GENDER DATA IN THE AUTOMATED ADMINISTRATIVE STATE

*Ari Ezra Waldman**

In myriad areas of public life—from voting to professional licensure—the state collects, shares, and uses sex and gender data in complex algorithmic systems that mete out benefits, verify identity, and secure spaces. But in doing so, the state often erases transgender, nonbinary, and gender-nonconforming individuals, subjecting them to the harms of exclusion. These harms are not simply features of technology design, as others have ably written. This erasure and discrimination are the products of law.

This Article demonstrates how the law, both on the books and on the ground, mandates, incentivizes, and fosters a particular kind of automated administrative state that binarizes gender data and harms gender-nonconforming individuals as a result. It traces the law's critical role in creating pathways for binary gender data, from legal mandates to official forms, through their sharing via intergovernmental agreements, and finally to their use in automated systems procured by agencies and legitimized by procedural privacy law compliance. At each point, the law mandates and fosters automated governance that prioritizes efficiency rather than inclusivity, thereby erasing gender-diverse populations and causing dignitary, expressive, and practical harms.

In making this argument, the Article challenges the conventional account in the legal literature of automated governance as devoid of

* Professor of Law, University of California, Irvine School of Law; Professor of Sociology, by courtesy, University of California, Irvine. Ph.D., Columbia University; J.D., Harvard Law School; A.B., Harvard College. Affiliated Fellow, Yale Law School Information Society Project. Thanks to the following scholars who offered helpful comments on multiple drafts (in alphabetical order): Victoria Adelmant, Kendra Albert, Chinmayi Arun, Jack Balkin, Sebastian Benthall, Elettra Bietti, danah boyd, Ryan Calo, Danielle Citron, Julie Cohen, Sam Cournoyer, Brenda Dvoskin, Woodrow Hartzog, Anna Lauren Hoffmann, Gordon Hull, Rebecca Ingber, Thomas Kadri, Tomer Kenneth, Aniket Kesari, Mehtab Khan, Karen Levy, Daniel Medwed, Ngozi Okidegbe, Jeremy Paul, Jon Penney, Neil Richards, Hilary Robinson, Blaine Saito, Morgan Klaus Scheuerman, Paul Schwartz, Jessica Silbey, Scott Skinner-Thompson, Alicia Solow-Niederman, Katherine Strandburg, Thomas Streinz, Karen Tani, and Kate Weisburd. Thank you also to participants in the 2023 Privacy Law Scholars Conference, New York University School of Law's Privacy Research Group, the Yale Law School Information Project, and faculty workshops and talks at Northeastern University, Texas A&M University School of Law, Temple University Beasley School of Law, University of California, Los Angeles School of Law, and University of California, Irvine School of Law, all of whom provided outstanding feedback. Special thanks to AJ Jarrett (they/them) and Maya Bornstein (she/her) for essential research assistance. None of the errors in this Article are automated; they are my own.

discretion, as reliant on technical expertise, and as the result of law stepping out of the way. It concludes with principles for reforming the state's approach to sex and gender data from the ground up, focusing on privacy law principles of necessity, inclusivity, and antisubordination.

INTRODUCTION

Sasha Costanza-Chock triggered the alarm when they walked through the full-body scanner at the Detroit Metro Airport.[1] They knew it would happen because it happens to transgender, nonbinary, and gender-nonconforming people all the time.[2] The machine deemed Sasha "risky" because their body, datafied into machine-readable code, differed from the pictures of bodies that trained the machine's algorithm.[3] Their breasts were too pronounced relative to data associated with "male," and their groin area deviated from data associated with "female."[4] Pulled out of the line for a physical body search, Sasha found themself in an awkward, humiliating, and potentially dangerous situation.

Toby P., a transgender man living in Colorado, was singled out by a different kind of automated administrative technology.[5] After Toby sustained a debilitating injury at work, his employer completed the required workers' compensation First Report of Injury Form by checking the box next to "Female," a designation that matched Toby's assigned sex at birth and the information in his human resources file.[6] The state's

---

1. Sasha Costanza-Chock, Design Justice, A.I., and Escape From the Matrix of Domination, J. Design & Sci. (July 16, 2018), https://doi.org/10.21428/96c8d426 [https://perma.cc/E2M3-WGW5] [hereinafter Costanza-Chock, Design Justice]; see also About, Sasha Costanza-Chock, Ph.D., https://www.schock.cc/?page_id=13 [https://perma.cc/JEQ3-JELT] (last visited Aug. 21, 2023).

2. See, e.g., Deema B. Abini, Traveling Transgender: How Airport Screening Procedures Threaten the Right to Informational Privacy, 87 S. Cal. L. Rev. Postscript 120, 135 (2014); Paisley Currah & Tara Mulqueen, Securitizing Gender: Identity, Biometrics, and Transgender Bodies at the Airport, 78 Soc. Rsch. 557, 562–66 (2011); Dawn Ennis, Her Tweets Tell One Trans Woman's TSA Horror Story, Advocate (Sept. 22, 2015), https://www.advocate.com/transgender/2015/9/22/one-trans-womans-tsa-horror-story [https://perma.cc/5FZS-6NKV]. For detailed definitions of "transgender," "nonbinary," "gender-nonconforming," and related terms, please see Jessica A. Clarke, They, Them, and Theirs, 132 Harv. L. Rev. 894, 897–99 (2019); Glossary of Terms: LGBTQ, GLAAD, https://www.glaad.org/reference/terms [https://perma.cc/7BHP-6Y2T] (last visited Aug. 21, 2023). In brief, transgender individuals are those whose sense of self or expression of their gender differs from their assigned sex at birth. Nonbinary individuals are those whose identities cannot be restricted to just "male" or "female." "Gender-nonconforming" is an umbrella term that can include nonbinary individuals, but it is used in this Article to refer to those who are genderqueer (those who challenge norms concerning sex, gender, and sexuality), genderfluid (those whose gender expressions or identities may change over time), or agender (those who do not adopt a traditional gender category and may describe their gender as the lack of one).

3. Costanza-Chock, Design Justice, supra note 1.

4. Id.

5. Toby's name has been changed to protect his anonymity as he and his lawyers determine how to proceed with a potential claim against the state.

6. Telephone Interview with Toby P. (May 22, 2022) (notes on file with the *Columbia Law Review*); Colo. Dep't of Lab., WC 1, Employer's First Report of Injury (2006), https://codwc.app.box.com/v/wc1-first-report-injury (on file with the *Columbia Law Review*).

automated fraud-detection system, which compares this claim form with information pooled from state databases, denied Toby's claim. The "system," Toby told me, "saw 'female' here and 'male' [everywhere else] . . . and figured something didn't match."[7] Seven months, twenty-five phone calls, sixteen refiled forms, and two demand letters later, Toby is still hurt and still without the compensation to which he is entitled. He is "basically bankrupt."[8]

Sasha and Toby fell through the cracks of the automated administrative state.[9] As government agencies turn to algorithms and artificial intelligence (AI) to administer benefits programs, detect fraud, and secure spaces, transgender, nonbinary, and gender-nonconforming individuals are put in situations where they can't win. They become "anomalies" or "deviants" in systems designed for efficiency.[10]

Technologies "have politics."[11] Just like race and gender hierarchies can be embedded into technological systems,[12] in this case it is

---

7. Telephone Interview with Toby P., supra note 6.

8. Id.

9. This Article uses the phrase "automated decisionmaking system" or "algorithmic decisionmaking system" to refer to the overall process in which a computational mechanism uses data inputs to make probabilistic, predictive conclusions or implements policy by software. See Ryan Calo, Artificial Intelligence Policy: A Primer and Roadmap, 51 U.C. Davis L. Rev. 399, 404–05 (2017) (noting that there is no one "consensus definition of artificial intelligence" but clarifying ways of understanding what scholars and industry mean by AI). This simplification is intentional: The Article focuses on the law's responsibility for trends in automation rather than the technical distinctions between different types of automated technologies. See AI Now Inst., Confronting Black Boxes: A Shadow Report of the New York City Automated Decision System Task Force 7 (Rashida Richardson ed., 2019), https://ainowinstitute.org/publication/confronting-black-boxes-a-shadow-report-of-the-new-york-city-automated [https://perma.cc/2K5X-GB3A] (defining algorithmic or automated decisionmaking systems as "data-driven technologies used to automate human-centered procedures, practices, or policies for the purpose of predicting, identifying, surveilling, detecting, and targeting individuals or communities").

10. See Toby Beauchamp, Going Stealth: Transgender Politics and U.S. Surveillance Practices 35–37 (2019); Sonia K. Katyal & Jessica Y. Jung, The Gender Panopticon: AI, Gender, and Design Justice, 68 UCLA L. Rev. 692, 710–11 (2021) (explaining that identity detection as a form of biometric surveillance treats some individuals as "anomalies" or outliers when they do not conform to gender binaries).

11. Langdon Winner, Do Artifacts Have Politics?, Dædalus, Winter 1980, at 121, 121 (explaining that technology embodies forms of power and authority).

12. There is a vast literature in this space. See, e.g., Safiya Umoja Noble, Algorithms of Oppression: How Search Engines Reinforce Racism (2018) (explaining how digital decisions made through systemic algorithms reinforce oppressive social relationships); Sarah Myers West, Meredith Whittaker & Kate Crawford, Discriminating Systems: Gender, Race, and Power in AI 8–9 (2019), https://ainowinstitute.org/wp-content/uploads/2023/04/discriminatingsystems.pdf [https://perma.cc/A4YD-UPPG] (outlining research findings that the AI sector has a lack of diversity among its professionals, which has led to discriminatory outcomes); Solon Barocas & Andrew D. Selbst, Big Data's Disparate Impact, 104 Calif. L. Rev. 671, 674–77 (2016) [hereinafter Barocas & Selbst, Big Data's Disparate Impact] (outlining various reports that have suggested "big data" has unintended

cisnormativity—the assumption that everyone's gender identity and presentation accord with their assigned sex at birth—that is designed into the automated systems that singled out Sasha and Toby. The underlying data that train machines to recognize males and females, the algorithms that identify anomalies in a person's body relative to that database, the forms inconsistently designed to collect sex and gender data in the first place, and the systems' restriction to only male/female options all reflect assumptions of gender as binary. Anyone who deviates from a normative, binary body is "risky" and singled out, potentially exposing them to harm. Those gender-nonconforming individuals who are also religious minorities, immigrants, people of color, or people with disabilities, and people who hold more than one minoritized identity, are multiply burdened.[13]

But this Article is not simply about the biases replicated and entrenched by AI and algorithmic technologies, a story deftly told by others and summarized in Part I. Nor is it just about gender as a tool of classification, a story as old as the nation.[14] This is a story about law. Specifically, this Article argues that the law has mandated, influenced, and guided the state to automate in a way that binarizes gender data, thereby erasing and harming transgender, nonbinary, and gender-nonconforming individuals.

The law's active role in the creation of this kind of automated state has been overlooked because the two dominant strands in legal scholarship on algorithmic technologies are focused elsewhere. One of

---

discriminatory effects); Joy Buolamwini & Timnit Gebru, Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification, 81 Proc. Mach. Learning Rsch. 1, 10–11 (2018), https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf [https://perma.cc/Q5VD-EF9F] (detailing how machine-learning technology can produce disastrous results in high-stakes circumstances, specifically when used in criminal matters); Pauline T. Kim, Data-Driven Discrimination at Work, 58 Wm. & Mary L. Rev. 857, 874–90 (2017) (describing how "training data," or data used to inform machines running algorithms, are often unknowingly infected with bias, creating discriminatory results that are especially harmful in the workplace). In a recent article, Professor Sonia Katyal and healthcare industry lawyer Jessica Jung focus almost entirely on the gender and racial biases of algorithmic technologies used by private, for-profit companies. Katyal & Jung, supra note 10. This Article adds to this literature with a different narrative, focusing on government uses of automated technology and the mostly underappreciated laws that are responsible for collecting and entrenching binary gender in government systems.

13. See, e.g., Patricia Hill Collins, Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment 221–38 (1990) (describing how minoritized populations experience oppression and domination on multiple levels); Kimberlé Crenshaw, Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color, 43 Stan. L. Rev. 1241, 1250–52 (1991) (outlining how all intersections of race and gender affect the social construct of identity).

14. See Gérard Noiriel, The Identification of the Citizen: The Birth of Republican Civil Status in France, *in* Documenting Individual Identity 28, 30–42 (Jane Caplan & John Torpey eds., 2001).

those strands sees automation and its harms flourishing in a regulatory void. Scholarship in this vein rightly argues that automated systems used by private, for-profit technology companies cause harm because "the law has offered insufficient protection."[15] Other scholars suggest that algorithmic technologies are built amidst "lawlessness," or the lack of regulation.[16]

A second important strand of law and technology scholarship focuses on how law can address automation's harms. This research explores how the technologies work, where they go wrong, and how we might use law to regulate them, fix them, and restore the status quo ex ante by holding technologies and those that use them accountable for discrimination, bias, and harm.[17] Few scholars have focused on how the law *creates* the

---

15. See Katyal & Jung, supra note 10, at 704 ("[G]ender panopticism has been facilitated by absences within privacy law, in that the law has offered insufficient protection to gender self-determination and informational privacy."); see also id. at 723, 760–61 (outlining forms of biometric surveillance technology that render nonbinary individuals outliers).

16. Shoshana Zuboff, The Age of Surveillance Capitalism 127–28 (2019). But see, e.g., Julie Cohen, Between Truth and Power 3 (2019) [hereinafter Cohen, Between Truth and Power] (arguing that informational capitalism itself is a construct of opportunistic economic actors using law to control the means of informational production); Amy Kapczynski, The Law of Informational Capitalism, 129 Yale L.J. 1460, 1465 (2020) (reviewing both texts); see also Bridget Fahey, Data Federalism, 135 Harv. L. Rev. 1007, 1013–14, 1036–39 (2022) [hereinafter Fahey, Data Federalism] (highlighting the "absence" of "major federal legislation" as one reason for rampant, unregulated data sharing among state agencies but noting the role of interagency agreements and other more informal legal instruments).

17. E.g., Dillon Reisman, Jason Schultz, Kate Crawford & Meredith Whittaker, Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability (2018), https://openresearch.amsterdam/image/2018/6/12/aiareport2018.pdf [https://perma.cc/Y3YY-BSTG]; Barocas & Selbst, Big Data's Disparate Impact, supra note 12; Danielle Keats Citron & Frank Pasquale, The Scored Society: Due Process for Automated Predictions, 89 Wash. L. Rev. 1 (2014); Danielle Keats Citron, Technological Due Process, 85 Wash. U. L. Rev. 1249 (2008) [hereinafter Citron, Technological Due Process]; Ignacio N. Cofone, Algorithmic Discrimination Is an Information Problem, 70 Hastings L.J. 1389 (2019); Kate Crawford & Jason Schultz, Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms, 55 B.C. L. Rev. 93 (2014); A. Michael Froomkin, Ian Kerr & Joelle Pineau, When AIs Outperform Doctors: Confronting the Challenges of a Tort-Induced Over-Reliance on Machine Learning, 61 Ariz. L. Rev. 33 (2019); James Grimmelmann & Daniel Westreich, Incomprehensible Discrimination, 7 Calif. L. Rev. Online 164 (2017), https://lawcat.berkeley.edu/record/1128018/files/GrimmelmannWestreich.final_.pdf [https://perma.cc/7QMW-AEDQ]; Meg Leta Jones, The Right to a Human in the Loop: Political Constructions of Computer Automation and Personhood, 47 Soc. Stud. Sci. 216 (2017); Margot E. Kaminski, Binary Governance: Lessons From the GDPR's Approach to Algorithmic Accountability, 92 S. Cal. L. Rev. 1529 (2019); Sonia K. Katyal, Private Accountability in the Age of Artificial Intelligence, 66 UCLA L. Rev. 54 (2019) [hereinafter Katyal, Private Accountability]; W. Nicholson Price II, Regulating Black-Box Medicine, 116 Mich. L. Rev. 421 (2017); Andrew D. Selbst & Solon Barocas, The Intuitive Appeal of Explainable Machines, 87 Fordham L. Rev. 1085 (2018); Alicia Solow-Niederman, Administering Artificial Intelligence, 93 S. Cal. L. Rev. 633 (2020).

automated administrative state,[18] and fewer still have focused on how the law constructs gender data in the automated state.[19] This Article fills that gap: Sasha's and Toby's stories are actively and indelibly framed, constructed, and sustained by law every step of the way.

