expanded its efforts to use AI, procurement has also taken center stage. Indeed, the Advancing American AI Act regulates little, but it does recognize the role that procurement plays in important data-related AI questions. *See* Advancing American AI Act, Pub. L. No. 117-263, § 7224(d)(1)(A)(iii), 136 Stat. 3668 (2022) (codified at 40 U.S.C. § 11301) (instructing the OMB to "address the ownership and security of data and other information created, used, processed, stored, maintained, disseminated, disclosed, or disposed of by a contractor or subcontractor on behalf of the Federal Government"); OFF. OF MGMT. & BUDGET, ADVANCING GOVERNANCE, INNOVATION, AND RISK MANAGEMENT FOR AGENCY USE OF ARTIFICIAL INTELLIGENCE 25 (2024) (instructing procurement officers to "consider contracting provisions that protect Federal information used by [AI] vendors . . . , so that such data is protected from unauthorized disclosure and use and cannot be subsequently used to train or improve the functionality of the vendor's commercial offerings without express permission from the agency").

[227] Steven J. Kelman, *Achieving Contracting Goals and Recognizing Public Law Concerns*, *in* GOVERNMENT BY CONTRACT 153, 153 (Jody Freeman & Martha Minow eds., 2009) (noting that the role of "administrative law" in federal contracting has "focus[ed on] . . . the treatment of those wishing to sell to the government, not on . . . members of the public . . . or others concerned about democratic governance, the exercise of state power, and respect for individual rights").

to make normative evaluations when acquiring data or data technology. Rather, the procurement process tends to take policy needs as exogenous and to simply seek tools that will maximize the agency's ability to meet its ends.

Delegating significant discretion to administrative agencies need not result in specialist policymaking without generalist oversight. Even absent constitutional controls of the kind that we describe for fiscal power, the President can exert control over agencies and direct their decision-making in ways cognizant of broader policy tradeoffs.[228] Executive Order 12,866, for example, requires that all "significant regulatory action"—that is, all major regulatory proposals by agency specialists—be reviewed by the administrative generalists in the Office of Information and Regulatory Affairs (OIRA) in the Executive Office of the President.[229] OIRA has ninety days to conduct such reviews.[230]

But this common form of generalist control is, we posit, significantly weaker with respect to data-related decisions. Although OMB regulations require agencies to submit their Privacy Act SORNs to OIRA for review, they provide the office only thirty days in which to conduct that review before the notices can be published in the Federal Register and, at the same time, the data collection can begin.[231] The short time period the office requests to conduct its review suggests that it is likely considerably less robust than its review of agency rules pursuant to Executive Order 12,866. As noted above, the Paperwork Reduction Act also requires review by OMB, but we think likely does little to yield generalist review of data decision-making. The Act invites the White House to assess the agency's evaluation of just one concern related to data collection—its burden on regulated parties—to the exclusion of the many other concerns that generalists ought to consider.

C. Decentralized Technical Administration

Just as there is no Appropriations Clause for data, there is no Treasury for data. The government's data is not held by one single institution, nor is it subject to centralized oversight by an

---

[228] Elena Kagan, *Presidential Administration*, 114 HARV. L. REV. 2245, 2363 (2000).

[229] Exec. Order No. 12,866, 58 Fed. Reg. 51,735 (Oct. 4, 1993), *as amended by* Exec. Order No. 14,094, 88 Fed. Reg. 21,879 (Apr. 11, 2023).

[230] *Id.*

[231] OFFICE OF MGMT. & BUDGET, OMB CIRCULAR NO. A-108, FEDERAL AGENCY RESPONSIBILITIES FOR REVIEW, REPORTING, AND PUBLICATION UNDER THE PRIVACY ACT 13–14 (2016).

institution that sets standard operational practices for agencies to follow. Rather, each agency governs its own data using a mishmash of technical systems.[232] Most importantly, each agency exerts ultimate control over its data, the structure of its databases, and the technologies it uses to manage and transfer that data.

This decentralized administration has several consequences. First, it shapes how government data moves into government reserves and between government agencies. It means that data sharing among agencies must be the result of a bilateral negotiation. Together, the sending and receiving agencies must decide whether (for example) one agency will allow another to read or consult its data, or instead to copy and transfer the sending agency's raw data files. To share raw data, rather than merely to permit access, is to empower the recipient to mine, use, and transfer that data, while multiplying privacy and security risks. On the other hand, to share only access to information is to restrict what a receiving agency can do with that data, perhaps at real cost to programs that could benefit from more comprehensive access.

As we explore below, private firms are increasingly using sophisticated data-sharing techniques like data escrows, data white rooms, data enclaves, and Hippocratic databases to allow companies with sensitive data assets to transfer data in much more controlled and calibrated ways.[233] But each of those techniques requires some centralized data infrastructure, which for now is lacking in the federal government.

---

[232] Without producing a map of the technologies the government uses, we cannot know the precise contours of its patchwork of systems, but we can see that such a patchwork exists in three significant data initiatives that would not be necessary were the government's technical systems more standardized. First, for decades, the government has pursued the goal of data "interoperability"—or the ability to integrate data sets and data systems that are technically distinct. *See, e.g.*, U.S. GOV'T ACCOUNTABILITY OFF., GAO-22-106175, PUBLIC HEALTH EMERGENCIES: DATA MANAGEMENT CHALLENGES IMPACT NATIONAL RESPONSE 1 (2022) (noting that "the federal government still lacks [an interoperable disease surveillance] . . . network"); U.S. DEP'T OF HOMELAND SEC., PRIVACY IMPACT ASSESSMENT UPDATE FOR THE BIOMETRIC INTEROPERABILITY BETWEEN THE U.S. DEPARTMENT OF HOMELAND SECURITY AND THE U.S. DEPARTMENT OF JUSTICE 2 (2011). Second, the government's effort, most recently articulated in a significant cyber security Executive Order to adopt a "[z]ero [t]rust" security approach, has been hampered by technical variation across agencies. *See* Exec. Order No. 14,028, § 3(a), 86 Fed. Reg. 26,633, 26,635 (May 12, 2021); U.S. GOV'T ACCOUNTABILITY OFF., GAO-23-106065, ZERO TRUST ARCHITECTURE 2 (2022) (highlighting technological variation as a barrier to a zero trust security approach). Finally, the need for data inventories, *see infra* notes 252–54, suggest a disorganized data approach.