The process begins at the source, where statutes mandate the collection of sex and gender data. As Part II describes, the law of gender data collection relies on assumptions of static gender, taps into uninformed perceptions of the gender binary as "common sense," and creates the conditions for civil servants to design forms with primarily binary gender questions. This creates binary gender data streams. Part III shows how interstate compacts and interagency contracts, all of which I collected from public records requests, require states to share datasets that include sex and gender. The law of gender data sharing looks outward and inward to privilege the gender binary: It has expressive effects that normalize the gender binary, conflationary effects that confuse the social aspects of gender with the biological aspects of sex, and interoperability effects that force the gender binary onto any agency that wants to realize the benefits of participating in shared data systems. Part IV demonstrates how automation mandates, agency policymaking by procurement, trade secrecy law, and privacy and data protection law actively encourage automation to improve efficiencies while preventing anyone from interrogating the underlying assumptions of the algorithms that use sex and gender data. This web of legal rules guides automation to exclude those outside the norm and erects barriers around automated tools that protect the gender binary from change.[20] In other words, the law forces an oversimplified legibility on its subjects, leaving those most marginalized at risk.[21]

---

18. But see Cohen, Between Truth and Power, supra note 16, at 48–74 (exploring the ways law, actively leveraged by interested economic actors, has created a "zone of legal privilege" around the activities of data-driven technologies); Alicia Solow-Niederman, YooJung Choi & Guy Van den Broeck, The Institutional Life of Algorithmic Risk Assessment, 34 Berkeley Tech. L.J. 705, 705–08 (2019) (arguing that risk assessment statutes create frameworks that constrain and empower policymakers and technical actors when it comes to the design and implementation of a particular instrument).

19. Of course, there has been scholarship on gender as a tool of administrative governance. See, e.g., Dean Spade, Normal Life: Administrative Violence, Critical Trans Politics, & The Limits of Law 73–93 (2015) [hereinafter Spade, Normal Life]. But this scholarship has not extended to consider the effects of algorithms and automation in the administrative state.

20. See Cohen, Between Truth and Power, supra note 16, at 49 (referring to how the law creates "zone[s] of legal privilege" around information-driven business models).

21. For how governments force this legibility on their subjects, see generally James C. Scott, Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed (1998) [hereinafter Scott, Seeing Like a State] ("[T]he legibility of a society provides the capacity for large-scale social engineering, high-modernist ideology provides the desire, the authoritarian state provides the determination to act on that desire, and an incapacitated civil society provides the leveled social terrain on which to build.").

This rich account of how law collects, shares, and uses sex and gender data in state-run automated systems offers several insights about automation and the automated state in general that challenge or add nuance to the conventional wisdom in the legal literature. Part V discusses four of those lessons.

The automated state is *discretionary*.[22] Scholars have argued that automation erodes traditional agency discretion, a pillar of the administrative state.[23] But this Article shows that civil servants have discretion to guide automation in ways that binarize gender data. The discretion may be buried, but its fingerprints are everywhere—in the design of data-collection forms, in the terms of data-sharing agreements, in the procurement of technologies, and in the design and completion of privacy impact assessments (PIAs).[24] Relatedly, the automated state is also driven by *stereotypes*.[25] Rather than merely shifting expertise from civil servants hired for their substantive knowledge to engineers with technological knowledge about how algorithms work, the automated state relies on both civil servants' and engineers' supposedly commonsense perceptions of sex and gender.[26] Because most people have traditionally presumed that sex and gender are the same and static, automated systems designed by engineers and used by the government reflect those stereotypes.

The automated state is also *managerial*.[27] Far from a product of the law stepping out of the way, the state's use of algorithmic decisionmaking processes represents the synthesis of the logics (and pathologies) of data-driven governance, risk assessment, public–private partnerships, and procedural compliance, leveraging the power of law and the state to achieve efficiency goals. By orienting algorithmic tools toward the neoliberal goal of targeted governance through risk assessments that are supposed to cover most people most of the time, the law singles out those outside the norm for disproportionate harm. Finally, and again, relatedly, the automated state is structurally *subordinating*.[28] Law infuses the

22. See infra section V.A.

23. See, e.g., Ryan Calo & Danielle Keats Citron, The Automated Administrative State: A Crisis of Legitimacy, 70 Emory L.J. 797, 804 (2021).

24. Impact assessments in the law and technology space document development rationales for new technologies and are supposed to keep certain values like privacy and fairness front of mind for those developing and using the technologies. See Andrew D. Selbst, An Institutional View of Algorithmic Impact Assessments, 35 Harv. J.L. & Tech. 117, 122 (2021). But see Ari Ezra Waldman, Industry Unbound 132–33 (2021) [hereinafter Waldman, Industry Unbound] (describing how impact assessments can be reduced to mere checkbox compliance).

25. See infra section V.B.

26. See infra section V.B.

27. See infra section V.C.

28. See infra section V.D.

government's data ecosystem with sex and gender information in a way that is both over- and underinclusive: It is overinclusive because it collects sex and gender data too often when not necessary; it is underinclusive because its reliance on the gender binary excludes transgender, nonbinary, and gender-nonconforming individuals from any of the benefits that could come from data's capacity to create insight.

This kind of automated state harms gender-diverse populations. But the reification of the gender binary in the automated state is not a niche concern; it harms anyone constrained by strict gender expectations.[29] Plus, those most dependent on government resources and thereby subject to the state's informational demands will bear the greatest burdens of the state's automated use of binary gender data streams.[30] This poses a particular problem for members of the LGBTQ+ community, approximately one million of whom are on Medicaid.[31] Nearly half of LGBT people of color live in low-income households.[32] Transgender people are nearly two and a half times more likely than non-transgender people to face food insecurity.[33] LGBT people have higher rates of unemployment than the general population.[34]

For some scholars and advocates, the solution to these problems is for the state to stop collecting sex and gender data.[35] But as various scholars

---

29. Feminist scholars have long argued that discrimination on the basis of gender nonconformity should be redressable. See, e.g., Mary Anne C. Case, Disaggregating Gender From Sex and Sexual Orientation: The Effeminate Man in the Law and Feminist Jurisprudence, 105 Yale L.J. 1, 2–4 (1995); Katherine M. Franke, The Central Mistake of Sex Discrimination Law: The Disaggregation of Sex From Gender, 144 U. Pa. L. Rev. 1, 3–5 (1995); Vicki Schultz, Reconceptualizing Sexual Harassment, 107 Yale L.J. 1683, 1774–88 (1998).

30. Cf. Khiara M. Bridges, The Poverty of Privacy Rights 9 (2017) [hereinafter Bridges, Poverty] ("[P]oor mothers have *traded* [their privacy] for a welfare benefit.").

31. See Kerith J. Conron & Shoshana Goldberg, Over Half a Million LGBT Adults Face Uncertainty About Health Insurance Coverage Due to HHS Guidance on Medicaid Requirements 1 (2018), https://williamsinstitute.law.ucla.edu/wp-content/uploads/LGBT-Medicaid-Coverage-US-Jan-2018.pdf [https://perma.cc/H7Q3-JS7X].

32. Bianca D.M. Wilson, Lauren Bouton & Christy Mallory, Racial Differences Among LGBT Adults in the US 2 (2022), https://williamsinstitute.law.ucla.edu/wp-content/uploads/LGBT-Race-Comparison-Jan-2022.pdf [https://perma.cc/3RYL-4XK7].

33. Kerith J. Conron & Kathryn K. O'Neill, Food Insufficiency Among Transgender Adults During the COVID-19 Pandemic 5 (2022), https://williamsinstitute.law.ucla.edu/wp-content/uploads/Trans-Food-Insufficiency-Update-Apr-2022.pdf [https://perma.cc/G5HE-RSYV].

34. Richard J. Martino, Kristen D. Krause, Marybec Griffin, Caleb LoSchiavo, Camilla Comer-Carruthers & Perry N. Halkitis, Employment Loss as a Result of COVID-19: A Nationwide Survey at the Onset of COVID-19 in US LGBTQ+ Populations, 19 Sexuality Rsch. & Soc. Pol'y 1855, 1860 (2022).

35. See, e.g., Lila Braunschweig, Abolishing Gender Registration: A Feminist Defence, 1 Int'l J. Gender Sexuality & L. 76, 86 (2020); Davina Cooper & Flora Renz, If the State Decertified Gender, What Might Happen to Its Meaning and Value?, 43 J.L. & Soc'y 483, 484 (2016); Ido Katri, Transitions in Sex Reclassification Law, 70 UCLA L. Rev. 636, 641

have shown, legibility comes with benefits as well as risks.[36] I don't know whether there is a way to get it right, to find the "Goldilocks Zone" for gender, data, and power, especially given the state's historic commitment to queer oppression and the historical aims of what James C. Scott might call top-down legibility.[37] But I would like to try. This Article offers a way to navigate the legibility dilemmas triggered by state gender data collection.

The Article's lessons about the automated state—its persistent reliance on civil servant discretion, its use of stereotypes and perceptions of common sense, its orientation toward efficiency, and its subordinating capacities—suggest that scholars and advocates ignore the liminal space between the law on the books and the law on the ground to our peril.[38] For sure, we can pass new laws that guarantee an "X" gender marker option; we can also litigate in court when state gender designations discriminate against those outside the gender binary. But "new categories

---

(2023); Anna James (AJ) Neuman Wipfler, Identity Crisis: The Limitations of Expanding Government Recognition of Gender Identity and the Possibility of Genderless Identity Documents, 39 Harv. J.L. & Gender 491, 543 (2016).

36. See Clarke, supra note 2, at 990 (noting the contextual need for the state to recognize gender diversity in some circumstances); Dean Spade, Documenting Gender, 59 Hastings L.J. 731, 814–15 (2008) [hereinafter Spade, Documenting Gender] (suggesting that the state should continue to collect gender data in the public health context). In the context of racial data, see, e.g., Melissa Nobles, Shades of Citizenship: Race and the Census in Modern Politics, at xi (2000) (arguing that racial data and racial enumeration by censuses advance concepts of race); Clara E. Rodríguez, Changing Race: Latinos, the Census, and the History of Ethnicity in the United States, at xiii (2000) (discussing the need for governmental race data to address past discrimination as balanced against the effect race data have on reification and racial identity); Cassius Adair, Licensing Citizenship: Anti-Blackness, Identification Documents, and Transgender Studies, 71 Am. Q. 569, 570 (2019) (discussing race markers on identification documents in American history and the movement to abolish their use); Nancy Leong, Judicial Erasure of Mixed-Race Discrimination, 59 Am. U. L. Rev. 469, 491–92 (2010) (describing activism in support of adding a multiracial category to the census); Naomi Mezey, Erasure and Recognition: The Census, Race and the National Imagination, 97 Nw. U. L. Rev. 1701, 1713–22 (2003) (evaluating the paradoxical nature of racial classification in the census given the tension between the government's power to recognize and its power to discipline); Nathaniel Persily, Color by Numbers: Race, Redistricting, and the 2000 Census, 85 Minn. L. Rev. 899, 903 (2001) (discussing the importance of census racial data accuracy for minority electoral representation); Naomi Zack, American Mixed Race: The U.S. 2000 Census and Related Issues, 17 Harv. BlackLetter L.J. 33, 35–37 (2001) (discussing the importance of the introduction of mixed-race identification in the 2000 Census but also identifying continuing problems with governmental classification).

37. See Scott, Seeing Like a State, supra note 21, at 65–73. On the state's orientation toward queer oppression, see generally George Chauncey, Gay New York: Gender, Urban Culture, and the Making of the Gay Male World, 1890–1940 (1994); Jonathan Ned Katz, The Invention of Heterosexuality (2007). On legibility, see Scott, Seeing Like a State, supra note 21, at 65–73.

38. This is known as "gap studies" in the sociolegal literature, and this Article is situated in that intellectual tradition. See Jon B. Gould & Scott Barclay, Mind the Gap: The Place of Gap Studies in Sociolegal Scholarship, 8 Ann. Rev. L. & Soc. Sci. 323, 324 (2012).

are not enough."[39] Nor will a statute "deprogram" a gender binary so embedded in our culture and in the technologies of private and state surveillance.[40] To protect transgender, nonbinary, and gender-nonconforming individuals from automation-based harms on a more systematic level, we can also develop the state's "gender competence."[41] That is, in addition to changing the law on the books, scholars and advocates can also help change how civil servants understand gender data and its value, limits, and powers.

These are the goals of Part VI, which wrestles with the live and pressing questions of the proper role of the state: Should the state ever collect and use gender data? If not, why? If so, how can the state do so in a way that serves the interests of gender-diverse populations rather than its own disciplinary interests? Resolving these questions is beyond the scope of this Article, but in a world in which the state does collect and use gender data, its role should be particularly narrow. Part VI offers three principles, familiar to privacy scholars, for building a future in which government uses of gender data and algorithmic technology foster rather than erode antisubordination goals. A *necessity* principle urges the state to ask whether it actually needs sex or gender data to achieve its goals and, if it does, to determine which one it needs. An *antisubordination* principle would limit sex and gender data collection to only those uses that benefit and support greater inclusion of gender-diverse populations. And an *inclusivity* principle would ensure that once the state decides to collect sex or gender data for emancipatory ends, it does so sensitively and in a contextually inclusive way.

Luckily, privacy law principles of data minimization—that one should only collect as much personal data as is necessary to achieve a stated purpose—and antisubordination—that law should disrupt traditional hierarchies of power enjoyed by data collectors—are capable of doing just that.[42] Part VI concludes with this Article's ultimate recommendation: The law on the books and the law on the ground should take gender diversity into account. The state should be able to collect, share, and use sex and gender data only when necessary to support a gender-inclusive

---

39. Laurel Westbrook & Aliya Saperstein, New Categories Are Not Enough: Rethinking the Measurement of Sex and Gender in Social Surveys, 29 Gender & Soc'y 534, 535–36 (2015).

40. See Rena Bivens, The Gender Binary Will Not Be Deprogrammed: Ten Years of Coding Gender on Facebook, 19 New Media & Soc'y 880, 895 (2017).

41. Kevin Guyan, Queer Data: Using Gender, Sex and Sexuality Data for Action 155 (2022).

42. Scott Skinner-Thompson, Privacy at the Margins 6 (2021) (noting that an antisubordination agenda requires consciousness of classifications and using them to "level up" those disadvantaged by traditional hierarchies of power); Spiros Simitis, Reviewing Privacy in an Information Society, 135 U. Pa. L. Rev. 707, 740 (1987) ("Personal information should only be processed for unequivocally specified purposes. Both government and private institutions should abstain from collecting and retrieving data merely for possible future uses for still unknown purposes.").

antisubordination agenda: to combat discrimination, to provide adequate healthcare, to guarantee benefits that have been traditionally denied, and to enable self-determination for gender-diverse populations.

To date, the law's role in creating an automated state that binarizes gender data has been mostly hidden from view. It is a puzzle of statutes, rules, interstate compacts, intergovernmental cooperation, procurement, street-level bureaucracy, and managerial policymaking, all of which is summarized in Table 1. This Article pieces that puzzle together. It relies on a mix of primary source materials, including a computationally derived novel dataset of more than 12,000 government forms scraped from state agency websites, documents obtained through public record requests, and first-person interviews with lawyers and government officials.

TABLE 1. LAW AND THE BINARIZATION OF GENDER DATA, SUMMARY

Law of Data Collection (examples)[43]	**Data binarized by . . .**
	Mediation by the state, which creates the data.
Statutes requiring sex/gender data collection (e.g., security, identity verification, distribution of benefits).	*Perceptions* of "common sense" about sex/gender, which govern form design.
Information primarily gathered through forms created by street-level bureaucrats.	*Path dependencies*, which ensure that forms remain the same over time.
	Assumption that gender is a static/secure identifier, which implies gender binary only.
Law of Data Sharing[44]	**Data binarized by . . .**
Data sharing required to realize security and efficiency benefits.	*Normalization* of the binary by dissemination.
Data sharing permitted at discretion of state agency leadership.	*Conflation* of sex and gender.
Interagency agreements.	*Interoperability*, which requires all data look to the same.
Interstate compacts.	

43. See infra Part II.

44. See infra Part III.

Law of Data Use[45]	Data binarized by . . .
Automation mandates.	*Efficiency* mandates, which mean binary design.
Efficiency mandates.	
Innovation, chief innovation offices.	*Managerialization* via innovation offices, which ensures narrow cost–benefit analysis.
Procurement.	*No interrogation* of design via procurement process.
Trade secrecy.	
Privacy law compliance (privacy impact assessments).	*Symbolic compliance*, which weaponizes PIAs to serve automation rather than privacy.

I. AUTOMATED ADMINISTRATIVE TECHNOLOGY AND ITS HARMS

In today's automated administrative state, algorithmic technologies offer governments new opportunities for gender-based classifications. Professor Sonia Katyal and healthcare industry lawyer Jessica Jung argue in the context of private, for-profit uses of algorithms and AI, anti-transgender bias and erasure are designed into these tools.[46] That is in line with the conventional account in much of the legal literature on algorithmic discrimination, which focuses primarily on technology's capacity to entrench historical racial and gender biases.[47] This Part briefly recounts that conventional account, focusing on how the design of algorithmic technologies used by the automated administrative state erases and causes harm to gender-diverse populations.

A. *Technologies in the Automated State*

Automated systems will sometimes use gender to apply rules in practice, like meting out benefits.[48] Other technologies use gender as data points in data-matching systems and as training data for data-mining systems. Data-matching systems compare two sets of data—for example,

45. See infra Part IV.

46. Katyal & Jung, supra note 10, at 700–01 (arguing that "invisibility" is the result of how AI and algorithmic technologies are built and function).

47. See supra note 17.

48. See Citron, Technological Due Process, supra note 17, at 1268.

demographic data provided on an application for unemployment benefits and a database with the applicant's motor vehicle records, voter registration, and information from private brokers—to determine if both datasets represent the same person.[49] If one or more data points do not match, the system flags the applicant as risky or fraudulent. This is what happened to nearly 50,000 people who applied for unemployment insurance in Michigan, which introduced an automated fraud-detection system in 2013.[50] The problem was that few of them actually committed fraud.[51] When the comparison data is incorrect or outdated, as was the case in Michigan, data-matching systems flag fraud where there is none.[52] In Michigan, the error caused profound harm. The state garnished wages and withdrew money from people's bank accounts, money that many victims are still trying to get back.[53]

Toby was harmed by a data-matching system. Fraud-detection software compared data on the employer's forms with data about Toby in state databases. Because those data did not match, Toby was accused of fraud. Sasha, on the other hand, was the victim of another cluster of algorithmic decisionmaking tools that use gender data—namely, data-mining systems.[54]

Data mining uses gender information as training data to "teach" an algorithm to find patterns and correlations in large datasets.[55] The algorithm then makes probabilistic predictions about the future.[56] For example, in the private commercial space, Amazon's recommendation algorithm mines our prior purchases, browser history, and latent characteristics to predict what we might buy next.[57] Google's search algorithm combines internet-wide data with information about our

---

49. Id. at 1260.

50. See Cahoo v. SAS Analytics Inc., 912 F.3d 887, 892 (6th Cir. 2019) (describing the faulty data-matching algorithm that caused the false determinations of fraud).

51. See Calo & Citron, supra note 23, at 827–29; Robert N. Charette, Michigan's MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold, IEEE Spectrum (Jan. 24, 2018), https://spectrum.ieee.org/riskfactor/computing/software/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold (on file with the *Columbia Law Review*).

52. Charette, supra note 51.

53. Calo & Citron, supra note 23, at 828–29.

54. See Citron, Technological Due Process, supra note 17, at 1260.

55. Solow-Niederman, supra note 17, at 639.

56. Morgan Klaus Scheuerman, Jacob M. Paul & Jed R. Brubaker, How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis and Image Labeling Services, 3 Proc. ACM on Hum.-Comput. Interaction, no. CSCW, art. 144, at 144:1, 144:2 (2019).

57. Allison J.B. Chaney, Brandon M. Stewart & Barbara E. Engelhardt, How Algorithmic Confounding in Recommendation Systems Increases Homogeneity and Decreases Utility, 12 Proc. ACM Conf. on Recommender Sys. 224, 224 (2018).

interests and prior searches to autocomplete our queries and arrange search results.[58]

Data mining enhances the state's power to leverage gender data to make decisions about people's lives.[59] Sex and gender have become data points in complex algorithms that try to predict recidivism in sentencing: "Female" is associated with lower rates of recidivism; "male" with higher.[60] The now-infamous Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) system, which assesses risk for use in parole decisions, also uses gender data in the same way.[61] Public and private employers use algorithms to assess job applicants.[62] An increasing number of jurisdictions use binary gender data to train complex algorithms meant to identify children who are at risk of committing future

---

58. See How Google Autocomplete Predictions Work, Google, https://support.google.com/websearch/answer/7368877?hl=en [https://perma.cc/BSV8-6BTL] (last visited Aug. 24, 2023).

59. Although this section is exclusively about the state's use of advanced technology to make policy decisions, there is a vast literature on how private companies use these kinds of automated systems to make decisions about credit, loan risks, housing, and much more. See, e.g., Frank Pasquale, Black Box Society: The Secret Algorithms that Control Money and Information 102 (2015) [hereinafter Pasquale, Black Box Society]; Citron & Pasquale, supra note 17, at 4 (describing algorithm use to score credit card applicants and rank job candidates' talent, among other uses); Katyal, Private Accountability, supra note 17, at 56 (describing algorithmic housing and hiring discrimination); Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson & Harlan Yu, Accountable Algorithms, 165 U. Pa. L. Rev. 633, 636 (2017) (describing algorithmic decisionmaking for loan and credit card applications). There is also a related literature about how algorithms exacerbate inequality and should trigger equal protection concerns. See, e.g., Virginia Eubanks, Automating Inequality 180–88 (2018); Barocas & Selbst, Big Data's Disparate Impact, supra note 12, at 673–74; Deborah Hellman, Sex, Causation, and Algorithms: Equal Protection in the Age of Machine Learning, 98 Wash. U. L. Rev. 481, 484 (2020) [hereinafter Hellman, Causation].

60. See State v. Loomis, 881 N.W.2d 749, 765 (Wis. 2016); see also Brian J. Ostrom, Matthew Kleiman, Fred Cheesman II, Randall M. Hansen & Neal B. Kauder, Nat'l Ctr. for State Cts. & Va. Crim. Sent'g Comm'n, Offender Risk Assessment in Virginia 74–76 (2002), http://www.vcsc.virginia.gov/risk_off_rpt.pdf [https://perma.cc/TAD4-7SLF] (providing calculations that demonstrate that their "results suggest that men had a higher probability of recidivating than women").

61. See Julia Dressel & Hany Farid, The Accuracy, Fairness, and Limits of Predicting Recidivism, Sci. Advances, no. eaao5580, Jan. 2018, at 1, 1; see also *Loomis*, 881 N.W.2d at 754–57; Sam Corbett-Davies, Emma Pierson, Avi Feller & Sharad Goel, A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased Against Blacks. It's Actually Not that Clear, Wash. Post (Oct. 17, 2016), https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/ [https://perma.cc/WH6P-3YQE]; Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kirchner, Machine Bias, ProPublica (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [https://perma.cc/V6CV-QJG8].

62. Kim, supra note 12, at 874–90 (emphasizing that employers' use of data analytic tools to identify employees' skills also disadvantages certain groups).

violence.[63] And law enforcement uses binary gender data in facial recognition tools to help identify persons of interest in criminal investigations.[64]

Data-matching and data-mining programs have several things in common that make them appear attractive for government agencies. Both automated systems use large datasets to identify patterns that might be illegible to humans but that are relevant to government agencies: fraud, eligibility, and risk assessment. Importantly, both systems are designed and marketed to reduce costs and increase efficiency.[65] As a result, automation taps into persistent norms that efficient government is "good" government that can do more with less.[66]

B. *Effects on Gender-Diverse Populations*

Data-matching systems pose unique problems for transgender and nonbinary people. Many have inconsistent identity documents because gender reclassification rules are labyrinthine and inconsistent.[67] Individuals may lack the money or time to meet onerous medical or surgical standards for updating birth certificates or driver licenses in certain jurisdictions.[68] Granted, transgender people could purposely answer questions to match their information on official documents. But

---

63. See, e.g., Nicole M. Muir, Jodi L. Viljoen, Melissa R. Jonnson, Dana M. Cochrane & Billie Joe Rogers, Predictive Validity of the Structured Assessment of Violence Risk in Youth (SAVRY) With Indigenous and Caucasian Female and Male Adolescents on Probation, 32 Psych. Assessment 594, 597 (2020).

64. See, e.g., Lynch v. State, 260 So. 3d 1166, 1169 (Fla. Dist. Ct. App. 2018) ("[T]he crime analyst testified [that] . . . [she] [t]urn[ed] to law-enforcement databases, . . . looked up those who had been previously arrested at the address . . . [and] then used a facial-recognition program that compared the photo officers took against photos in law-enforcement databases.").

65. See, e.g., Charette, supra note 51.

66. See Brooke D. Coleman, The Efficiency Norm, 56 B.C. L. Rev. 1777, 1786–95 (2015) (describing and critiquing the tendency to associate efficiency and cost cutting with good government).

67. See Paisley Currah, Sex Is as Sex Does: Governing Transgender Identity 76–98 (2022) ("Individuals whose gender identity differs from what is traditionally associated with the sex assigned to them at birth may be included or excluded from systems of sex classification."); Katri, supra note 35, at 656–95 (examining American sex reclassification law); Spade, Documenting Gender, supra note 36, at 733–34 (same).

68. Many lack the financial means to access appropriate healthcare. But the socioeconomic marginalization of transgender people and, in particular, trans people of color exacerbates the problem. Sandy E. James, Jody L. Herman, Susan Rankin, Mara Keisling, Lisa Mottet & Ma'ayan Anafi, The Report of the 2015 U.S. Transgender Survey 5 (2016), https://transequality.org/sites/default/files/docs/usts/USTS-Full-Report-Dec17.pdf [https://perma.cc/KVV9-AQ8E]; see also Transgender L. Ctr., Transgender Health and the Law: Identifying and Fighting Health Care Discrimination (2004), http://transgenderlawcenter.org/wp-content/uploads/2012/07/99737410-Health-Law-Fact.pdf [https://perma.cc/V4K8-BTJM].

lying on government forms is a crime.[69] Identifying yourself as something you're not resurrects gender dysphoria.[70] Plus, intentional self-misidentification on one form fails to solve the problem created by data-matching and data-mining algorithms: The vast reach of data-matching databases and data inputs creates the risk that any inconsistency on any form completed at any time could trigger an accusation of fraud.[71]

Transgender, nonbinary, and gender-nonconforming individuals also face increased risk from automated systems designed to turn the body into code in the most efficient way possible.[72] Machines designed for efficiency make conclusions that cover most people most of the time. They "stylize reality";[73] models make assumptions about the world to make data more legible and easier to manipulate.[74] As a result, they have trouble correctly identifying people who do not meet social expectations associated with their assigned gender at birth.[75] If training data is binary or based on cisnormative expectations of how males and females are supposed to look,[76] as was the case with the full-body scanner that flagged Sasha as a security risk, those who exist outside the gender binary are treated as outliers.[77] Similar harms can affect people of color, especially when AI is trained on mostly white faces and expected to make predictions about how Black or Asian individuals should look. That is how facial recognition

---

69. See, e.g., IRS, U.S. Department of Treasury, U.S. Individual Income Tax Return Form 1040 (2022), https://www.irs.gov/pub/irs-pdf/f1040.pdf [https://perma.cc/WD72-EM6Z] ("Under penalties of perjury, I declare that . . . to the best of my knowledge and belief, [the information I provided is] true, correct, and complete.").

70. "Gender dysphoria" refers to clinical distress associated with one's sex assigned at birth. Am. Psychiatric Ass'n, Diagnostic and Statistical Manual of Mental Disorders 455–56 (5th ed. 2013).

71. Currah & Mulqueen, supra note 2, at 559 (stating that providing inconsistent information during the air travel process may create false security risk alerts).

72. Kathryn Conrad, Surveillance, Gender, and the Virtual Body in the Information Age, 6 Surveillance & Soc'y 380, 382–85 (2009) (referring to tools like iris scanners, digital fingerprinting, and facial recognition as the "informatization of the body" by the state (internal quotation marks omitted) (quoting Irma van der Ploeg, Genetics, Biometrics and the Informatization of the Body, 43 Ann Ist Super Sanità 44, 44 (2007) (It.))).

73. Mar Hicks, Hacking the Cis-tem: Transgender Citizens and the Early Digital State, 41 IEEE Annals Hist. Computing, no. 1, 2019, at 20, 29.

74. George E.P. Box & Norman R. Draper, Empirical Model-Building and Response Surfaces 74 (1987).

75. See Scheuerman et al., supra note 56, at 144:14–144:15.

76. See id. at 144:17.

77. See Kendra Albert & Maggie Delano, Algorithmic Exclusion, *in* Handbook of Critical Studies of Artificial Intelligence 538, 540 (Simon Lindgren ed., 2023) ("[M]ethods [used] to remove outliers from particular datasets may result in indirect exclusion of particular groups of people . . . .").

technology classifies the eyes of Asian faces as "closed" or misidentifies Black women at higher rates than white women.[78]

Plus, data-mining systems need training data, all of which come from a time (even in the very recent past) when transgender, nonbinary, and gender-nonconforming people were barely recognized in the public consciousness.[79] This "increase[s] the influence of the past"—one dominated by the gender binary (as well as white supremacy and homophobia, among other exclusionary ideologies)—on the future.[80] The process is also iterative and self-reinforcing: Data inputs reflect the gender binary; algorithmic technologies output new data that reflect the gender binary; those data are then added back to better train the automated system, thereby amplifying and replicating the gender assumptions built into the algorithm itself.[81]

The exclusion of gender diversity also stems from the social contexts in which algorithmic technologies are designed. The people who design automated decisionmaking systems and the corporate organizations in which they do their work are notoriously unrepresentative; they skew cisgender, heterosexual, and white.[82] The lived experiences of that limited slice of the population are more likely than others to make their way into the political, distributional, and technical decisions in design.[83]

C. *Harms of Erasure*

Automated decisionmaking systems harm marginalized populations in at least four related ways. The first two are practical. First, algorithmic tools create repeated moments of vulnerability for transgender and nonbinary individuals with inconsistent identity documents. Every airport or doctor's visit, every job or benefits application, every background check, every vote, every interaction with the police, every plan to start a business, and every identity verification demand triggers a larger system of technological surveillance designed, from the ground up, to erase or

---

78. See Buolamwini & Gebru, supra note 12, at 10–11; Selina Cheng, An Algorithm Rejected an Asian Man's Passport Photo for Having "Closed Eyes", Quartz (Dec. 7, 2016), https://qz.com/857122/an-algorithm-rejected-an-asian-mans-passport-photo-for-having-closed-eyes/ [https://perma.cc/YBZ8-G3FX].

79. Indeed, as the sociotechnical scholar Os Keyes found in a review of hundreds of published studies at the intersection of AI and gender, every single one reified the gender binary. Os Keyes, The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition, 2 Proc. ACM on Hum.-Comput. Interaction, no. CSCW, art. 88, at 88:1, 88:2 (2018).

80. Hellman, Causation, supra note 59, at 487.

81. Katyal & Jung, supra note 10, at 710.

82. Kate Crawford, Opinion, Artificial Intelligence's White Guy Problem, N.Y. Times (June 25, 2016), https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html (on file with the *Columbia Law Review*).

83. Id.

misgender anyone outside the norm.[84] Second, and relatedly, the pervasive danger of vulnerability causes chilling effects. To avoid situations likely to include misgendering, many transgender individuals choose to avoid those situations entirely, opting themselves out of daily life, government benefits, and opportunity.[85] Interviews with transgender individuals describe a "continuous assault upon our existence, well-being, opportunity, and potential" and a "process of cisgendering reality" whereby "only cisgender people may move freely without punishment, shock, and stigmatization coming from others," among other similar expressions of harm.[86] This may be one reason why transgender, nonbinary, and gender-nonconforming individuals report higher rates of depression, suicidal ideation, loneliness, and underemployment than the general population.[87]

Third, exclusion comes with dignitary harms as well. Institutional erasure tells gender-nonconforming individuals that they do not count, that their identities do not matter, and that their humanity does not exist. This exclusion is then broadcast throughout the data ecosystem, affecting the views of everyone who encounters binary gender data.[88]

Fourth, and finally, algorithms and automated systems more generally amplify these harms, creating powerful expressive effects. Because they rely on data inputs to make predictive policy decisions about the future, algorithms replicate and entrench old biases.[89] Popular trust in computers as infallible make those predictions harder to challenge.[90] Beyond merely amplifying old harms, automation privileges decisionmaking based exclusively on quantifiable variables, ignoring value-based, qualitative, and human rights considerations that defy neat clustering into numerical

---

84. Chan Tov McNamarah, Misgendering, 109 Calif. L. Rev. 2227, 2234–35 (2021) (arguing that misgendering and misrecognition are part of a pattern of subordination that denigrates the personhood of transgender and nonbinary people).

85. Currah & Mulqueen, supra note 2, at 560.

86. J.E. Sumerau & Lain A.B. Mathers, America Through Transgender Eyes 3–4 (2019).

87. James et al., supra note 68, at 5–6.

88. See Taylor Flynn, Instant (Gender) Messaging: Expression-Based Challenges to State Enforcement of Gender Norms, 18 Temp. Pol. & C.R.L. Rev. 465, 466 (2009). For more on expressive effects of law on gender, see infra section III.D. For a more general account of expressive effects of the law, see Danielle Keats Citron, Law's Expressive Value in Combating Cyber Gender Harassment, 108 Mich. L. Rev. 373, 404–14 (2009) [hereinafter Citron, Expressive Value]; Cass R. Sunstein, On the Expressive Function of Law, 144 U. Pa. L. Rev. 2021, 2022–24 (1996); Matthew Tokson & Ari Ezra Waldman, Social Norms in Fourth Amendment Law, 120 Mich. L. Rev. 265, 279–84 (2021).