[233] *See, e.g.*, Siyuan Xia, Zhiru Zhu, Chris Zhu, Jinjin Zhao, Kyle Chard, Aaron J. Elmore, Ian Foster, Michael Franklin, Sanjay Krishnan & Raul Castro Fernandez, *Data*

Second, decentralized administration shapes which personnel *within* government make data-related decisions—not only infrastructural choices but also policy choices. We hypothesize that consequential decisions about the systems used to manage data are made not by senior policymakers (or even, in many cases, program managers) but by IT officials within agencies who build the systems that facilitate data sharing and negotiate data-sharing agreements with sister agencies.

Third, this practice also affects how actors both inside and outside of government exert control over the government's data. In the context of fiscal power, for instance, the Treasury provides a salient focal point for fiscal policy, acting as a trusted conduit of information regarding the government's fiscal assets, a home in which to centralize the government's technocratic fiscal expertise, and a place to lay blame if security or use problems arise. By contrast, it would be nearly impossible to aggregate information about data-related infrastructure—whether security, data system design, or data-sharing technology—across agencies.

Finally, decentralization of data administration shapes—and, we think, distorts—our understanding of what data the government has, as we explore further below.

D. Bilaterally Negotiated Movements

Although President Trump has ordered government data to be aggregated across "information silos"—and the impact of that order remains to be seen—government data movements between branches and agencies have not historically been centrally managed.[234] Congress has authorized agencies to "make available to another agency[ ] information obtained by a collection of information if the disclosure is not inconsistent with applicable law."[235] As a result, agencies have generally negotiated data exchanges bilaterally. Because of the discretion individual data custodians have to share or withhold data, data has come to be treated, in practice, as something of a proprietary asset, which can be deployed as the policy (and personal) interests of its custodian dictate. This sets data in striking contrast to other governmental

Station: Delegated, Trustworthy, and Auditable Computation to Enable Data-Sharing Consortia with a Data Escrow, 15 PROC. VLDB ENDOWMENT 3172, 3183 (2022).

[234] *See supra* note 7.

[235] 44 U.S.C. § 3510(a). The Director of the OMB may also in theory "direct an agency to make available [information] to another agency." *Id.*

assets, which are the property of the government and managed for the public benefit.

These bilaterally negotiated data flows mean that data can, at times, move with ease across agencies to create new data assets of mutual benefit to the sender and recipient—which may have significantly greater value and capacity than the raw and diffuse data initially collected. Consider the federal government's decade-long effort to develop and use facial recognition technology. The FBI and the Department of Homeland Security have aggregated hundreds of millions of photographs from across federal, state, and local agencies for the purpose of conducting facial recognition searches. Agencies rely on individual grants of authority for their authority to share or receive data and set the contours of that data sharing through data use agreements that have been identified by researchers and reported on by newspapers but have never been systematically disclosed to the public.[236] But Congress has not organically structured the joint data pool by statute like how it might authorize and determine the parameters of policy programs of similar stature.

The bilaterally negotiated nature of the government's data flows does not always result in too much data sharing. It can also yield the opposite result—the husbandry of data by agencies in situations where cross-governmental data sharing would be useful or desirable. By placing significant control in the hands of data custodians within agencies, basic political economics would predict that data flows that are inconvenient, embarrassing, unhelpful, or simply burdensome to the current data custodians are unlikely to materialize.[237]

---

[236] *See* GAO, FACIAL RECOGNITION TECHNOLOGY, *supra* note 64, at 15 (reporting that the "Automated Biometric Identification System (IDENT) included roughly 836 million facial images" from "visa application[s], passport[s], mug shot[s], and others"); Fahey, *Data Federalism*, *supra* note 4, at 1025 ("The FBI pools 'hundreds of millions of photos' from state DMVs for use in facial recognition searches." (quoting U.S. GOV'T ACCOUNTABILITY OFF., GAO-19-579T, FACE RECOGNITION TECHNOLOGY 18 (2019)); U.S. GOV'T ACCOUNTABILITY OFF., GAO-23-105607, FACIAL RECOGNITION SERVICES: FEDERAL LAW ENFORCEMENT AGENCIES SHOULD TAKE ACTIONS TO IMPLEMENT TRAINING, AND POLICIES FOR CIVIL LIBERTIES 6 n.19 (2023) [hereinafter GAO, FACIAL RECOGNITION SERVICES] ("The photos in a facial recognition service's gallery may be drawn from various sources, including public web sites.").

[237] Scholars, for example, have been critical of agencies for not making better use of existing race data to evaluate the disparate impacts of governmental programs, as President Joe Biden envisioned in his Executive Order "Advancing Racial Equity and Support for Underserved Communities Through the Federal Government." *See* Exec. Order No. 13,985, 86 Fed. Reg. 7,009 (Jan. 29, 2021). Some, we think incorrectly, lay blame with the formal legal requirements of the Privacy Act, *see* Jennifer King, Daniel Ho, Arushi Gupta, Victor

The Privacy Act, for its part, nominally adds an additional layer of regulation for data flows of personal identifying information by requiring the written consent of the data subject before transferring data outside its originating agency, but the Act makes two sweeping exceptions that eclipse the rule: It exempts transfers that are consistent with the routine uses disclosed in the system's SORN, which, as noted above, are often so generic as to impose little serious constraint on data use; and it exempts transfers "for a civil or criminal law enforcement activity if the activity is authorized by law."[238] The agency that shares information must "keep an accurate accounting" of that transfer but, consistent with the Privacy Act's focus on the individual, those accountings are only available at the request of the individual whose data has been shared.[239] They are not disclosed publicly in the kind of summary form that would disclose how data moves throughout the government and for what purposes. Consistent with other areas in which agencies must coordinate their activities, many intergovernmental data-sharing projects are structured through individually negotiated memoranda of understanding or data sharing agreements, which are likewise not affirmatively disclosed.[240]

Y. Wu & Helen Webley-Brown, et al., *Privacy-Bias Tradeoff: Data Minimization and Racial Disparity Assessments in U.S. Government*, 2023 PROC. ACM CONF. ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 492. But in our assessment, the Privacy Act is more bluster than substance. It is, as noted above, legally easy (if bureaucratically irritating) to comply with the Privacy Act, given the absence of judicial review and the resulting option to use vague and general SORNs to justify nearly any data use. An understanding of the extraordinary control that data custodians exert over the movements of their data (individualized control that has no analogue in the fiscal context, where funds are public goods that must be centrally appropriated) suggests that data *under*sharing is more likely shaped by the basic incentives of data custodians. Data sharing that might embarrass or simply does not benefit a custodian (like the transfer of race data to a sister agency to diagnose disparate treatment) would be unlikely absent a regime of central control.