89. E.g., Cathy O'Neil, Weapons of Math Destruction 3, 7–8 (2016); Pasquale, Black Box Society, supra note 59, at 14–15; Katyal, Private Accountability, supra note 17, at 69.

90. Scholars call this "automation bias." See Ryan Calo, Modeling Through, 71 Duke L.J. 1391, 1417 (2022) [hereinafter Calo, Modeling]; Citron, Technological Due Process, supra note 17, at 1271–72.

values. In other words, whereas inconsistencies in documents could have once been resolved through civil servant discretion, machines programmed to see only ones and zeros transform data input errors or inconsistencies into grounds for benefit denials, fraud accusations, and discrimination.

To most scholars, technology is the root cause of these harms; law seems absent from this story of automation and discrimination. Legal scholars who see law as a means of holding states and technology companies accountable for harms caused by automated decisionmaking systems tend to gloss over the things that created the conditions necessary for automation in the first place. Indeed, because it focuses on legal redress after algorithmic harm, much of the algorithmic accountability literature skips right to descriptions of legal responses to harm.[91] Some scholars merely note that algorithmic policymaking is becoming "more common."[92] Others acknowledge that the rise of automation stems from austerity.[93] Although tight budgets are undoubtedly the products of law, this legal narrative of the rise of the automated administrative state is thin.

Automated systems that apply rules, match identities, and mine for patterns need data to function; states need to find or purchase those data from somewhere. System designers also need instructions about what categories of data to include in the system. They need principles, values, directions, goals, and budgets with which to build automated tools for the state to use. In particular, the state must decide whether, when, and how to collect gender data; whether, when, and how to share it; and whether, when, and how to use it. At each stage—collection, sharing, and automated use—binary gender data's pathway is laid, brick by brick, by law and, more specifically, by a legal regime designed primarily for efficiency. The next three Parts describe this pathway and how it erases gender-diverse populations and causes the above harms.

II. LAW AND THE COLLECTION OF BINARY GENDER DATA

Gender data's path begins with laws that require states to collect gender data. It is difficult to estimate how many state laws require individuals to provide their sex or gender to engage in daily life; even targeted searches return thousands of hits. The examples discussed below

---

91. See supra note 17; see also Frank Pasquale, The Second Wave of Algorithmic Accountability, LPE Project (Nov. 25, 2019), https://lpeproject.org/blog/the-second-wave-of-algorithmic-accountability/ [https://perma.cc/P68K-K87D] (referring to this scholarship as the "first wave," following similar terminology used in the feminist movement).

92. Hellman, Causation, supra note 59, at 484.

93. E.g., Calo & Citron, supra note 23, at 800; Citron, Technological Due Process, supra note 17, at 1259; see also Robert Brauneis & Ellen P. Goodman, Algorithmic Transparency for the Smart City, 20 Yale J.L. & Tech. 103, 114 (2018) (discussing how tight budgets impel municipalities to use private technology companies for their automation needs).

are paradigmatic of the law's role in triggering many gender data streams. After describing some of these laws, this Part then shows that even though the law rarely states *how* the information should be collected, the law's underlying assumptions and practical implementation act as a filter that makes binary gender data streams most likely.

A. Statutory Gender Data-Collection Mandates

Almost all states use individuals' sex and gender data in several administrative areas.[94] Thirty-seven states require driver license or identification card applicants to provide their sex.[95] Eight states ask for gender.[96] Ten states have statutes requiring sex data on voter registration

94. This Part recites some of the ways sex and gender data are used. It does not support their use. Indeed, using sex or gender to classify populations has been deftly criticized in the sociolegal literature. See, e.g., Heath Fogg Davis, Beyond Trans: Does Gender Matter? 17 (2017); Wipfler, supra note 35, at 493.

95. See Alaska Stat. § 28.15.061(b)(1) (2023); Ariz. Rev. Stat. Ann. §§ 28-3158(C), -3165(F) (2023); Ark. Code Ann. § 27-16-701(b)(1) (2023); Colo. Rev. Stat. § 42-2-107(2)(a)(I) (2023); Del. Code tit. 21, § 2711(b) (2023); Ga. Code Ann. § 40-5-25(c) (2023); Idaho Code § 49-306(3) (2023); 625 Ill. Comp. Stat. Ann. 5/6-106(b) (West 2023); Iowa Code § 321.182.1.a (2023); Ky. Rev. Stat. Ann. § 281A.140(1)(b) (West 2023); La. Stat. Ann. § 32:410.A(3)(a)(viii) (2023); Md. Code Ann., Transp. § 16-106(b)(1) (West 2023); Mass. Gen. Laws Ann. ch. 90F, § 8(3) (West 2023); Mich. Comp. Laws Ann. § 257.307(1)(a) (West 2023); Minn. Stat. § 171.06.3(1) (2023); Miss. Code Ann. § 63-1-19(1)(a) (2023); Mo. Ann. Stat. § 302.171(1) (West 2023); Mont. Code Ann. § 61-5-107(2) (West 2023); Nev. Rev. Stat. Ann. § 483-290.1(d) (West 2023); N.M. Stat. Ann. § 66-5-9(C) (2023); N.Y. Veh. & Traf. Law § 502.1 (McKinney 2023); N.C. Gen. Stat. § 20-7(b1)(3) (2023); N.D. Cent. Code § 39-06-07.2 (2023); Ohio Rev. Code Ann. § 4506.07(A)(1) (2023); Okla. Stat. tit. 47, § 6-106.B.3 (2023); Or. Rev. Stat. § 807.050(1) (West 2023); 31 R.I. Gen. Laws § 31-10.3-18(b) (2023); S.C. Code Ann. § 56-1-80(A)(3) (2023); S.D. Codified Laws § 32-12-3 (2023); Tenn. Code Ann. § 55-50-321(c)(1)(A) (2023); Tex. Transp. Code Ann. § 521.142(c)(1) (West 2023); Utah Code § 53-3-205(8)(a)(i)(C) (2023); Va. Code Ann. § 46.2-323(B) (2023); Wash. Rev. Code Ann. § 46.20.091(1)(c) (West 2023); W. Va. Code Ann. § 17B-2-6(c) (LexisNexis 2023); Wis. Stat. & Ann. § 343.14(2)(b) (2023); Wyo. Stat. Ann. § 31-7-111(b)(ii) (2023).

96. See Cal. Veh. Code § 12800(a)(1) (2023); Conn. Gen. Stat. Ann. § 14-36h(a) (West 2023); Fla. Stat. Ann. § 322.14(1)(a) (West 2023); Haw. Rev. Stat. Ann. § 286-111(d) (West 2023); Ind. Code Ann. § 9-24-9-2(a)(3) (West 2023); Kan. Stat. Ann. § 8-240(c) (West 2023); Neb. Rev. Stat. § 60-484(3) (2021); N.H. Rev. Stat. Ann. § 263:5(II)(b) (2023). The remaining state laws are silent. For a brief discussion of the differences yet entanglements between sex and gender, please see infra notes 222–225 and accompanying text.

applications;[97] three collect gender data.[98] All states require applicants to present a form of identification in order to register to vote, and all driver licenses and state identification cards must include sex designations under federal law.[99] Statutes governing birth and death certificates all mandate the inclusion of sex data.[100] And five states still require parties to disclose their sex on marriage license applications.[101]

Sex and gender data are also statutorily required in more targeted areas of social and professional life. Firearm licenses require sex or gender.[102] Prospective state employees, licensed professionals, and foster parents, among others, have to provide their sex for background checks.[103] Licensure for for-hire and private carrier vehicle drivers,[104] chiropractors,[105] private detectives,[106] medical cannabis caregivers,[107]

---

97. See Ala. Code § 17-4-36(a) (2023); Alaska Stat. § 15.07.060(a)(1) (2023); Fla. Stat. Ann. § 97.052(2)(i) (West 2023); Ga. Code Ann. § 21-2-417(1)(c)(5) (2023); Idaho Code § 34-411(1)(a) (2023); 10 Ill. Comp. Stat. Ann. 5/5-7 (West 2023); Iowa Code § 48A.11(1)(g) (2023); Kan. Stat. Ann. § 25-2309(b)(4) (West 2023); La. Stat. Ann. § 18:104(B)(1) (2023); S.C. Code Ann. § 7-5-170(2) (2023); see also James et al., supra note 68, at 233–35 ("[Transgender] respondents reported not being registered to vote because they wanted to avoid anti-transgender harassment by election officials . . . and because they thought their state's voter identification law would stop them from voting . . . .").

98. See N.M. Stat. Ann. § 1-5-19(B) (2023); Va. Code Ann. § 24.2-418(A) (2023); W. Va. Code Ann. § 3-2-5(d)(8) (LexisNexis 2023).

99. See Adair, supra note 36, at 587–88 (explaining how sex markers are universally mandated by the federal 2005 Real ID Act).

100. See Lisa Mottet, Modernizing State Vital Statistics Statutes and Policies to Ensure Accurate Gender Markers on Birth Certificates: A Good Government Approach to Recognizing the Lives of Transgender People, 19 Mich. J. Gender & L. 373, 381–83 (2013).

101. See Colo. Rev. Stat. § 14-2-105(1)(a) (2023); Del. Code tit. 13, § 122(a) (2023); 750 Ill. Comp. Stat. Ann. 5/202(a)(1) (West 2023); Minn. Stat. § 517.08(1a)(1) (2022); Or. Rev. Stat. § 106.041(2)(b) (West 2023).

102. Examples of laws requiring sex data in order to carry a firearm include Ark. Code Ann. § 5-73-310(1) (2023); Haw. Rev. Stat. Ann. § 134-2 (West 2023); 430 Ill. Comp. Stat. Ann. 65/6(a) (West 2023); Ind. Code Ann. 35-47-2-3(e) (West 2023); Iowa Code §§ 724.10, .17 (2023); Mass. Gen. Laws Ann. ch. 140, § 123 (West 2023); N.J. Stat. Ann. § 2c:58-3(e) (West 2023); S.C. Code Ann. § 23-31-215(E)(3) (2023); Wis. Stat. & Ann. § 175.60(5) (2021–2022). Those requiring gender data include, for example, Cal. Penal Code §§ 30900(b)(3), (c)(3), 33850(a)(1) (2020); Ky. Rev. Stat. Ann. § 237.110(20)(b)(2) (West 2023); Mo. Ann. Stat. § 571.205(4)(1) (West 2023); N.M. Stat. Ann. § 29-19-5(A)(1) (2005); N.Y. Penal Law § 400.00(5) (McKinney 2023); Va. Code Ann. § 18.2-308.04(E) (2023); Wash. Rev. Code Ann. § 9.41.070(4) (West 2022).

103. E.g., Ala. Code § 34-25B-13(a)(1) (2023); Del. Code tit. 24, §§ 1205, 1313, 5507 (2023); Haw. Rev. Stat. Ann. §§ 421I-12, 514B-133(a) (West 2023); 70 Ill. Comp. Stat. Ann. 1205/8-23(a), 3605/28b (West 2023); 18 Pa. Stat. and Cons. Stat. Ann. § 6111(b)(1.1)(iii) (West 2023); Tenn. Code Ann. § 62-26-208(a)(1)(B) (2023).

104. 70 Ill. Comp. Stat. Ann. 3605/28b.

105. Mo. Ann. Stat. § 331.030(2) (West 2023); Tenn. Code Ann. § 63-4-109(b) (2023).

106. Minn. Stat. § 326.3382(a)(1) (2022).

107. Utah Code § 26B-4-214(5)(b) (2023).

commercial fishers,[108] home solicitation salespersons,[109] anyone "engaged in the business of collecting secondhand building materials for resale,"[110] and precious metals dealers all require sex data in some states.[111] Organ donors must be issued identification cards that list their sex.[112] Anyone in Illinois and Missouri whose job requires them to work with explosives has to provide their sex to obtain a license.[113] Collection agents in Arkansas and bail enforcement agents in Delaware can be licensed only if they provide their sex.[114] If minors want to work in the District of Columbia or Puerto Rico, their permit or certificate must have, among other things, their sex.[115] This section could go on and on.[116]

B. *Mandating the Gender Binary at Data Collection*

Although these laws mandate sex and gender data collection, it is rare for a law to explicitly detail how to collect the data, what answer options to provide, how to phrase the question, or whether forms should explain why the information is required. Therefore, it is at least theoretically possible that these laws could catalyze gender data streams that respect diverse

---

108. Cal. Fish & Game Code § 7851 (2023).

109. Fla. Stat. Ann. § 501.022(2)(c) (West 2023).

110. Va. Code Ann. § 59.1-118 (2023).

111. Va. Code Ann. § 54.1-4108(B) (2023).

112. Md. Code Ann., Transp. § 12-301(g)(2)(iii) (West 2023).

113. 225 Ill. Comp. Stat. Ann. 210/2002 (West 2023); Mo. Ann. Stat. § 319.306(1)(4) (West 2023).

114. Ark. Code Ann. § 17-24-302(a)(2) (2023); Del. Code tit. 24, § 5507(c) (2023).

115. D.C. Code § 32-208 (2023); P.R. Laws Ann. tit. 29, § 436 (2012).

116. Other areas where state law requires the collection of sex and gender data include public-facing reports and applications for scholarships, loan forgiveness, and appointed government positions. E.g., Cal. Gov't Code § 12011.5(n)(1)(A)–(B) (2023) (requiring release of gender data of all applicants to state judicial positions); Cal. Welf. & Inst. Code § 11024(a) (2023) (breaking down Medi-Cal enrollees by gender); Conn. Gen. Stat. Ann. § 10-95k(a) (West 2023) (requiring the board of technical colleges to deliver biennial reports to the state's General Assembly Committee on Education, including "the number accepted and the number enrolled reported by race and sex"); Fla. Stat. Ann. Sup. Ct. Jud. Nominating Comm'n Rules Proc. § II (West 2023) (applicants for state judicial appointments); 20 Ill. Comp. Stat. Ann. 2610/11.5(b) (West 2023) (requiring the Illinois State Police Merit Board to have a gender breakdown for individuals promoted in their reports); 110 Ill. Comp. Stat. Ann. § 932/20(f) (West 2023) (loan repayment and scholarships for healthcare workers); N.J. Stat. Ann. § 44:15-2(b) (West 2023) (requiring report of low-income elderly residents to break down population by gender); id. § 52:17B-4.11(a)(1)–(5) (requiring breakdown of police forces by gender); 40 R.I. Gen. Laws § 8.7-9(c)(2) (2023) (requiring Rhode Island's health department to report on the sex breakdown of individuals with disabilities on Medicaid); Tex. Gov't Code Ann. § 411.193 (West 2023) (making reports of gun licenses issued the previous month, broken down by gender, available to the public); Va. Code Ann. § 30-394(A) (2023) (requiring gender data to apply to be a citizen commissioner on the Virginia Redistricting Commission).

gender identities. To be sure, some laws do.[117] But three features of statutory gender data mandates tend to binarize whatever data are collected: the source of the data, the assumption that gender data are a useful securitizing tool, and the law's practical implementation. The first two concerns are discussed here; the third is detailed in the next section.

The first feature of state gender data-collection mandates that tilts the data toward the gender binary is that much of the data is created by the state in the first place. It is commonly presumed that sex and gender data are raw materials in what Professor Julie Cohen calls the "biopolitical public domain," or a "source of raw materials about people framed as inputs into productive, informationalized activity."[118] These data are biopolitical because they are information about people used for classification and, therefore, have political and distributive consequences; they are also presumed to be in the public domain—namely, there for the taking within a legal construct of privilege, or "conduct as to which no one has a right to object."[119] The biopolitical public domain is a foundational premise of the information economy and the automated state. It asserts that certain data are raw, that no previous claims to those data exist, and that they can be collected, used, and mixed with labor and turned into something productive.[120]

But gender designations are not raw. They are mediated by the state before and after birth: at Medicaid recipients' prenatal appointments with healthcare providers, during which physicians designate the fetus's sex; at birth, when physicians or bureaucrats complete birth certificates and Live Birth Worksheets; and at schools, where nurses designate sex or gender on immunization and health forms. By the time Sasha walked through the full-body scanner and Toby submitted his workers' compensation claim, they had both been designated by the state as male or female.[121] The presumed power of official documents to verify identity derives precisely from "the authority of the institution that issued it," not from the

117. For example, in 2018, New York City's health department added the nonbinary gender category "X" to birth certificates, so the department built a new form to reflect the new option. See Certificate Corrections, N.Y.C. Health, https://www.nyc.gov/site/doh/services/certificate-corrections.page [https://perma.cc/HHP7-UHWA] (last visited Aug. 24, 2023). California law states that residents "shall choose their gender category of female, male, or nonbinary" on a driver license application. Cal. Veh. Code § 12800(a)(2) (2023). Therefore, that gender data stream will, by statute, include data on nonbinary individuals. See Cal. Dep't Motor Vehicles, Form DL 329S, Gender Category Request (Jan. 2019), https://www.dmv.ca.gov/portal/uploads/2020/03/dl329S.pdf [https://perma.cc/3PNP-S7C7].

118. Cohen, Between Truth and Power, supra note 16, at 48.

119. Id. at 49.

120. Id. at 50–52.

121. Spade, Normal Life, supra note 19, at 14.

documents' inherent accuracy or the law's respect for self-identification.[122] In other words, state laws that require gender data collection are relying on the state's determinations of a person's gender, which historically have been binary.[123]

In addition to assuming that sex and gender data are raw and accurate, a regime that uses sex and gender data to verify identity, assess risk, and maintain security also assumes that sex and gender are effective at achieving these goals. But the only way these data could be effective is if they were unchanging descriptions of individuals. If they weren't, gender data would do a poor job at ensuring that the people applying for jobs or benefits or licenses are who they say they are. Security systems use retinal scans instead of, say, hair color for the same reason: The former relies on data that rarely, if ever, change; the latter can change on a whim. One is a more permanent marker of identity than the other.[124] Of course, sex and gender designations can change. Therefore, the only people for whom gender data can help predict whether a given person is committing fraud are cisgender people. In this way, the state's mere use of sex and gender data as securitizing, identification-verifying tools necessarily implies cisnormativity.

C. *Entrenching the Gender Binary Through Form Design*

This leads to the third feature of statutes' capacity to binarize gender data—namely, their implementation in practice through official government forms. We fill out forms to obtain identification cards, purchase license plates, practice licensed professions, record vaccinations for schoolchildren, and obtain government-sponsored healthcare, among myriad other aspects of everyday life. Forms were supposed to give Toby access to compensation after being injured on the job. Forms' ubiquity means that they have an outsized effect on how we perceive and understand the law.[125]

Forms are also where the state collects data to classify people by race, gender, ethnicity, disability, and myriad other demographic characteristics. The design of those forms determines what the state's gender data will

---

122. Irma van der Ploeg, Written on the Body: Biometrics and Identity, 29 Comput. & Soc'y, no. 1, 1999, at 37, 38.

123. Jane Caplan, "This or That Particular Person": Protocols of Identification in Nineteenth-Century Europe, *in* Documenting Individual Identity 49, 52 (Jane Caplan & John Torpey eds., 2001).

124. Not that we should rush to use retinal scans and other biometric data. See, e.g., Danielle Keats Citron, Reservoirs of Danger: The Evolution of Public and Private Law at the Dawn of the Information Age, 80 S. Cal. L. Rev. 241, 250–53, 255 (2007) (noting that "[t]he release of biometric information from a database will engender serious harm as criminals can use such data to impersonate individuals").

125. See Patricia Ewick & Susan S. Silbey, The Common Place of Law: Stories From Everyday Life 30–34 (1998) (introducing the concept of "legal consciousness").

look like. That is a type of power exercised by what political scientist Michael Lipsky called "street-level bureaucrats."[126] Street-level bureaucrats are frontline civil servants with the least formal authority but the most discretion to determine how the law is implemented.[127] For example, in Professor Lipsky's canonical account, street-level bureaucrats decide how to achieve the best interests of children in foster care, flexibly apply rules to send lifesaving benefits to those in need, and evaluate patient medical needs to secure care.[128] Frontline workers also determine precisely how to begin the large, free-flowing system of gender data among government agencies at the local, state, and federal levels.[129] The law of sex and gender "remains an abstraction" until these frontline workers carry it out and apply it in real life,[130] communicating with the public through the gender questions and answer options they create.[131] When they exercise this discretion to collect sex and gender information in certain ways, gender-box designers are effectively "making law" in the most practical sense.

Gender questions on most government forms are limited to male/female answer options.[132] This is because form designers work in

---

126. Michael Lipsky, Street-Level Bureaucracy 3 (2d ed. 2010). Granted, traditional street-level bureaucrats have often been defined by their face-to-face interactions with the public. Id. at 3–4. But their choices affect the practical implementation of the law. Mark Bovens & Stavros Zouridis, From Street-Level to System-Level Bureaucracies: How Information and Communication Technology Is Transforming Administrative Discretion and Constitutional Control, 62 Pub. Admin. Rev. 174, 181 (2002). Form designers have at least three characteristics in common with street-level bureaucrats: They exercise discretion, they shape policy through their discretionary acts, and they sit in social and organizational contexts that may affect their work. They exercise discretion because even when formal law requires an agency to collect sex or gender data, the law rarely says anything about *how* the agency should collect it. See Evelyn Z. Brodkin, Reflections on Street-Level Bureaucracy: Past, Present, and Future, 72 Pub. Admin. Rev. 940, 943 (2012) (reviewing Lipsky, supra) (noting that policy is "indeterminate").

127. Steven Maynard-Moody & Michael Musheno, State Agent or Citizen Agent: Two Narratives of Discretion, 10 J. Pub. Admin. Rsch. & Theory 329, 333 (2000). Such discretion is inevitable because it is inherent to both street-level work specifically and "all acts of administration" generally. Id. at 338–39.

128. See Lipsky, supra note 126, at 3 (providing examples of roles street-level bureaucrats inhabit in public service agencies).

129. See Fahey, Data Federalism, supra note 16, at 1078–79 (documenting the ways mid- to line-level bureaucrats are part of a larger system of data exchange between agencies).

130. See Bernardo Zacka, When the State Meets the Street: Public Service and Moral Agency 16 (2017).

131. Koen P.R. Bartels, Public Encounters: The History and Future of Face-to-Face Contact Between Public Professionals and Citizens, 91 Pub. Admin. 469, 476–77 (2013) (describing the performative nature of interactions between public officials and citizens).

132. Ari Ezra Waldman, The Gender Box: Heterogeneity and Inclusivity in State Collection of Sex and Gender Data 20 (n.d.) (unpublished manuscript) (on file with the *Columbia Law Review*) [hereinafter Waldman, The Gender Box] (empirically measuring the extent to which sex and gender questions on government forms permit answers beyond the gender binary).

organizational contexts in which a combination of social forces incentivizes inertia.[133] These include complex decisionmaking processes that make change difficult, social networks of colleagues that help civil servants "learn the ropes" and maintain the status quo, the perception that expertise is irrelevant to gender question design, and intergovernmental dependencies that constrain design options.[134] These pressures, combined with norms against politicization of the bureaucracy,[135] status quo biases and path dependencies,[136] the urge to simplify information for superiors, and decades-long trends toward digitization and automation,[137] all encourage form designers to restrict sex and gender questions to male/female answer options.[138]

---

133. Ari Ezra Waldman, Opening the Gender Box: Legibility Dilemmas and Gender Data Collection on U.S. State Government Forms, 49 Law & Soc. Inquiry (forthcoming 2023) (manuscript at 14) (on file with the *Columbia Law Review*) [hereinafter Waldman, Opening].

134. See, e.g., Deneen M. Hatmaker, Hyun Hee Park & R. Karl Rethemeyer, Learning the Ropes: Communities of Practice and Social Networks in the Public Sector, 14 Int'l Pub. Mgmt. J. 395, 396 (2011) (explaining how an organization's "socialization tactics" function to inculcate organizational values in newcomers); Rebecca Ingber, The Obama War Powers Legacy and the Internal Forces that Entrench Executive Power, 110 Am. J. Int'l L. 680, 696–98 (2016) (describing how deliberative bodies that operate on a "consensus model" can stifle dissent); Gillian E. Metzger, Administrative Constitutionalism, 91 Tex. L. Rev. 1897, 1900, 1929–30 (2013) (describing how federal agencies' complex web of interactions with the public and other governmental bodies helps construct constitutional meaning); Nadine Raaphorst & Kim Loyens, From Poker Games to Kitchen Tables: How Social Dynamics Affect Frontline Decision Making, 52 Admin. & Soc'y 31, 32–34 (2020) (arguing that the complexity of multiprofessional social interactions directly affects frontline decisionmaking); Gerald E. Caiden, Excessive Bureaucratization: The J-Curve Theory of Bureaucracy and Max Weber Through the Looking Glass, Dialogue, Summer 1985, at 21, 31–32 (explaining how careerism in bureaucratic organizations can result in "self-perpetuating" systems where "mediocrity predominates").

135. See Ingber, supra note 134, at 687; Jon D. Michaels, An Enduring, Evolving Separation of Powers, 115 Colum. L. Rev. 515, 541–44 (2015).

136. See Graham Allison & Philip Zelikow, Essence of Decision: Explaining the Cuban Missile Crisis 148–49 (Longman 2d ed. 1999) (1971) (defining path dependency in the context of government decisions); William Samuelson & Richard Zeckhauser, Status Quo Bias in Decision Making, 1 J. Risk & Uncertainty 7, 8 (1988) (finding that "decision makers exhibit a significant status quo bias"); Philip J. Weiser, Entrepreneurial Administration, 97 B.U. L. Rev. 2011, 2028–29 (2017) (describing path dependency as a barrier to entrepreneurial approaches to agency work). There are related path dependencies in the formal law, as well. See Oona A. Hathaway, Path Dependence in the Law: The Course and Pattern of Legal Change in a Common Law System, 86 Iowa L. Rev. 101, 104–05 (2001) (applying path dependence theory to the common law doctrine of stare decisis).

137. See Paul Schwartz, Data Processing and Government Administration: The Failure of the American Legal Response to the Computer, 43 Hastings L.J. 1321, 1322–25 (1992) (proposing principles of data protection law to counter the rise of the digitization of personal data by the government).

138. This does not exclude the reality that transphobia pervades social and legal institutions. See Gayle S. Rubin, Thinking Sex: Notes for a Radical Theory of the Politics of Sexuality, *in* Culture, Society and Sexuality, A Reader 150, 158 (Richard Parker & Peter

As a result, even if state laws simply require an agency to collect sex and gender data generally, the forms the agency uses to collect that data will most often reflect the gender binary. Consider, for example, how state boards of elections and secretaries of state implement voter registration laws. Of the seventeen states that explicitly require or request that citizens designate their sex or gender when registering to vote, fourteen use forms with only male/female options.[139] And of the remaining thirty-four

---

Aggleton eds., 2d ed. 2007) (identifying "transsexuals" as one of the "most despised sexual castes"); see also Riki Anne Wilchins, Read My Lips: Sexual Subversion and the End of Gender 230 (1997) (defining transphobia as the "fear and hatred of changing sexual characteristics").

139. Compare Ala. Code § 17-4-36(a) (2022) (requiring the reporting of registered voters' sex to the Secretary of State), with Ala. Sec'y of State, State of Alabama Voter Registration Form (July 5, 2022), https://www.sos.alabama.gov/sites/default/files/voter-pdfs/nvra-2.pdf [https://perma.cc/Q74K-NKFP] (providing only male/female options under sex). Compare Alaska Stat. § 15.07.060(a)(1) (2023) (requiring reporting of the applicant's sex during voter registration), with Alaska Div. of Elections, State of Alaska Voter Registration Application, https://www.elections.alaska.gov/doc/forms/C03-Fill-In.pdf [https://perma.cc/5PZ6-A6M3] (last updated May 12, 2021) (listing only "Male" and "Female" as options under "Gender"). Compare Fla. Stat. Ann. § 97.052(2)(i) (West 2023) (mandating the voter registration application to include a question on the applicant's sex), with Fla. Dep't of State, Florida Voter Registration Application (Oct. 2013), https://files.floridados.gov/media/704795/dsde39-english-pre-7066-20200914.pdf [https://perma.cc/XQS8-BEJD] ("Gender: M, F"). Compare Ga. Code Ann. § 21-2-417(1)(c)(5) (2023) (requiring the voter identification card to list voters' sex), with Ga. Sec'y of State, State of Georgia Application for Voter Registration, https://sos.ga.gov/sites/default/files/forms/GA_VR_APP_2019.pdf [https://perma.cc/3SVQ-B89C] (last visited Aug. 24, 2023) ("Gender: Male, Female"). Compare Idaho Code § 34-411(1)(a) (2023) (requiring individuals wanting to register to vote to provide proof of identity, including their sex), with Idaho Sec'y of State, Idaho Voter Registration Form (2022), https://sos.idaho.gov/elections/forms/voter_registration.pdf [https://perma.cc/M7VW-D8PB] ("Male/Female"). Compare 10 Ill. Comp. Stat. Ann. 5/5-7 (West 2023) (necessitating that applicants provide information about their sex to determine their identification for registering to vote), with Ill. State Bd. of Elections, Illinois Voter Registration Application (Oct. 2022), https://elections.il.gov/electionoperations/votingregistrationforms.aspx [https://perma.cc/6DJG-8YA7] ("Sex: M, F, X"). Compare Iowa Code § 48A.11(1)(g) (2023) (asserting that voter registration forms in Iowa must have an option for voter registration applicants to provide their sex), with Iowa Sec'y of State, State of Iowa Official Voter Registration Form, https://sos.iowa.gov/elections/pdf/voteapp.pdf [https://perma.cc/5Y5U-VHCE] (last updated Dec. 28, 2022) ("Sex: Male, Female"). Compare Kan. Stat. Ann. § 25-2309(b)(4) (West 2023) (enabling the collection of information about applicants' sex to register them as voters and prevent voter fraud), with Kan. Sec'y of State, Kansas Voter Registration Application, https://www.kssos.org/forms/elections/voterregistration.pdf [https://perma.cc/VCV9-H85N] (last updated Oct. 8, 2020) ("Male/Female"). Compare La. Stat. Ann. § 18:104(B)(1) (2023) (allowing applicants to provide information about their sex either of their own volition or after being prompted for additional information), with La. Sec'y of State, Louisiana Voter Registration Application, https://www.sos.la.gov/ElectionsAndVoting/PublishedDocuments/ApplicationToRegisterToVote.pdf [https://perma.cc/R9WP-46N2] (last updated June 2019) ("Sex: M, F"). Compare N.M. Stat. Ann. § 1-12-7.3(A)(2) (2023) (requiring that the voter registration checklist include the voter's gender), with N.M. Sec'y of State, Register to Vote (2015), https://portal.sos.state.

jurisdictions (including the District of Columbia) where the law is silent on whether sex or gender data are required to register to vote, five nevertheless have binary male/female options on their forms,[140] three ask

---

nm.us/ovr/VRForms/VRFormEnglishFinal.pdf [https://perma.cc/UV4J-34M4] ("Gender: ___"). Compare N.C. Gen. Stat. § 163-82.4(a)(6) (2023) (mandating the inclusion of gender in North Carolina's voter registration form), with N.C. State Bd. of Elections, North Carolina Voter Registration Application (Apr. 2023), https://dl.ncsbe.gov/Voter_Registration/NCVoterRegForm_06W.pdf [https://perma.cc/9NZC-KHTT] ("[O]ptional[:] Gender: Male, Female"). Compare S.C. Code Ann. §§ 7-5-170, -185 (2023) (instructing applicants to provide their sex in their application prior to officially becoming registered to vote), with S.C. Election Comm'n, South Carolina Voter Registration, https://scvotes.gov/wp-content/uploads/2023/08/SEC-FRM-1301-202305-VR-by-Mail-web-1.pdf [https://perma.cc/7TJT-YVKG] (last visited Oct. 7, 2023) ("Sex: Male, Female"). Compare Tenn. Code Ann. § 2-2-116 (2023) (maintaining that each applicant must provide their sex prior to being registered to vote), with Tenn. Sec'y of State, Tennessee Mail-In Application for Voter Registration, https://sos-tn-gov-files.s3.amazonaws.com/forms/ss-3010.pdf [https://perma.cc/RA4N-9SF2] (last updated Sept. 2020) ("Sex: M, F"). Compare Tex. Elec. Code Ann. § 13.122(a)(6) (West 2023) (necessitating space in the voter registration application form for applicants to fill out their sex), with Tex. Sec'y of State, Texas Voter Registration Application, https://www.sos.state.tx.us/elections/forms/vr-with-receipt.pdf [https://perma.cc/9LH9-47ZM] (last visited Aug. 24, 2023) ("Gender (Optional): Male, Female"). Compare Va. Code Ann. § 24.2-418(A) (2023) (mandating that voter registration applicants in Virginia provide their gender information), with Virginia Voter Registration Application (July 2020), https://www.elections.virginia.gov/media/formswarehouse/veris-voter-registration/applications/VA-NVRA-1-Voter-Registration-Application-rev-4_1-(1).pdf [https://perma.cc/2AMZ-G2A3] ("Gender: ___"). Compare W. Va. Code Ann. § 3-2-5(d)(8) (LexisNexis 2023) (allowing West Virginia voter registration applications to ask about gender but clarifying that applicants may not be rejected for choosing not to provide this information), with W. Va. Sec'y of State, West Virginia Voter Registration Application (June 2023), https://sos.wv.gov/FormSearch/Elections/Voter/mail%20in%20voter%20registration%20application.pdf [https://perma.cc/K9RF-B83G] (listing "gender" as an optional field and providing only "M, F" options). Compare Wyo. Stat. Ann. § 22-3-108(b)(viii) (2023) (allowing, but not mandating, applicants to provide their gender when registering to vote in the state of Wyoming), with Wyo. Sec'y of State, Wyoming Voter Registration Application and Change Form (Mar. 2020), https://sos.wyo.gov/Forms/Elections/General/VoterRegistrationForm.pdf [https://perma.cc/4MTK-8U4G] ("[O]ptional[:] Male, Female").