238 5 U.S.C. § 552a(b). Data sharing between agencies and Congress is also always permitted. *Id.* § 552a(b)(9).

239 *Id.* § 552a(c)(1).

240 *See* Fahey, *Data Federalism*, *supra* note 4, at 1040–45. For a discussion of more formalized agreements, see *id.* at 1039 (describing required agreements for certain computer matching programs pursuant to 5 U.S.C. § 552a(a)(8), (o)). For agency-specific examples, *see, e.g.*, *U.S. Census Bureau Statement on Data Sharing Agreements*, U.S. CENSUS BUREAU (Mar. 7, 2019), https://perma.cc/SC3A-7SRL (describing the Bureau's "routine data sharing agreements with many federal and state agencies, such as the Social Security Administration, Internal Revenue Service, Centers for Medicare and Medicaid Services and U.S. Department of Housing and Urban Development"); *IRS Information Sharing Programs*, U.S. INTERNAL REVENUE SERVICE (last updated Feb. 14, 2025), https://perma.cc/NE9K-5D6H ("[IRS i]nformation sharing utilizes agreements to strengthen

Returning to the facial recognition database, the Government Accountability Office is the best governmental source of information about the government's use of facial recognition—and the data that supports it—but the Office's various reports themselves have the tone of investigative journalism as it struggles to understand of the scale, scope, and tenor of one of the government's most significant data programs with the minimal information about data's movements around government that are available *even* to the government itself.[241]

E. Tentative Measurement

The government's money is gathered in Treasury accounts held by the Federal Reserve Bank of New York and is subject to standardized and detailed accounting for public inspection.[242] The government's land, for its part, is tracked and surveyed by the Bureau of Land Management in the Department of the Interior, which discloses not just metes and bounds but also information about mineral deposits, wildlife, and improvements.[243] But our diffuse and accretive system for holding and documenting government data means that our knowledge of the government's data assets is—in contrast to money and land—both decentralized and nonstandardized. We have no systematic, consistent, and aggregated way of measuring the data held by our government. Indeed, the government's focus—such as it is—on providing prospective information about its data assets has generally focused on identifying data that can be disclosed publicly without implicating state interests in privacy, trade secrets, security, internal deliberation, and the like.[244] But patchwork disclosure of government-held data,

---

relationships and collaboration."); *Data Sharing Agreements*, CTRS. FOR MEDICARE & MEDICAID SERVS. (last updated Sept. 10, 2024), https://perma.cc/CA87-B883.

241 *See* GAO, FACIAL RECOGNITION SERVICES, *supra* note 236, at 36 (noting that only three out of seven agencies studied had privacy and civil liberties policies governing facial recognition); U.S. GOV'T ACCOUNTABILITY OFF., GAO-22-106100, FACIAL RECOGNITION TECHNOLOGY: FEDERAL AGENCIES' USE AND RELATED PRIVACY PROTECTIONS 4 (2022) (similar); GAO, FACIAL RECOGNITION SERVICES, *supra* note 236, at 14 ("[T]he FBI could not fully account for searches it conducted using two [facial recognition] services."); *id.* at 41 ("CBP provides staff access to facial recognition services but does not have information on the number of staff that use . . . [them] or how often.").

242 *See supra* notes 136–38.

243 *See supra* note 139.

244 The Electronic Freedom of Information Act Amendments of 1996, for example, required agencies to disclose repeatedly requested records. *See* Pub. L. No. 104-231, § 4, 110 Stat. 3048, 3049 (codified at 5 U.S.C. § 552(a)(2)). So too the OPEN Government Data Act, Pub. L. No. 115-435, tit. II, 132 Stat. 5529, 5534–44 (2019) (codified in scattered sections

while useful for researchers whose data needs those disclosures address, does little to tell the public in consolidated and legible form what data the government holds and, most importantly, how that data is used by the government itself.[245]

The few data tracking measures that Congress has required agencies to undertake are both ineffective and incomplete. The Privacy Act—the 1974 statute that requires agencies to publish notices of new databases in the Federal Register—mandates that agencies disclose plans to collect data ex ante but does not require them to disclose the data they in fact collected ex post. And, as elaborated above, the incentives to make disclosures sweeping and imprecise are strong.[246] The result is a suite of decentralized documents describing in highly equivocal and gestural terms the data that government *may* possess, but that gives us little sense of what it *in fact* possesses.

The Privacy Act does authorize individuals to seek access to their own records in many cases.[247] But the statutory regime envisions this right being exercised piecemeal—primarily when an individual suffers an adverse event because of the government's use of her data, believes the record to be inaccurate, and wishes to inspect and correct it.[248] The Privacy Act was "designed to provide *individuals* with more control over the gathering, dissemination, and accuracy of agency information about themselves,"[249] not as a general-purpose knowledge-gathering statute allowing members of the public to ascertain "what their Government is up to."[250] In practice, a person who intended to develop an account only of the data the government holds about *her* would have to sift through thousands of SORNs, going agency by agency, notice by notice, and surmising whether her data might have been collected by the systems so described. The same is true of the Freedom of

---

of 44 U.S.C.), created an online data disclosure repository, but agencies have been slow to fully comply with those disclosure mandates. *See* U.S. GOV'T ACCOUNTABILITY OFF., GAO-22-104574, OPEN DATA: ADDITIONAL ACTION REQUIRED FOR FULL PUBLIC ACCESS 21–22 (2021) [hereinafter GAO, OPEN DATA]; *see also* Morten & Kapczynski, *supra* note 82, at 515 (noting the deficits in data disclosure at the Food and Drug Administration and calling for "more comprehensive data publicity").