140. See Conn. Sec'y of State, State of Connecticut Mail-In Voter Registration, https://portal.ct.gov/-/media/SOTS/ElectionServices/ElectForms/electforms/ED-671-En-8x10-No-code.pdf [https://perma.cc/7HVS-JHLA] (last updated Sept. 2015) ("Gender: Male, Female"); Ind. Sec'y of State, Indiana Voter Registration Application (Mar. 2023), https://forms.in.gov/Download.aspx?id=9341 [https://perma.cc/R5PP-A5N4] ("Gender: Female, Male"); Ky. State Bd. of Elections, Commonwealth of Kentucky Mail-In Voter Registration Form, https://elect.ky.gov/registertovote/Documents/SBE%2001%20406%20Mail%20In%20Voter%20Registration%20Application.pdf [https://perma.cc/AE87-CPF5] (last updated Mar. 2020) ("Female, Male"); Mo. Sec'y of State, Missouri Voter Registration Application, https://s1.sos.mo.gov/cmsimages/ElectionGoVoteMissouri/register2vote/Adair.pdf [https://perma.cc/835P-RDLJ] (last updated Nov. 2022) ("Male, Female"); N.J. Div. of Elections, New Jersey Voter Registration Application, https://www.state.nj.us/state/elections/assets/pdf/forms-voter-registration/68-voter-registration-english.pdf [https://perma.cc/J7XY-HYNS] (last updated Jan. 9, 2020) ("Gender (Optional): Female, Male").

registrants to select gendered salutations,[141] and only five include the option to select "Unspecified/Other" in response to a question about gender.[142] Civil servants made these forms, and the result of their work means that—as broad-based empirical studies have shown—the gender binary is for the most part entrenched at the implementation level.[143]

Of course, governments do not collect all this information on their own. They also buy it from the private sector.[144] Gender data purchased on the open market are also likely to reflect the gender binary. Despite high-profile examples of digital platforms adding multiple checkboxes to answer gender questions,[145] those same platforms only allow advertisers to target users based on binary gender categories (male, female, or all).[146]

---

141. See Ark. Sec'y of State, Arkansas Voter Registration Application, https://www.sos.arkansas.gov/uploads/elections/ArkansasVoterRegistrationApplication.pdf [https://perma.cc/SKP9-HNH9] (last updated Jan. 24, 2019); Cal. Sec'y of State, Classification—Voter Registration Application, https://covr.sos.ca.gov/ [https://perma.cc/E7D8-PHES] (last visited Nov. 5, 2023) ("[Optional] Prefix: Mr., Mrs., Ms. Ms."); State of Connecticut Mail-In Voter Registration, supra note 140.

142. See Md. State Bd. of Elections, Maryland Voter Registration Application, https://elections.maryland.gov/voter_registration/documents/English_Internet_VRA.pdf [https://perma.cc/39CA-AKKA] (last updated Mar. 2023) ("Gender: Male, Female, Unspecified or Other"); Mich. Sec'y of State, State of Michigan Voter Registration Application, https://www.michigan.gov/sos/- /media/Project/Websites/sos/Elections/Election-Forms/Voter-Registration-FormEnglish.pdf [https://perma.cc/H6E8-28DS] (last updated July 2023) ("Female (f), Male (m), Non-binary (x)"); N.Y. Bd. of Elections, New York State Voter Registration Form, https://www.elections.ny.gov/NYSBOE/download/voting/voteregform-eng-fillable.pdf (on file with the *Columbia Law Review*) (last visited Nov. 5, 2023) ("[O]ptional . . . Gender: ____"); Pa. Dep't of State, Pennsylvania Voter Registration Application & Mail-in Ballot Request, https://www.vote.pa.gov/Resources/Documents/Voter_Registration_Application_English.pdf [https://perma.cc/RE44-QQDF] (last visited Aug. 24, 2023) ("Gender[:] Female (F), Male (M), Non-Binary/Other (X)"); Wash. Sec'y of State, Washington State Voter Registration Form, https://www.sos.wa.gov/sites/default/files/2023-07/VRF_English.pdf?uid=6546b07c67589 [https://perma.cc/M7FJ-MHDW] (last updated Mar. 2023) ("[G]ender: ____").

143. Waldman, The Gender Box, supra note 132, at 5 (discussing how civil servants play a role in pre-determining the options on administrative forms).

144. See Julie E. Cohen, The Inverse Relationship Between Secrecy and Privacy, 77 Soc. Rsch. 883, 885 (2010) (discussing how the federal government acquires data from private entities); Joel Reidenberg, The Transparent Citizen, 47 Loy. U. Chi. L.J. 437, 452 (2015) (same); Sara Morrison, A Surprising Number of Government Agencies Buy Cell Phone Location Data. Lawmakers Want to Know Why, Vox (Dec. 2, 2020), https://www.vox.com/recode/22038383/dhs-cbp-investigation-cellphone-data-brokers-venntel [https://perma.cc/23BD-7AB3] (same).

145. Rhiannon Williams, Facebook's 71 Gender Options Come to UK Users, Telegraph (June 27, 2014), http://www.telegraph.co.uk/technology/facebook/10930654/Facebooks-71-gender-options-come-to-UK-users.html [https://perma.cc/79S3-FBRT] (discussing the seventy-one gender options available to Facebook users).

146. Facebook EEOC Complaints: Charge of Discrimination, ACLU (Sept. 18, 2018), https://www.aclu.org/legal-document/facebook-eeoc-complaint-charge-discrimination

They recode nonbinary individuals within the gender binary on the back end.[147] The private sector also packages clusters of users into categories based on gender.[148] We know little about the secretive data broker industry, so we can only surmise that it is likely that data brokers follow the gender binary as well.

Even if it were possible to systematically make gender data collection more inclusive (for many reasons discussed below, doing so is not the answer to the harms caused by gender data collection by the state[149]), the law is not done binarizing gender data streams after mandating collection. As the next Part describes, the law also determines how that data will be shared in the automated state, privileging the gender binary along the way.

III. LAW AND THE SHARING OF BINARY GENDER DATA

Data from official government forms replicate and spread throughout the automated administrative state. As Professor Bridget Fahey notes, data are nonrivalrous and complementary: The same data can be used by multiple agencies without interfering with anyone's access, and datasets increase in value as they increase in size by giving the state the means to learn more about the people it surveils.[150] Large datasets are now cheap to store and easy to copy. They are even easier to use now that sophisticated AI systems are just a procurement contract away.[151] Gender data are no different.

But the replication of binary gender data across state agencies and across states is not merely a feature of modern technology. It is also a product of the law. In addition to requiring the collection of gender data, state law often requires agencies to share the data with other departments, spreading the gender binary across government bureaucracies. State agencies agree to share gender data with each other under memorandums of understanding (MOUs).[152] There are also interstate compacts and federal funding rules that require states to share data with other states, coordinated bureaucracies, and the federal government. These data-

---

[https://perma.cc/LU59-6FQ7] (explaining that Facebook offers the gender categories of "All," "Male," and "Female").

147. Bivens, supra note 40, at 891–93 (explaining Facebook's invisible gender-recoding process).

148. Bruce Schneier, Data and Goliath 63 (2015).

149. See infra notes 355–374 and accompanying text.

150. Fahey, Data Federalism, supra note 16, at 1072–73.

151. See infra section IV.B.

152. Arkansas, Delaware, Kentucky, Tennessee, and Virginia allow only residents of those states to submit public records requests or receive documents. See Ark. Code Ann. § 25-19-105(a)(1)(A) (2023); Ky. Rev. Stat. Ann. § 61.872(1)–(3) (West 2021); Tenn. Code Ann. § 10-7-503(a)(2)(A) (2023); Va. Code Ann. §§ 2.2-3700, -3701 (2023); Op. Att'y Gen. No. 96-IB01, at 2 (Del. Jan. 2, 1996), 1996 WL 40922 (interpreting Del. Code Ann. tit. 29, §§ 10001, 10003 (1995) to apply "only to Delaware citizens").

sharing mandates, agreements, and MOUs include gender information that has already been binarized at the front end by perceptions of common sense and frontline civil servants. By sharing those data, the law entrenches and normalizes the gender binary, conflates sex and gender, and creates data-driven systems that function only on binary gender data.

A. *Laws and Rules Requiring Gender Data Sharing*

On the premise that larger and more detailed datasets are more valuable than smaller ones,[153] many state laws either require interagency data sharing about individuals or permit agencies to enter into data sharing agreements in order to achieve administrative goals. Many of these laws focus on children and families. For instance, Pennsylvania requires agencies to share the "contents of county agency, juvenile probation department, drug and alcohol, mental health and education records" about any child in protective services "to enhance the coordination of case management" and "disposition."[154] This dataset includes demographic information about the child.[155] Louisiana law envisions the creation of data-sharing agreements among state agencies "involved in the assessment, diagnosis, treatment, care, or rehabilitation of children."[156] Those health records include sex data.[157] So too would any data shared among state and federal agencies to implement health exchanges under the Affordable Care Act.[158]

Criminal justice laws frequently include gender data-sharing mandates. California's Monthly Arrest and Citation Register includes binary gender in its "personal characteristics."[159] The state's Juvenile Court and Probation Statistical System tracks the binary sex of everyone passing through the state juvenile criminal justice system.[160] And the California Youth Authority's Offender-Based Information Tracking System

153. See Fahey, Data Federalism, supra note 16, at 1073.

154. 42 Pa. Stat. and Cons. Stat. Ann. § 6352.2 (West 2023).

155. Id.

156. La. Child. Code Ann. art. 545 (2023).

157. Id.

158. See, e.g., Colo. Rev. Stat. Ann. § 10-22-106(2) (2023); Ind. Code Ann. §§ 27-19-1-4(3), -3-3(e)(2) (West 2023); Md. Code Ann., Ins. § 31-106 (West 2023); Va. Code Ann. § 38.2-6512 (2023).

159. See Letter from Danielle Brousseau, Staff Servs. Manager I, Cal. Just. Info. Servs. Div., to author (Sept. 26, 2022) (on file with the *Columbia Law Review*) [hereinafter Brousseau Letter] (noting that categories of gender data collected are "male and female" only); Data Portal, Open Just., https://openjustice.doj.ca.gov/data [https://perma.cc/867D-ET7J] (last visited Aug. 24, 2023) (select "Arrests - CSV" under "Criminal Justice Data"); see also John L. Worrall & Pamela Schram, Sch. Behav. & Soc. Scis., Cal. State Univ., San Bernardino, Evaluation of California's State-Level Data Systems for Incarcerated Youth 20 (2000), https://sor.senate.ca.gov/sites/sor.senate.ca.gov/files/ctools/%7B3F3F9617-9598-4DD5-AA4F-E5DCFD8A8A67%7D.PDF [https://perma.cc/WW33-VYTP].

160. Worrall & Schram, supra note 159, at 21; Brousseau Letter, supra note 159.

extracts the binary sex of juvenile offenders across all California jurisdictions from the state's Automated Criminal History System.[161]

California also has many statutorily created education- and health-related data-sharing programs that limit gender data to the binary. The state's Longitudinal Pupil Achievement Data System collects discipline and achievement data on all students in both general and special education programs.[162] Its demographic dataset includes gender.[163] And the state's Cradle to Career Data System Act authorized the creation of a system-wide database that uses gender, among other data points, to help students and families successfully transition from California K–12 schools to college and the workforce.[164] Notably, California includes a nonbinary gender option in annual reports about students who graduate from the state's public schools and meet state university entry requirements.[165]

Then there are laws that require regulatory agencies to use data-sharing agreements to enforce the law and to verify identity. The Louisiana Gaming Control Board is authorized by state law to enter into agreements that would, among other things, share information from workers' "personal history forms" to ensure they are who they say they are.[166] Those forms only allow workers to enter "M" or "F" in response to a question about sex.[167] And Montana requires its chief elections official to enter into data-sharing agreements with the state's department of motor vehicles to "verify voter registration information."[168] Both departments collect only

---

161. Worrall & Schram, supra note 159, at 21; Brousseau Letter, supra note 159.

162. California Longitudinal Pupil Achievement Data System (CALPADS), Cal. Dep't of Educ., http://www.cde.ca.gov/ds/sp/cl/ [https://perma.cc/P25B-M5HJ] (last visited Aug. 24, 2023).

163. See CALPADS Background/History, Cal. Dep't of Educ., https://www.cde.ca.gov/ds/sp/cl/background.asp [https://perma.cc/5BKG-XSCG] (last visited Aug. 24, 2023); Data Reports by Topic, Cal. Dep't of Educ., https://www.cde.ca.gov/ds/ad/accessdatasub.asp [https://perma.cc/LN56-2RYP] (last visited Aug. 24, 2023) (providing information "disaggregated by race/ethnicity, gender, and program subgroup").

164. Cal. Educ. Code §§ 10850–10874 (2022); see also California Cradle-to-Career Data System, State of Cal., https://c2c.ca.gov/ [https://perma.cc/97UW-E5PF] (last visited Aug. 24, 2023).

165. Data Rep. Off., Cal. Dep't of Educ., 2020–21 Four-Year Adjusted Cohort Graduation Rate: Statewide Report, DataQuest, https://dq.cde.ca.gov/dataquest/dqcensus/CohRate.aspx?cds=00&agglevel=state%20&year=2020-21 [https://perma.cc/VM39-VSL2] (last visited Aug. 24, 2023) (allowing filter by "male," "female," "nonbinary," or "missing").

166. La. Stat. Ann. § 27:45(A), (C) (2023).

167. See Email from Margot Lassit, La. Gaming Control Bd., to author (July 14, 2022) (on file with the *Columbia Law Review*) (confirming that only male/female options are accepted); see also La. Gaming Control Bd., Multijurisdictional Personal History Disclosure Form, https://dpsweb.dps.louisiana.gov/gamingforms.nsf/fdcf9e5f850b2bc78625731b006934c6/7cf5e544b362ce49862575830062fd2b/$FILE/Multi%20Jurisdictional%20Personal%20History%20Diclosure%20Form.pdf [https://perma.cc/9ZG6-N4L5] (last visited Aug. 24, 2023).

168. Mont. Code Ann. § 13-2-107(3)(a) (West 2023).

binary sex data.[169] In Oklahoma, leaders at several state agencies have arranged to share gender data with the State Election Board, including the Department of Health (death records), court clerks (lists of convicted felons), and the Department of Public Safety (voter registration).[170]

These data-sharing laws create what Professor Fahey calls "data pools": aggregations of information collected for a variety of purposes by other agents of the state.[171] Data pools "aggregate power and diffuse access" by allowing more state agencies to more intensively track, surveil, and verify identities.[172] When the laws sweep in sex and gender data, they do not always specify what that data should look like; rather, that depends on how the state agency decided to collect the data in the first place and how technical systems are programmed to use the data in the end. As we have seen, because the vast majority of that data is collected along binary lines, data-sharing mandates replicate the gender binary throughout the government's larger data ecosystem.

B. *Interagency Agreements*

Interagency data-sharing agreements supplement statutory data-sharing mandates, replicating binary gender in the same way. Although many statutes permit data-sharing agreements involving the transfer of personal data,[173] engaging with other departments and other states is often up to the agencies themselves. This type of lawmaking is more informal but no less binding on agency behavior. And many of these agreements include gender data to be used for a variety of purposes—identifying individuals and detecting fraud, conducting research, or implementing the law—or, in some cases, for no stated purpose at all.[174] In almost all cases, the agreements are broad and traffic in binary gender data.

---

169. See Motor Vehicle Div., State of Montana Application for Class D Driver License or Identification Card, https://dojmt.gov/wp-content/uploads/11-1400-Application-for-Class-D-Driver-License-and-Application-for-Identification-Card-0723v2-Fillable.pdf [https://perma.cc/K29C-J6YV] (last visited Aug. 24, 2023).

170. Okla. Stat. tit. 26, § 4-109.3A (2023) (voter registration); Id. § 4-120.3A (death records); Id. § 4-120.4A (felons).

171. Fahey, Data Federalism, supra note 16, at 1012.

172. Id.

173. E.g., Conn. Gen. Stat. Ann. § 9-50c(a) (West 2023) ("The Secretary of the State may enter into an agreement to share information or data with any other state . . . ."); 105 Ill. Comp. Stat. Ann. 13/25(b) (West 2023) (providing that "[a]ny State agency, board, authority, or commission may enter into a data sharing arrangement" as part of implementing the Longitudinal Education Data System Act); Wash. Rev. Code Ann. § 50A.25.070(1) (West 2022) ("The department may enter into data-sharing contracts and may disclose records and information deemed confidential to state or local government agencies . . . .").

174. This section is based on the results of public records requests sent to three departments—the chief election division, the motor vehicle division, and the division that administers professional licensure—in forty-five states and the District of Columbia.

Many state agencies share binary gender data with the goal of detecting fraud and verifying identity.[175] Departments of motor vehicles (DMVs) and those in charge of elections and voter registration share data frequently to verify identity for benefits programs.[176] DMVs share data with boards of elections to assist with voter registration.[177] To verify identities, DMVs distribute binary gender data to fishing and hunting licensure divisions,[178] organ donor registries,[179] departments of veterans' affairs,[180] police departments,[181] municipal courts dealing with traffic violations,[182]

Arkansas, Delaware, Kentucky, Tennessee, and Virginia only allow residents of those states to submit public records requests and receive documents, see supra note 152; therefore, those states were excluded. Additional research could cover additional divisions of state government.

175. See, e.g., Driver License Data Verification System Jurisdiction Service Agreement Between Vt. Dep't of Motor Vehicles and Am. Ass'n of Motor Vehicle Adm'rs cls. 1 & 3(B)(xii) (Aug. 23, 2018) (on file with the *Columbia Law Review*) (sharing driver license data, including gender, with the American Association of Motor Vehicle Administrators (AAMVA), a nonprofit that provides participating states with a nationwide database against which to verify the identities of those seeking licenses); Data Licensing Agreement for Driver Record Information Between [Wash.] Dep't of Licensing and [Wash.] Emp. Sec. Dep't 12 (Mar. 19, 2019) (on file with the *Columbia Law Review*) [hereinafter Wash. Driver Record Agreement] (sharing license data, including gender, "for the purposes of fraud investigations").

176. E.g., Memorandum of Understanding Between Iowa Dep't of Transp., Motor Vehicle Div., and Iowa Dep't of Hum. Servs. 1 (July 19, 2022) (on file with the *Columbia Law Review*); Memorandum of Understanding Between R.I. Dep't of State and R.I. Div. of Motor Vehicles 1 (June 13, 2016) (on file with the *Columbia Law Review*) [hereinafter R.I. DMV Agreement]; Wash. Driver Record Agreement, supra note 175, at 12.

177. E.g., R.I. DMV Agreement, supra note 176, at 1; Data Sharing Memorandum of Understanding Between Vt. Dep't of Motor Vehicles and Vt. Sec'y of State 1 (Sept. 8, 2021) (on file with the *Columbia Law Review*).

178. E.g., Memorandum of Understanding Between Iowa Dep't of Transp., Motor Vehicle Div., and Iowa Dep't of Nat. Res. 4 (Oct. 1, 2021) (on file with the *Columbia Law Review*) [hereinafter Iowa DNR MOU].

179. E.g., Contract for Acquisition of Records in Bulk for Permissible Purposes Between Idaho Transp. Dep't and DonorConnect 2 (Oct. 12, 2021) (on file with the *Columbia Law Review*) [hereinafter Idaho DonorConnect Contract].

180. E.g., Memorandum of Agreement Between the Idaho Transp. Dep't and the Idaho Div. of Veteran Servs. 1 (Sept. 17, 2020) (on file with the *Columbia Law Review*).

181. E.g., Memorandum of Agreement for Use of Records Among N.C. Dep't of Transp., Div. of Motor Vehicles, Dep't of N.C. Pub. Safety, State Highway Patrol, and Interplat Solutions, Inc. 11 (Sept. 5, 2022) (on file with the *Columbia Law Review*); Wash. State Dep't of Licensing, DSC-425-009, Moxee Police Dep't, Driver and Plate Search (DAPS) and Driver Information and Internet Query System (IHPS) Agency Access Request 3 (Oct. 5, 2016) (on file with the *Columbia Law Review*). Public records requests resulted in more than 217 identical or similar agreements with different police departments and federal investigative units.

182. E.g., Interagency Data Sharing Agreement Between [Wash.] Dep't of Licensing and Wash. State Admin. Off. of the Cts. 13 (July 9, 2019) (on file with the *Columbia Law Review*).

and departments of social services.[183] And all of these agreements include gender data.

States that share borders with Canada or Mexico exchange all data on Enhanced Driver's Licenses with the Department of Homeland Security for border security purposes.[184] DMVs also share gender data with departments, like those responsible for enforcing child support orders, that can order driver's license suspensions for people who fail to meet their obligations.[185] When the departments originating the data collect only binary sex and gender information, only male/female data can be shared.

A second cluster of interagency agreements that share gender data focuses on research. Rhode Island shares voter registration data, including the identification information provided at registration, with Brown University's Rhode Island Innovative Policy Lab for research into how voter identification requirements impact registration and turnout rates.[186] Iowa shares binary sex and gender data with the University of Northern Iowa to "assist in identifying any health disparities . . . for those seeking treatment for problem gambling and/or substance abuse disorders."[187] In both cases, sex and gender data are exclusively binary.

State agencies also share sex and gender data with divisions of criminal justice, schools, and health to, among other things, "carry[] out . . . investigations [and] prosecutions of criminal offenses."[188] In Washington State, for example, the automobile licensing division shares gender data with all "authorized criminal justice authorities throughout the state" for general use.[189] North Carolina's FAST Program, which facilitates the state department of health's provision of social services to families, has collected gender data from the state's DMV since 2013.[190]

---

183. E.g., Memorandum of Agreement for Secure Online Access to Information Between Mo. State Emps.' Ret. Sys. (MOSERS) and Mo. Dep't of Revenue 1 (Mar. 12, 2014) (on file with the *Columbia Law Review*).

184. E.g., Addendum to the Memorandum of Agreement Between State of Vt. and DHS 1 (Mar. 15, 2017) (on file with the *Columbia Law Review*).

185. E.g., Memorandum of Agreement Between Idaho Transp. Dep't and Idaho Dep't of Health & Welfare l–2 (n.d.) (on file with the *Columbia Law Review*).

186. Cooperation and Data Sharing Agreement Between R.I. Innovative Pol'y Lab at Brown Univ. and R.I. Dep't of State 6 (May 30, 2018) (on file with the *Columbia Law Review*).

187. Monitoring and Evaluation Contract, Special Conditions for Contract #5882BH11 Between Iowa Dep't of Pub. Health and Univ. of N. Iowa 3–4 (Sept. 27, 2021) (on file with the *Columbia Law Review*).

188. Data Sharing Agreement Between Wash. Dep't of Licensing and Wash. Att'y Gen.'s Off. 15 (Mar. 9, 2020) (on file with the *Columbia Law Review*).

189. Contract Between Wash. State Dep't of Licensing and State of Wash. Admin. Off. of the Cts. 5 (Sept. 30, 2017) (on file with the *Columbia Law Review*).

190. Memorandum of Understanding Between N.C. Div. of Motor Vehicles and N.C. Dep't of Health & Hum. Servs. attachs. 1, 2 (Sept. 25, 2013) (on file with the *Columbia Law Review*).

These are just a handful of examples available through public record requests. But data-sharing agreements are common arrangements among a variety of agencies. Including agreements signed between 2016 and 2022, the Florida Department of Highway Safety and Motor Vehicles is currently a party to at least 1,172 active data-sharing agreements with state agencies, agencies in other states, the federal government, or private entities.[191] The Washington Department of Licensing has data-sharing agreements for driver data—which include gender—with at least 349 other agencies.[192]

C. *Interstate Compacts and Data Federalism*

There are also explicit intergovernmental dependencies that spread sex and gender data throughout the government data ecosystem.[193] For instance, state agencies have agreed to share binary sex and gender data with other departments and the federal government to determine eligibility for public benefits programs, including the Tenant Rental Assistance Certification System (TRACS) and the Supplemental Nutrition Assistance Program (SNAP).[194] Federal funding for state agencies involved in coordinating foster care programs is also tied to a long-running data-sharing agreement in which states must report children's sex as either "male" or "female."[195]

---

191. See Fla. Dep't of Highway Safety & Motor Vehicles, Florida Data-Listing Unit MOUs (Aug. 9, 2022) (on file with the *Columbia Law Review*) (listing 1,172 active agreements with contract effective dates between 2016 and 2022).

192. See Washington Dep't of Licensing, DIAS Account List (n.d.) (on file with the *Columbia Law Review*) (listing 349 accounts).

193. Professor Fahey chronicled many of these but did not focus on whether—or how—they shared gender data. See Fahey, Data Federalism, supra note 16, at 1016–29.

194. See Computer Matching Agreement Among HHS, Admin. for Child. & Fams., Off. of Child Support Enf't, and State Agency Administering the Supplemental Nutrition Assistance Program 7 (Aug. 16, 2021), https://www.hhs.gov/sites/default/files/acf-snap-cma-2111.pdf [https://perma.cc/3RSW-CXBC] (noting that "sources of records used" in the matching program include "information collected by the state agency in its administration of SNAP"); Off. of Hous., HUD, Tenant Rental Assistance Certification System (TRACS): Privacy Impact Assessment 8 (2009), https://www.hud.gov/sites/documents/TRACS.PDF [https://perma.cc/6PRM-3GZR] (noting that "Gender/sex" is collected by the TRACS systems) [hereinafter TRACS PIA]. SNAP applications collect sex data with only male/female answer options. See, e.g., N.Y. State Off. of Temp. & Disability Assistance, SNAP Application/Recertification 3 https://otda.ny.gov/programs/applications/4826.pdf [https://perma.cc/X83C-M3D7] (last visited Aug. 24, 2023) (noting that applicants and members of their household should designate their sex only as "M" or "F"); Tex. Health & Hum. Servs., Your Texas Benefits: Getting Started 3–5 (June 22, 2022), https://yourtexasbenefits.com/GeneratePDF/StaticPdfs/en_US/H1010_June_22_FINAL.pdf [https://perma.cc/8DQA-V6JP] (asking applicants to select either "male" or "female").

195. See About AFCARS, HHS, https://www.acf.hhs.gov/cb/resource/about-afcars [https://perma.cc/ED6N-BXEX] (last visited Aug. 24, 2023) (describing the Adoption and Foster Care Analysis and Reporting System (AFCARS)); see also 45 C.F.R. § 1355.44(b)(2) (2020) ("Child's sex. Indicate whether the child is 'male' or 'female.'").

All states participate in the CDC's National Notifiable Disease Surveillance System (NNDSS), a "passive surveillance system" that collects data from state health departments on incidents or outbreaks of more than 120 diseases.[196] The NNDSS collects gender data chaotically: Each division within the CDC designs sample forms for the reportable diseases in its portfolio. Its Adult and Pediatric HIV/AIDS Confidential Case Report Forms, which are used in at least eleven states, asks for individuals' "sex assigned at birth" with "male," "female," and "unknown" answer options, as well as "gender identity" with a variety of inclusive options.[197] Many of the CDC's other disease surveillance forms ask for "sex" with just three answer options,[198] and its Multisystem Inflammatory Syndrome Associated With COVID-19 Form asks for "sex" but provides only "male" and "female" answer options.[199]

Twenty-five states and the District of Columbia are part of the Electronic Registration Information Center (ERIC), a nonprofit corporation that helps states improve voter roll accuracy and increase

---

196. Sandra Roush, Enhancing Surveillance, *in* Manual for the Surveillance of Vaccine-Preventable Diseases ch. 19-1 (5th ed. 2011); see also Lawrence Gostin, Public Health Law: Power, Duty, Restraint 296 (2d ed. 2008); National Notifiable Diseases Surveillance System (NNDSS), CDC, https://www.cdc.gov/nndss/ [https://perma.cc/GZK6-SN7B] (last visited Aug. 24, 2023).

197. See CDC, Adult HIV Confidential Case Report Form (Nov. 2019), https://www.cdc.gov/hiv/pdf/guidelines/cdc-hiv-adult-confidential-case-report-form-2019.pdf [https://perma.cc/5YDV-5224]; CDC, Pediatric HIV Confidential Case Report Form (Nov. 2019), https://www.cdc.gov/hiv/pdf/guidelines/cdc-hiv-pediatric-confidential-case-report-form-2019.pdf [https://perma.cc/AXR4-B7XL]. West Virginia uses the CDC's Pediatric and Adult HIV Report Forms. W. Va. Dep't of Health & Hum. Res. Bureau for Pub. Health, Adult HIV Confidential Case Report Form (Nov. 2019), https://oeps.wv.gov/hiv-aids/Documents/lhd/adultHIVcaseReport_fillable.pdf [https://perma.cc/FPW3-RM2E]; W. Va. Dep't of Health & Hum. Res. Bureau for Pub. Health, Pediatric HIV Confidential Case Report Form (Nov. 2019), https://oeps.wv.gov/hiv-aids/Documents/lhd/pediatricHIVcaseReport_Fillable.pdf [https://perma.cc/4UPA-J5JA].

198. Under federal vocabulary standards for electronic health information set by the HHS Secretary, "[b]irth sex must be . . . attributed as follows: (i) Male. M, (ii) Female. F, (iii) Unknown . . . . UNK . . . ." 45 C.F.R. § 170.207(n) (2022) (emphasis omitted). For examples of CDC disease surveillance forms that follow this standard, see CDC, OMB No. 0920-0728, Babesiosis Case Report Form (2016), https://www.cdc.gov/parasites/babesiosis/resources/50.153.pdf [https://perma.cc/3WX8-3NJ8]; CDC, OMB No. 0920-0728, Brucellosis Case Report Form (2021), https://www.cdc.gov/brucellosis/pdf/case-report-form.pdf [https://perma.cc/3CBM-72EG]; CDC, Meningococcal Disease Surveillance Worksheet (2021), https://www.cdc.gov/ncird/surveillance/downloads/Meningococcal-Worksheet-2021-annot-508.pdf [https://perma.cc/8W39-XUJ8]; CDC, OMB No. 0920-0728, Tularemia Case Investigation Report (2016), https://www.cdc.gov/tularemia/resources/TularemiaCaseReportForm.pdf [https://perma.cc/ZV39-6W9S].

199. CDC, Multisystem Inflammatory Syndrome Associated With SARS-CoV-2 Infection Case Report (2022), https://www.cdc.gov/mis/pdfs/MIS-C_case-report-form.pdf [https://perma.cc/6NUA-SN6X].

access to voter registration.[200] Twenty of the current twenty-six ERIC members explicitly collect sex or gender data during the voter registration process, and all of them collect it when individuals apply for driver licenses.[201] Only two of those states allow gender designations other than "male" or "female".[202]

The National Crime Information Center (NCIC), which "anchors the intergovernmental exchange of information for day-to-day policing,"[203] allows law enforcement to cross-check information on license plates and identifications with various law enforcement databases. Within the NCIC system, the Interstate Identification Index (III) includes, among other things, a person's "sex" with male, female, and unknown coding options.[204] Similarly, the National Instant Criminal Background Check System, which allows federal or state agents to run background checks on individuals before firearm purchases, leverages only binary sex information for identity verification purposes.[205] And the National Adult Mistreatment Reporting System gathers information about perpetrators of elder abuse, including the genders of victims. This data is reported annually, broken down by "men" and "women."[206]

Several interstate compacts include gender data and privilege the gender binary.[207] For instance, all fifty states and the District of Columbia are part of the Interstate Compact on Juveniles, a contract that has been

---

200. Who We Are, Elec. Registration Info. Ctr., https://ericstates.org/who-we-are/ [https://perma.cc/SBS4-LQL3] (last visited Sept. 12, 2023).

201. See supra notes 95–98 and accompanying text.

202. See supra notes 95–98 and accompanying text.

203. Fahey, Data Federalism, supra note 16, at 1022.

204. Nat'l Crime Info. Ctr. (NCIC), FBI, DOJ, https://irp.fas.org/agency/doj/fbi/is/ncic.htm [https://perma.cc/39S4-KSZP] (last updated June 2, 2008) (describing the National Instant Criminal Background Check System, which allows federal or state agents to run background checks on individuals before firearm purchases, taps into the NCIC and III, and leverages sex information for identify verification purposes); see also FBI, DOJ, Interstate Identification Index/National Fingerprint File Operational Technical Manual, ch. 2, at 1, 7–9 (2005) (coding only for "male," "female," and "unknown").

205. See FBI, DOJ, National Instant Criminal Background Check System Operational Report 2020–2021, at 6 (2022), https://www.fbi.gov/file-repository/nics-2020-2021-operations-report.pdf/view [https://perma.cc/86PB-WNXW].

206. Nat'l Adult Maltreatment Reporting Sys., Adult Maltreatment Report 2020, at 22 (2020), https://acl.gov/sites/default/files/programs/2021-10/2020_NAMRS_Report_ADA-Final%20%281%29.pdf [https://perma.cc/M5M6-AH7U].