245 *See* GAO, OPEN DATA, *supra* note 244, at 21–25.

246 *See supra* note 211 and accompanying text.

247 *See* 5 U.S.C. § 552a(d)(1). There are several exemptions including, for example, for data "compiled in reasonable anticipation of a civil action or proceeding." *Id.* § 552a(d)(5).

248 *See id.* § 552a(d).

249 Greentree v. U.S. Customs Serv., 674 F.2d 74, 76 (D.C. Cir. 1982) (emphasis in original).

250 Nat'l Archives & Recs. Admin. v. Favish, 541 U.S. 157, 171–72 (2004).

Information Act, which likewise requires an individual to submit records requests to each agency individually and tailor those requests to records held by that agency.[251]

In 2018, Congress instructed agencies to begin to fill the data-accounting gap by producing a "clear and comprehensive understanding of the data assets in the possession of [each] agency."[252] Motivated by an underspecified goal of helping the government better use its data assets (not, importantly, to help the public better understand or control the government's data assets), Congress's mandate was narrow. It instructed agencies to create "data inventories"—relatively rudimentary assessments of data, which describe the existence of a data set and its basic contours, but do not quantify its contents, describe its value, or convey its uses.[253] Data inventories, in short, are not a system equivalent to accounting for money, or surveying for land. They are descriptive at the highest level. Even so, the government has been slow to implement that limited mandate. A 2022 report noted that agencies had made only limited progress—sufficiently little to warrant an agency by agency status report.[254]

Even at its best, however, a data inventory does not convey the content necessary to understand how the government gathers, shares, and uses its data—or how access to data is distributed among governmental actors. It cannot, for instance, tell an individual who within government holds data about her and what data they possess. Nor can it tell the public when and how data has been used for new or controversial purposes—whether civil data is used for criminal law enforcement, for example; or what administrative data has been transferred from agencies like the Internal Revenue Service, the U.S. Census Bureau, and the Department of Health and Human Services for immigration enforcement; or what data has been used to train artificial intelligence systems. It cannot tell us what analytical strategies agencies have applied to their data, nor what knowledge they have produced. Nor do inventories track how data moves among governmental actors or trace where the data in the possession of a particular agency came from.

---

[251] *See supra* note 82 and accompanying text.

[252] OPEN Government Data Act § 202(d)(1), 132 Stat. at 5538 (codified at 44 U.S.C. § 3511(a)(2)).

[253] *See id.*

[254] *See Enterprise Data Inventories*, CHIEF DATA OFFICER COUNCIL (Apr. 2022), https://perma.cc/5ZJA-MFUA.

These early efforts at data inventories are fainthearted. The government has not expressed a commitment to transparency about the data it holds or articulated a strategy for getting there. We continue to need for data what the Accounting Clause provides for money: a legible and auditable understanding of the government's data assets, including their form, use, transformations, movement, security, and access. In the next Part, we begin to suggest an alternative path to getting there.

IV. RESTRUCTURING DATA POWER

The previous Part helps explain why so much data policy happens outside public view, without popular engagement, and in ways illegible even to interested voters. Although public law is broadly concerned with structuring power, data generally lacks the structures of control that ensure other instruments of power are popularly responsive. In this Part, we identify three paths to expanding popular control over the government's data and the power it confers. These proposals largely accept the constitutional baseline as a given, asking instead whether there are other ways to adjust the structural mechanisms to which data is subject. This is not because we view the constitutional baseline as settled: In the coming years, we think there may be opportunities to liquidate the Constitution's information powers and their application to digital data. But because transformative constitutional change is not a realistic component of data's immediate structural future, we focus here on statutory, regulatory, and technical ways to assert greater control over government data.

We begin with data's permissive default and the outsized role played by specialists, rather than generalists, in deciding what data to obtain and how it should be used. We identify a range of statutory and regulatory reforms to enhance generalist oversight of data policy and to close the gaps between the treatment of ordinary administrative activity and data in Congress and the Executive Branch.

But our emphatic view is that we cannot address the deficits identified in the previous Part with legal rules alone. Data is unique, and its technical nature presents unique technical opportunities. The government can control data through ordinary legal tools, like institutional arrangements and procedural requirements. But the possibilities for exerting control over data can also be internally determined by technical design: by structuring data

systems, frameworks, architectures, and code. Correctly calibrating those systems can play an essential role in facilitating popular control over the government's data. We are not the first to make this point. Writing about the early internet, Professor Lawrence Lessig denaturalized the sharp division between law and technical design by emphasizing the lawlike function of the internet's basic architecture. The Internet, in Lessig's account, is "governed" not just by law or markets but by law, norms, markets, and technical systems acting together.[255] Likewise, to improve governmental stewardship of data, we need an integrated legal-technical approach, one characterized by what we call *law-informed engineering*. We should control public data not through off-the-shelf systems procured by IT departments, on the one hand, and through antiquated legal rules designed imperfectly to maneuver and control those systems' capabilities, on the other, but through data systems that are *designed* as systems of government structure.

To that end, we also propose two solutions for increasing popular control over data that are legal-technical, rather than merely legal, in nature. The first addresses the significant gap in accounting for (and thus in public knowledge about) the government's data assets. The government's money and land are measured using accounting and surveying techniques that are borrowed from, and standard across, the public and private sectors. But the problem of data valuation is a complex technical challenge, one in its earliest stages even in the private sector. We propose a different approach for data, one that can bring transparency to the government's data assets, suggesting that what data scientists call *provenance* can be reimagined to serve the function for data that accounting serves for money and surveying serves for land.

We then propose a second solution, one aimed at the challenges associated with data's disordered movements among governmental agencies and the lack of centralized administrative oversight. Drawing on the literature regarding controlled data transfer, we argue that the government should make use of the *data escrow*, an infrastructure for sharing access to and insights from data held by decentralized custodians in ways that are tightly calibrated to joint objectives, finely controllable, and institutionally salient.

---

[255] LAWRENCE LESSIG, CODE: VERSION 2.0, at 234 (2006).

A. Statutory and Regulatory Reforms

We begin by considering the classic tools of public law: statutory and regulatory reform. We identify a handful of proposals for thoughtfully restructuring the government's data power and increasing popular control over public data. First, there are several ways Congress could enhance its control over government data through legislative craft. Although easier said than done—and said in many areas to boot—Congress should assert its constitutional prerogatives to structure and regulate data more robustly. Especially when faced with a President willing to assert broad and inherent power to control the government's data supply, Congress should conscientiously draft database enabling acts—statutes that authorize the creation and regulate the use of federal databases—rather than let agencies and the Executive Branch draw on broad and outdated sources of authority to create and use data resources.

That conscientious drafting could, for example, make more regular use of sunset provisions, like the five-year sunset in § 702 of the Foreign Intelligence Surveillance Act, to flip the default on data programs and force agencies making data-gathering and data-use decisions to return to Congress for affirmative consent.[256] Relatedly, Congress could use specific language (both legally and technically) in authorizing data gathering and use, thus tethering data programs more concretely to democratic authority and discouraging agencies from claiming authority from general, technologically outdated enabling acts. This, too, would have a default-shifting effect, both indirectly (in that it might gradually shift agencies' understanding of what they can do absent specific statutory authority) and directly (in that, as technology shifts, agencies would need to return to Congress for further guidance).

Second, Congress should amend the Privacy Act of 1974 to ensure more thoughtful, transparent, and deliberative administrative policymaking about data. That Act, as we discuss above, was designed for a single specific purpose: to provide individuals with notice that their data is being collected and of how it will be used. It was not designed to facilitate rational, publicly responsive policymaking by requiring agencies to justify data programs or engage with the public in designing them. Privacy, to be sure, should play a significant role in data policy. But the Privacy Act is a predictably inadequate mechanism for facilitating public

---

[256] *See supra* notes 203–04 and accompanying text.

engagement in, and control over, data policymaking. Notice-and-comment procedures are no panacea, of course, but scholarly efforts to imagine new forms of participation in administrative processes have never been more energetic.[257] What's clear is that data policy has an accountability deficit, and, in our view, more robust procedures would help ensure more robust public participation.

Third, Congress should expand its institutional capacity to oversee data policy. Our discussion of fiscal power illustrates that multiple, overlapping generalist bodies can create a system of public transparency, contestation, and engagement in stewarding the fisc. And an important part of that ecosystem is the institutional capacity of the Executive Branch and of Congress to participate in the creation and oversight of fiscal power, as we show. When, for example, President Richard Nixon's effort to impound appropriated funds threatened Congress's control over appropriations, Congress's retort was an institutionalist one: The Congressional Budget and Impoundment Control Act of 1974 created the powerful Congressional Budget Office to more robustly superintend federal expenditures.[258] The Executive Branch's Office of Management and Budget, the House and Senate budgeting committees, and Congress's muscular fiscal bureaucracy in the office of Comptroller General likewise create institutional capacity to steer and contest fiscal choices across the branches. By contrast, Congress's capacity to superintend the federal government's data policy is comparatively anemic, as its late-in-the-game efforts to require inventories of data assets and data uses (efforts that have yielded only partial compliance) suggest.[259]

Fourth, turning to the Executive Branch, the Office of Management and Budget and the Office of Information and Regulatory Affairs—the entities within the White House that oversee the Privacy Act and the E-Government Act—could devise better internal rules (without any legislative intervention) for identifying major data policy choices, whether they implicate privacy or not. Those offices could use the familiar administrative rubric of Executive Order 12,866, which required cost-benefit analysis for all "significant" rules.[260] The President could likewise require agencies to

---

[257] *See generally* Nicholas Bagley, *The Procedure Fetish*, 118 MICH. L. REV. 345 (2019). *But see* Nikhil Menezes & David E. Pozen, *Looking for the Public in Public Law*, 92 U. CHI. L. REV. 971, 1020 n.226 (2025) (collecting literature on new forms of participation in administrative policymaking).

[258] *See supra* notes 123–27 and accompanying text.

[259] *See supra* notes 252–54 and accompanying text.

[260] Exec. Order No. 12,866, § 6(a)(3)(B), 58 Fed. Reg. 51,735, 51,741 (Sept. 30, 1993).

designate significant data collections and significant new data uses and to meet a heightened justificatory burden for those actions.[261]

Finally, all three branches of the federal government, as well as the law schools that train their legal professionals, should build expertise at the intersection of law and computer science. President Ronald Reagan's initial effort to require cost-benefit analysis for administrative rules helped prompt the rise of the law and economics movement and the resulting class of economically trained lawyers (and legally literate economists) now common in government and legal academia.[262] As data becomes a raw material essential to government programs—and as applications for transformative machine learning tools powered by that data rapidly multiply—literacy in law and computer science will likewise be essential to ensuring a healthy and publicly responsive ecosystem for stewarding that power.

Indeed, technical design is as much a part of data's structural ecosystem as is the body of legal rules canvassed in this Article. To illustrate that point, the next two subparts draw on new research in computer science to suggest large-scale technical projects and reforms that, we think, would help reorganize data's structural ecosystem for the better.

B. Data Accounting

As we have argued, measurement is a linchpin of popular control over instruments of power, as the Constitution's Accounting Clause for fiscal power illustrates.[263] Knowing only that the government has gathered or used an instrument of power in the abstract does not allow the public to make informed judgments about the government's access to, and use of, those instruments—the extent of the power those instruments confer on the government. Knowing that the government has imposed a military draft, for example, might tell us something about the government's theory of how the costs of military campaigns should be spread across the population, or something about how the government assesses its future military needs. But it cannot itself tell us about the government's military capacity. For that, we need to measure the

---

[261] This would, importantly, reach beyond the requirements of the Paperwork Reduction Act—which instructs agencies to quantify the time burden of complying with new information collections—but not other policy considerations. *See supra* notes 217–19 (describing the Paperwork Reduction Act).

[262] *See* Exec. Order No. 12,291, 46 Fed. Reg. 13,193 (Feb. 17, 1981).

[263] *See supra* notes 145–48 and accompanying text.

products of the draft: the number of boots on the ground and where they have been deployed.

Although data is increasingly seen as an "asset" and a "currency," there is perhaps no contrast more stark than in the systems the government uses to measure its data, on the one hand, and its money, on the other.[264] By constitutional command, the government must render an accounting of its revenues and expenditures.[265] And the long shadow of that requirement has prompted an elaborate system through which the public can obtain detailed and timely knowledge of the size, source, and allocation of the funds government raises and how it uses them to expand governmental capacity.[266] There exists, by contrast, no mechanism that would allow us to measure the government's data assets and how they expand government capacity.[267] We do not know the scale, form, or allocation—at individual or aggregated levels—of the government's data. And we have no standard accounting practices from which government and civil society alike might distill and explain to the public how the government uses its data, so that voters can shape data policy proactively instead of when (and if) they learn of its failures.

But the government is not wholly to blame. There is no generally accepted accounting framework for data in the private sector.[268] Data, in its current form, is a relatively new kind of asset for firms, and firms, no less than the government, are navigating the challenge of measuring and tracking their data stores so that they can exploit them for internal business purposes, value them in market transactions, and comply with government regulations. Because of the growing significance of data to private markets, there are energetic debates among economists and data scientists

---

[264] Memorandum from Russell T. Vought, Acting Director, Off. of Mgmt. & Budget, to Heads of Exec. Dep'ts & Agencies at 1 (June 4, 2019) ("[This] Strategy will enable Government to fully leverage data as a strategic asset."); Fahey, *Data Federalism*, *supra* note 4, at 1072 (calling data an "intergovernmental currency").

[265] *See supra* note 145 and accompanying text.

[266] *See supra* note 147 and accompanying text.

[267] As we elaborate above, the government's fainthearted 2019 effort to "inventory" its databases is not, we think, a meaningful attempt to measure the government's data assets or how they are used in the way that accounting acts as a measure for fiscal assets. *See supra* notes 252–54 and accompanying text.

[268] *See* Keith Atkinson & Ronald McGaughey, *Accounting for Data: A Shortcoming in Accounting for Intangible Assets*, 10 ACAD. ACCT. & FIN. STUD. J. 85, 93 (2006).

(in which one of us is a participant) over the complex problem of how to value data.[269]

Valuing data is difficult for multiple reasons. For one, data's value is highly *contextual*—it is shaped by its fit with the objectives its controllers have for it. It is also, in some respects, a *nonrival* good, one that can be used without being depleted. It is also a *complementary* good: Two pieces of well-matched data can be far more valuable than each is alone, requiring valuation strategies that can accommodate current complementarities and ideally predict future ones. Finally, data is *nonfungible*, in that each piece of data must be independently valued, with its value dependent on nuances about its quality, any alterations that have been made to it, and the integrity of its collection and storage.

Nevertheless, the government could do what it does for money and borrow one of the (preliminary) data valuation strategies under discussion in the private sector. But we are skeptical that the task of measuring data for the purposes of democratic transparency is amenable to private, market-based valuation strategies. Understanding what a government data store might fetch on a private data market does not necessarily convey its value to the *government* or provide even a heuristic sense of how it might enable new or more efficient governmental projects.[270]

We suggest, instead, that a coherent approach to measuring and tracking the government's data for transparency purposes

---

[269] *See generally, e.g.*, Veldkamp, *supra* note 80; Mike Fleckenstein, Ali Obaidi & Nektaria Tryfona, *A Review of Data Valuation Approaches and Building and Scoring a Data Valuation Model*, 5 HARV. DATA SCI. REV., no. 1, 2023; Jian Pei, *A Survey on Data Pricing: From Economics to Data Science*, 34 IEEE TRANSACTIONS ON KNOWLEDGE & DATA ENG'G 4586 (2020); Raul Castro Fernandez, Pranav Subramaniam & Michael J. Franklin, *Data Market Platforms: Trading Data Assets to Solve Data problems*, 13 PROC. VLDB ENDOWMENT 2150 (2020); Carol Corrado, Jonathan Haskel, Massimiliano Iommi & Cecilia Jona-Lasinio, *Measuring Data as an Asset* (Org. for Econ. Coop. & Dev. Econ. Dep't, Working Paper No. 1731, 2022); Dylan G. Rassier, Robert J. Kornfeld & Erich H. Strassner, *Treatment of Data in National Accounts*, U.S. BUREAU OF ECON. ANALYSIS (May 2019), https://perma.cc/K8SE-X9TM; Debbie Salzberger, Nikiforos Iatrou, Gideon Kwinter & Erin Keogh, *Data, Not Data: Uncovering the Implications of Data in Merger Reviews*, 52 U. MEM. L. REV. 969 (2021) (discussing data valuation in the mergers and acquisitions context); Amanda Parsons & Salomé Viljoen, *Valuing Social Data*, 124 COLUM. L. REV. 993 (2024) (discussing data valuation in the taxation context); Omri Marian, *Taxing Data*, 47 B.Y.U. L. REV. 511 (2022) (same).

[270] It is possible to use the same accounting strategies for money in both the public and private sectors because the two are highly porous: The government turns money into state capacity by procuring policy-supporting goods and services on the private market. But the government does not similarly need to sell data to private actors in order to apply it to governmental ends. It can collect and apply data to public projects without market intermediaries and resulting price signals.

should look beyond existing conversations about data valuation. The goal in assessing the size and significance of the government's data assets is (we think) to help the public understand how those assets change and expand governmental capacity, so that voters can guide that capacity to desired objectives and "incapacitate" it when they so desire.[271] And the best evidence of how data shapes government capacity is how government has *in fact* used that data.

That evidence, we think, can be derived from a technique tailored to data and used today for various purposes in data science: what computer scientists call *data provenance*, or data lineage.[272] Think of provenance as a self-surveilling capacity—a body of data about data. Like the provenance of a piece of art, for example, data's provenance includes a record of where that data originated; where it has been since; and who has held, accessed, and altered it.[273] But it can, in theory, also reveal how data has been processed and used: It can log normalization, standardization, and transformation; data sharing and data combinations; data queries; and data's use in algorithms and predictive instruments. It is increasingly possible to reliably record provenance when data remains in a single software system, like a database (or a larger "data lake"). But it remains a significant technical challenge to track provenance as data moves across systems, teams, and institutions, making it a promising path for future research—especially for data scientists interested in governmental application.[274]

Provenance has not, to our knowledge, been previously understood as a tool of public transparency for governmental data. It is most commonly used in data sciences to allow troubleshooting, ensure data validity, and detect data security breaches. Health science research, for example, requires sophisticated data

---

[271] *See* Levinson, *Incapacitating the State*, *supra* note 68, at 182.

[272] *See* Amarnath Gupta, *Data Provenance*, *in* ENCYCLOPEDIA OF DATA SYSTEMS 608, 608 (Ling Liu & M. Tamer Özsu eds., 2009) ("The term 'data provenance' refers to a record trail that accounts for the origin of a piece of data (in a database, document or repository) together with an explanation of how and why it got to the present place.").

[273] *See Provenance*, NAT'L INST. OF STANDS. & TECH., https://perma.cc/4K2B-WXX6 (defining "provenance" as the "chronology of the origin, development, ownership, location, and changes to a system or system component and associated data," which "may also include personnel and processes used to interact with or make modifications to the system, component, or associated data").

[274] *See generally, e.g.*, Neoklis Polyzotis, Sudip Roy, Steven Euijong Whang & Martin Zinkevich, *Data Lifecycle Challenges in Production Machine Learning: A Survey*, 47 SIGMOD REC. 17 (2018); Kiran-Kumar Muniswamy-Reddy, David A. Holland, Uri Braun & Margo Seltzer, *Provenance-Aware Storage Systems*, 2006 USENIX ANN. TECH. CONF. 43 (2006).

provenance—researchers need a record of when a health measurement was collected, from whom, where, and under what conditions; how that data has been stored; who has accessed it; and how it has been matched or aggregated with other data files to verify the integrity of their observations and allow them to be reproduced. Data provenance is also a well-established security mechanism: By tracking who has accessed data, data custodians can know whether any such access has been unauthorized and can make credible representations about the past security of a particular data set.

But data provenance can also be reimagined as a data transparency device in the public sector, facilitating visibility and control over public data. By tracking in granular detail what data government collects, where it is stored, who accesses it, and how it is used, government could develop a genealogy of its data and data-use practices. (Think of both individual uses—hovering over a cell in an Excel spreadsheet and seeing a record of that data's history—and aggregated uses—running a query of government-wide provenance records to locate all data used to train AI or all data held by agencies accessed by the Internal Revenue Service, for example.) Imagining a regime of data transparency rooted in provenance would extend the government's growing data-gathering and data-manipulation capabilities to the project of self-knowledge—offering a data-driven picture of the government's own data capacity.

Since Professor James C. Scott's celebrated book *Seeing Like a State*, the idea that governments produce information about their citizens to render them legible and capable of being governed—and that defects in how states see their citizens can, by the same logic, impede governance projects—has become part of the basic political science lexicon.[275] Reversing the lens, we should think of accounting for the federal fisc, the surveying of public land holdings, and other techniques for measuring the government's instruments of power as tools that help the public to "see like citizens"—that render the *state* capable of being controlled by us. Data provenance, we argue, could become for data what accounting is to money and surveying is to land.

---

[275] *See generally* SCOTT, *supra* note 2.

C. Data Movement Controls

As we describe above, instruments of power are distinctive in part because they are mobile—they can be moved, shared, and aggregated, changing their value, their security, and the projects to which they can be applied. If voters are to assert control over how government uses those instruments, they must be able to control where they are held, how they are allocated, and whether and how they are moved.

To return briefly to money, federal funds are centrally managed by the Department of the Treasury and held in Treasury accounts at the Federal Reserve Bank of New York.[276] This administrative centralization is an essential feature of the government's fiscal ecosystem. The Treasury can exert cohesive control over inflows and outflows of funds from those accounts, systematize and oversee the process of accounting for them, and produce consistent and legible reports for branches and agencies of government, as well as the public. Data, by contrast, is held and managed disjointedly by the hundreds of different government entities and institutions that collect it, with thin measurement and disordered movements. Observers concerned about the concentration of the government's knowledge about individuals—about the advent of a kind of all-knowing surveillance capability—might be tempted to celebrate this arrangement as a form of salutary decentralization. That instinct is a reasonable one because decentralization is an important tool that constitutions use to organize the government's legal power.[277]

But our current data decentralization is less appealing than it might seem at first glance. Although data's custody is decentralized, the underlying data is liquid, movable, and combinable. And unlike instruments of power that are depletable, data's nonrival character allows it to be multiplied and shared without being diminished. An agency can both share its data and retain use of it (in contrast to money or land). Data sharing under conditions of decentralization, therefore, often means data *duplication*. Decentralizing data management while encouraging data sharing can, in short, multiply rather than divide the power data confers

---

[276] *See supra* note 136 and accompanying text.

[277] Indeed, decentralization is the basic design feature of both the separation of powers and federalism. *See* Gregory v. Ashcroft, 501 U.S. 452, 458 (1991) (arguing that the separation of powers vertically and horizontally reduces the "risk of tyranny and abuse").

on government by creating duplicate data stores for each custodian to use at her own access point.[278]

Congress and the OMB, moreover, have long encouraged data sharing subject to the procedural framework articulated in the Privacy Act, which can be bureaucratically irritating (in its emphasis on bilateral negotiation) but is also at bottom legally permissive.[279] And because the value of data is frequently enhanced by its combination with other relevant data, agencies have a strong incentive, in many cases, to share their data with sister agencies.[280]

The FBI, for example, controls a powerful facial recognition database but provides access—in ways that range from clearly disclosed to confidential—to the State Department, the Department of Homeland Security, state and local policing agencies, and other governmental entities, thereby multiplying access points to state power even as it remains at the periphery.[281] It is possible, we think, to technically and legally reorganize data flows across government to calibrate what data and data insights are shared and duplicated; to facilitate greater process, consideration, and popular control over those flows; and to make use of automated techniques for tracking where data is and where it has been.

Moreover, we can do this through a mechanism that is neither a (frighteningly) centralized database of all government data, nor the existing combination of decentralized data custody and haphazard bilateral data sharing. What is needed instead is a standardized roadway of sorts that lets agencies share carefully calibrated forms of data using a trackable, controllable, and politically accountable infrastructure. Here again we can draw on technical literatures in data science, and specifically the robust conversations in data science about secure data-sharing methods.[282] As with data provenance, we can reimagine these tools for the public sector and with enabling popular data controls in mind.

---

[278] *See* Fahey, *Data Federalism*, *supra* note 4, at 1021–22 (developing a similar claim about data sharing across levels of government).

[279] *See supra* notes 238–40; *see also* Memorandum from Jacob J. Lew, Director, Off. of Mgmt. & Budget, to Heads of Exec. Depts & Agencies at 1 (Dec. 20, 2000), ("Agencies should work together to determine what data sharing opportunities are desirable, feasible, and appropriate.").

[280] *See supra* note 237 (theorizing some of the conditions under which agencies may have incentives to share—or withhold—the data they control).

[281] *See supra* note 236 and accompanying text.

[282] *See generally, e.g.*, Yehuda Lindell, *Secure Multiparty Computation*, 64 COMMC'NS ACM 86 (2021) (describing "secure multiparty computation"—a technique that permits multiple parties to perform computations on joint data); Peter Kairouz et al., *Advances and Open Problems in Federated Learning*, 4 FOUNDS. & TRENDS MACH. LEARNING, no. 1–2,

To see how, consider the data-sharing model that one of us has theorized and developed: the *data escrow*. The data escrow is a data system that uses a third-party intermediary to combine data, conduct delegated computation, and release only the agreed-to data products in the agreed-to form.[283] The escrow integrates and builds on other advancements in data sharing.[284]

In the governmental context, the escrow could be a data system managed by an agency like the General Services Administration or a dedicated administrative entity. The data escrow would mitigate several significant risks of the federal government's current data-sharing approach, in which data ownership is decentralized and data access is bilaterally but erratically negotiated.

First, by using an intermediary to combine data, an escrow system permits access by multiple users while avoiding the need to repeatedly duplicate the underlying data. If multiple agencies, for instance, require access to the Department of Health and Human Services' National Directory of New Hires—an extraordinary database that holds quarterly wage data on most working Americans—the intermediary can facilitate that access without duplicating any of the underlying raw data.[285]

Next, the intermediary in an escrow system can perform delegated computation and, in so doing, mitigate security and privacy risks associated with sharing data in the blunt way. Most basically, when two or more agencies need only a subset of one another's data—like data about individuals who appear in both agencies' databases but not about individuals who appear in just one of those databases—delegated computation can perform that matching function and conditionally release only the relevant records. Immigration and Customs Enforcement (ICE) may, for example, wish to consult the FBI-managed National Crime Information Center—which aggregates crime information across hundreds of city, state, and federal jurisdictions—to learn whether a particular noncitizen has committed one of the class of

---

2021 (collecting literature on "federated learning," a technique for training machine learning models on multiple data assets without moving them to a central location); Nikolaj Volgushev, Malte Schwarzkopf, Ben Getchell, Mayank Varia, Andrei Lapets & Azer Bestavros, *Conclave: Secure Multi-Party Computation on Big Data*, 14 EUROSYS CONF., 2019.

283 *See* Xia et al., *supra* note 233, at 1.

284 *See generally id.*; Kairouz et al., *supra* note 282.

285 *See* OFF. OF CHILD SUPPORT SERVS., A GUIDE TO THE NATIONAL DIRECTORY OF NEW HIRES 4–6 (2024) (cataloguing authorized users).

criminal offenses that renders her removable.[286] Rather than giving either the FBI or ICE complete access to the database, the escrow can develop a matched list for conditional release.

Delegated computation can also minimize privacy risks by performing more complex computational functions without transferring (or even providing any person with access to) raw data files. For example, President Joe Biden's Executive Order on Advancing Racial Equity directed agencies to "assess whether, and to what extent, its programs and policies perpetuate systemic barriers to opportunities and benefits for people of color and other underserved groups."[287] Many agencies, however, lack the data to conduct those assessments, and scholars have blamed privacy rights for the failure of the agencies that *do* hold racial data to share it.[288] A data escrow could be programmed to match and conduct analyses on an aggregated database without exposing the underlying data to the privacy exposures associated with interagency data transfers.

A data escrow system would also support the project described above—to track and measure the government's data assets by providing an auditable digital trail of the data movements that are now so opaque. The escrow, for example, was designed in part to allow companies to train machine learning models on data held by a range of different companies without transferring the underlying raw data.[289] If the machine learning models mentioned earlier were trained on government data held in escrow, we would not need the kind of inventory of AI that the government has now undertaken to learn what tools it has already employed.[290]

Finally, a data escrow or similar system could act as an institutional focal point and draw the kind of focused attention to data sharing, access, and combination that the Treasury invites for money. And it could provide a centralized procedural space for cross-agency negotiations over whether to share or withhold data, on what terms, and with what risk tolerances, rather than leaving such decisions to the largely unrecorded bilateral interactions that today characterize data sharing.

---

[286] *See* 8 U.S.C. § 1227(a)(2) (describing the categories of criminal offenses that render a noncitizen eligible for removal); Fahey, *Data Federalism*, *supra* note 4, at 1022–24 (describing the National Crime Information Center).

[287] *See* Exec. Order No. 13,985 § 1, 86 Fed. Reg. at 7,009.

[288] *See supra* note 237.

[289] Xia et al., *supra* note 233, at 1.

[290] *See supra* note 220.

CONCLUSION

This Article has trained a bird's-eye lens on the structures through which the federal government stewards data—the institutions, processes, default rules, movement controls, and accounting measures it uses (or does not use) to facilitate collective control over the data that is increasingly a source of governmental power. The goal of this initial account is to draw new attention to, and interest in, thinking structurally about data and in bringing public data back within public control.

But structural thinking about data is also relevant beyond federal constitutional law and the problem of democratic control. Data federalism, as one of us has previously written, inverts the structural logic of U.S. federalism and continues to be a zone of significant intergovernmental fights. In international relations, structural thinking about data will be necessary to resolve difficult questions about data sovereignty—a nation's normative claim to regulate data connected to its people or territory—which requires the legal and technical importation of national boundaries into data networks that lack them. And "indigenous data sovereignty" has become an important component of self-definition and self-determination, one that requires a reimagining of the structural distribution of data control among the federal government, states, and native nations. As data's capacity grows, so too will the need for new legal and technical tools to structure the power it confers on all levels and forms of government.