207. Interstate compacts are binding agreements between states. Bridget A. Fahey, Federalism by Contract, 129 Yale L.J. 2326, 2351 (2020). They are both statutes and contracts: statutes in each jurisdiction; contracts between them. Frederick L. Zimmermann & Mitchell Wendell, The Law and Use of Interstate Compacts 1 (1961). The Supreme Court has long held that interstate compacts are interpreted according to contract law principles but remain "law[s] of the United States." Tarrant Reg'l Water Dist. v. Herrmann, 569 U.S. 614, 627 n.8 (2013) (internal quotation marks omitted) (quoting Virginia v. Maryland, 540 U.S. 56, 66 (2003)).

adopted as law regulating the interstate movement of minors under court supervision or who have run away to another state.[208] The Compact requires those staffing its administrative body, the Interstate Commission for Juveniles, to "establish a system of uniform data collection on information pertaining to juveniles."[209] Therefore, the Commission, not individual states, dictates how the data should be gathered.[210] Six of the Compact's ten approved forms ask for sex, with "male," "female," and "unknown" answer options.[211] All participating jurisdictions must follow that protocol.

D. *Entrenching the Gender Binary at Data Sharing*

Just like the law of data collection, data-sharing mandates and more informal interagency agreements entrench the gender binary by making similar assumptions about gender data as static, secure identifiers. But the law of data sharing goes further. It solidifies the gender binary throughout the government's data ecosystem in three ways: Data-sharing agreements have expressive, conflationary, and interoperability effects.

As it spreads gender data, data-sharing law generates expressive and normalizing effects, framing how anyone who sees and uses the data understands sex and gender.[212] As many scholars have argued, law is an instrument of norm production that influences people's behavior indirectly by signaling what society thinks is right or wrong.[213] In other words, law has an "expressive function"[214] that creates "cultural consequences."[215] Professor Dan Kahan has argued that "gentle nudge[s]"

---

208. Christopher Holloway, DOJ, Interstate Compact on Juveniles 1 (2000), https://www.ojp.gov/pdffiles1/ojjdp/fs200012.pdf [https://perma.cc/7CGG-8EUA].

209. Interstate Compact for Juveniles art. I, cl. J (2014), https://juvenilecompact.org/sites/default/files/ICJRevisedLanguage.pdf [https://perma.cc/QZB3-2F8G]; see also id. art. III, cl. K; id. art. IV, cl. 19.

210. See Approved Forms, Interstate Comm'n for Juvs., https://www.juvenilecompact.org/forms [https://perma.cc/E7YP-698M] (last visited Aug. 24, 2023) (detailing that states must use Commission-approved information systems when collecting data pursuant to the Interstate Compact for Juveniles).

211. Id. (listing Commission-approved forms, including six that require sex data: Forms I, II, III, IV, and VII).

212. Flynn, supra note 88, at 466.

213. See, e.g., Citron, Expressive Value, supra note 88, at 377; Sunstein, supra note 88, at 2022–24.

214. Sunstein, supra note 88, at 2024; see also Deborah Hellman, The Expressive Dimension of Equal Protection, 85 Minn. L. Rev. 1, 39–40 (2000) (arguing that "to treat people with equal concern, government must attend to the expressive dimension of its actions").

215. Richard H. Pildes, The Unintended Cultural Consequences of Public Policy: A Comment on the Symposium, 89 Mich. L. Rev. 936, 938 (1991); see also Elizabeth S. Scott, Social Norms and the Legal Regulation of Marriage, 86 Va. L. Rev. 1901, 1902–03 (2000); Tokson & Waldman, supra note 88, at 281.

can incrementally change existing social norms by encouraging individuals to "revise upward" or downward "their judgment of the degree of condemnation warranted by the conduct in question."[216] Data streams created and maintained by law are no different. The more binary gender data spreads, the more people will encounter it, and the more power it will have to reify sex and gender as binary and static. In this way, laws that spread binary gender data normalize it as true and correct; they facilitate elision between frequency and propriety, nudging us to think that the things we see often—male/female-only categories—are the normal, commonsense ways to conceptualize and classify by sex and gender.[217]

Many of these agreements also conflate sex and gender. For instance, although the Iowa DMV collects sex data only from applicants for licenses and identification cards,[218] its data-sharing agreement with the state's Department of Natural Resources refers to sharing gender data.[219] Idaho makes the same mistake in its MOU with the state's organ donor registry.[220] More than half of the relevant interagency agreements provided under public records requests conflate sex and gender.[221]

Doing so helps reify the gender binary. Sex is primarily a matter of chromosomes or genital anatomy; gender is primarily a matter of social expectations and performance.[222] Sex and gender are undoubtedly

---

216. Dan M. Kahan, Gentle Nudges vs. Hard Shoves: Solving the Sticky Norms Problem, 67 U. Chi. L. Rev. 607, 610–11 (2000).

217. Normalization is cognitive slippage from statistical frequency to moral propriety; it is a process through which common things come to be understood as acceptable, ordinary, and, ultimately, good. See Adam Bear & Joshua Knobe, Normality: Part Descriptive, Part Prescriptive, 167 Cognition 25, 25 (2017) [hereinafter Bear & Knobe, Normality]. Political scandals are good examples of this phenomenon. As psychologists Adam Bear and Joshua Knobe have written, when a politician "continues to do things that once would have been regarded as outlandish, [their] actions are not simply coming to be regarded as more typical; they are coming to be seen as more normal[,] . . . as less bad and hence less worthy of outrage." Adam Bear & Joshua Knobe, Opinion, The Normalization Trap, N.Y. Times (Jan. 28, 2017), https://www.nytimes.com/2017/01/28/opinion/sunday/the-normalization-trap.html (on file with the *Columbia Law Review*); see also Diane Vaughan, The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA 77–195 (1996) (demonstrating how routinized decisions that violated rules and norms came to be normalized as part of engineering and testing work).

218. Iowa Code § 321.182 (2023).

219. Iowa DNR MOU, supra note 178, at sched. A.

220. Compare Idaho Code § 49-306 (2023), with Idaho DonorConnect Contract, supra note 179, at 2. It could be argued that this change from sex to gender reflects bureaucratic discretion or an agency exercising its delegated power to implement the law through its unique expertise. See Edward H. Stiglitz, Delegating for Trust, 166 U. Pa. L. Rev. 633, 635 (2018) (noting that the primary justification for the administrative state is agency expertise).

221. See supra section III.B.

222. See Glossary of Terms: Transgender, GLAAD, https://glaad.org/reference/trans-terms/ [https://perma.cc/RAB5-GFZ4] (last visited Aug. 24, 2023).

entangled; each influences the other.[223] But smashing them together without a second thought "forcibly homogenize[s] human personalities" and "validates hetero-patriarchy" by associating gender with the biological definition of sex.[224] Conflating the two concepts can deny the existence of masculine or androgynous women and feminine or androgynous men.[225]

Data-sharing law also creates interoperability effects. In computer science and engineering, interoperability refers to the capacity of technical systems to interact, connect, and function together.[226] Interoperability can be an anticompetitive barrier to information flow: App Store mobile apps will only run on Apple's operating system, giving the company significant influence over individuals' downstream technology purchases;[227] Facebook made Instagram interoperable with itself but not with Twitter.[228] But from the government's perspective, interoperability is a key driver in law enforcement data sharing.[229] When disparate technologies in a federal system are integrated, authorities have more data to use, more surveillance capacity, and seamless, efficient access to information. Indeed, interoperability in law enforcement intelligence data systems is actually federal law.[230]

---

223. Kristen W. Springer, Jeanne Mager Stellman & Rebecca M. Jordan-Young, Beyond a Catalogue of Differences: A Theoretical Frame and Good Practice Guidelines for Researching Sex/Gender in Human Health, 74 Soc. Sci. & Med. 1817, 1818–19 (2012).

224. Francisco Valdes, Queers, Sissies, Dykes, and Tomboys: Deconstructing the Conflation of "Sex," "Gender," and "Sexual Orientation" in Euro-American Law and Society, 83 Calif. L. Rev. 1, 7, 8 (1995).

225. See Dylan Vade, Expanding Gender and Expanding the Law: Toward a Social and Legal Conceptualization of Gender that Is More Inclusive of Transgender People, 11 Mich. J. Gender & L. 253, 265 (2005).

226. See, e.g., John Palfrey & Urs Gasser, Interop: The Promise and Perils of Highly Interconnected Systems 1–18 (2012) (defining "interoperability" as a "normative theory identifying" the "optimal level of interconnectedness").

227. Jonathan Todd, Real Reasons Behind Apple's Strong Opposition to Interoperability Confirmed, Interoperability News (Apr. 16, 2021), https://interoperability.news/2021/04/real-reasons-behind-apples-strong-opposition-to-interoperability-confirmed [https://perma.cc/757V-LUFJ] (explaining that Apple's "opposition to interoperability" stemmed from the company's desire to "keep users of Apple's services locked in to its own 'walled garden' of iOS devices").

228. Leena Rao, Instagram Photos Will No Longer Appear in Twitter Streams at All, TechCrunch (Dec. 9, 2012), https://techcrunch.com/2012/12/09/it-appears-that-instagram-photos-arent-showing-up-in-twitter-streams-at-all [https://perma.cc/7P4S-3LUQ] (explaining that Facebook made Instagram inoperable with Twitter to "drive more traffic to the web experience for Instagram").

229. See DOJ & DHS, Fusion Center Guidelines: Developing and Sharing Information in a New Era 37–38, 65 (2023), https://bja.ojp.gov/sites/g/files/xyckuh186/files/media/document/fusion_center_guidelines0.pdf [https://perma.cc/2FX5-32TC] (lamenting the lack of interoperability across law enforcement capabilities and signaling the role of fusion centers in creating interoperability).

230. See 8 U.S.C. § 1722(a)(2) (2018) ("[T]he President shall develop and implement an interoperable electronic data system to provide current and immediate access to

But because the benefits of interoperability hinge on system integration, any state wishing to participate in data-sharing systems must conform its data-collection practices to the designs of interagency databases. For instance, if they want to participate in the National Driver Register (NDR) Problem Driver Pointer System (PDPS), a database of information about those whose driving privileges have been revoked, suspended, or canceled,[231] states can collect and share only binary sex information from DMV records because the PDPS is designed with only "male" and "female" options for sex.[232] Therefore, regardless of how state agencies might decide to collect gender data within a vague statutory mandate, data-sharing agreements force those agencies to follow the designed-in limits of the databases and technological systems that use gender data. What is more, decades-old systems are difficult to change. Inclusivity at the data-sharing stage would require not only more nuanced agreements that might dictate inclusive data collection but also wholesale refactoring of the underlying databases to accept that inclusive data. That is a tall order.

IV. LAW AND THE USE OF BINARY GENDER DATA

Having collected and pooled gender data, street-level bureaucrats in state agencies then exercise their discretion to use those data. Indeed, sex and gender have long but checkered histories as classification tools.[233]

---

information in databases of Federal law enforcement agencies and the intelligence community . . . .").

231. See The National Driver Register (NDR) and Problem Driver Pointer System (PDPS), Nat'l Highway Traffic Safety Admin., DOT, https://www.nhtsa.gov/research-data/national-driver-register-ndr [https://perma.cc/U8ZT-L9WT] (last visited Aug. 24, 2023).

232. See, e.g., Nat'l Highway Traffic Safety Admin., DOT, National Driver Register Frequently Asked Questions 1 (2020), https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/national_driver_register_faq_081920_v2_tag.pdf [https://perma.cc/QFD8-G9X3] ("The records submitted to the NDR consist of the following identifying information: name, date of birth, sex, driver license number, and reporting State."); S.C. Dep't of Motor Vehicles, DL-107A, Request for National Driver Register Information on a Current or Prospective Employee (Oct. 2020), https://www.scdmvonline.com/-/media/Forms/DL-107A.ashx [https://perma.cc/9KJG-UWFW] (including "Sex: [Blank]"); Dep't of Motor Vehicles, State of Vt. Agency of Transp., Request for National Driver Register File Check on Current or Prospective Employee, https://dmv.vermont.gov/sites/dmv/files/documents/VN-191-National_Driver_Register_File_Check.pdf [https://perma.cc/7KGM-YKXV] (last visited Aug. 24, 2023) (same). But see Va. Dep't of Motor Vehicles, DL-56, National Driver Register File Check, Individual Request (July 1, 2020), https://www.dmv.virginia.gov/webdoc/pdf/dl56.pdf [https://perma.cc/7EPR-B28G] (including "Sex: Male, Female, Non-Binary").

233. Courts have a history of using gender (and race) data to calculate injured persons' future lost earning capacities. Martha Chamallas, Civil Rights in Ordinary Tort Cases: Race, Gender, and the Calculation of Economic Loss, 38 Loy. L.A. L. Rev. 1435, 1438–39 (2005). Many areas of family law still expect spouses to conform to social expectations associated

Even automated processing of gender data by the state is not new.[234] But AI-driven automation makes things qualitatively different today.[235]

This Part tells the legal story behind how and why automated technologies in the administrative state tend to rely on and reify the gender binary. With the growth of what Professor Aziz Huq called the "allocative state," state agencies that have to distribute benefits are incentivized to use AI to determine eligibility, detect fraud, and calculate entitlements.[236] Enforcement obligations and backlogs have pushed agencies to use AI to predict violations of the law.[237] These developments in law coincide with trends in the political economy of the state: Statutorily imposed austerity, budgetary constraints, and the significant increase in state data collection and sharing have pressured state and local governments to automate.[238]

But the law does more than restrict budgets and get out of the way of innovation.[239] The reality is that the law actively binarizes gender data at

---

with their sex assigned at birth. See Clare Huntington, Staging the Family, 88 N.Y.U. L. Rev. 589, 628–29 (2013). States use gender data to separate people in homeless shelters, drug treatment facilities, foster homes, domestic violence shelters, and prisons. See Spade, Documenting Gender, supra note 36, at 735–36, 752–53; see also Lisa Mottet & John M. Ohle, Transitioning Our Shelters: A Guide to Making Homeless Shelters Safe for Transgender People 1–6 (2003).

234. During what technology historian Mar Hicks calls the "prehistory of algorithmic bias," room-sized computing systems allocated welfare-state resources along gender lines. Hicks, supra note 73, at 27–30.

235. David Freeman Engstrom, Daniel E. Ho, Catherine M. Sharkey & Mariano-Florentino Cuéllar, Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies 9 (2020) (noting the importance of AI in making governance more effective); Kroll et al., supra note 59, at 636 ("[T]he accountability mechanisms and legal standards that govern decision processes have not kept pace with technology.").

236. Aziz Z. Huq, Constitutional Rights in the Machine-Learning State, 105 Cornell L. Rev. 1875, 1894–99 (2020); see also, e.g., Cahoo v. SAS Analytics Inc., 912 F.3d 887, 892, 895 (6th Cir. 2019) (challenging the erroneous termination of unemployment benefits by AI); K.W. v. Armstrong, 180 F. Supp. 3d 703, 708 (D. Idaho 2016) (challenging state use of an algorithm to determine in-home care benefits); Ark. Dep't of Hum. Servs. v. Ledgerwood, 530 S.W.3d 336, 339 (Ark. 2017) (challenging an algorithm used to assess disability care).

237. See Engstrom et al., supra note 235, at 22 (describing AI tools used by the SEC to identify potential securities law violations).

238. Citron, Technological Due Process, supra note 17, at 1259 (referring to budget shortfalls as motivating the government to automate).

239. See, e.g., Anupam Chander, How Law Made Silicon Valley, 63 Emory L.J. 639, 647–69 (2014) (arguing that immunity from liability, copyright safe harbors, and weak privacy law allowed technology companies to thrive in the United States); Mihailis E. Diamantis, The Extended Corporate Mind: When Corporations Use AI to Break the Law, 98 N.C. L. Rev. 893, 899–900 (2020) (noting that the "lack[] [of] a theory of liability" and the "legal loophole left by respondeat superior" allow corporations to use AI to violate the law); Katyal & Jung, supra note 10, at 760–63 (arguing that automated surveillance tools that discriminate arose in a void left by privacy law). Nor is it clear that deregulation spurs innovation or that regulation stifles it. See, e.g., Yafit Lev-Aretz & Katherine J. Strandburg, Privacy Regulation and Innovation Policy, 22 Yale J.L. & Tech. 256, 275–76 (2020).

use by directly mandating and indirectly incentivizing agencies to automate their administrative functions to improve efficiency and to rely on more and more data as the basis for effective governance.

A. *Mandating Automation: The Law on the Books*

For decades, states have explicitly required agencies to automate their work to increase efficiency. In 1979, Virginia established an automation fund to "fully automate[]" the entire system of vital statistics.[240] California required all counties and its department of health to automate the process "that accepts and screens applications for benefits under the Medi-Cal program" to streamline identity verification and eligibility determinations.[241] The state also made new county grant-reporting requirements contingent on implementing the "necessary automation to implement" the law efficiently[242] and required the Student Aid Commission to "develop an automated system to verify a student's status as a foster youth to aid in the processing of applications for federal financial aid."[243] The Colorado Public Assistance Act incentivized counties to use the state's automated case management and child support systems rather than spending additional funds on their own.[244] Arizona and West Virginia, among many other states, require their agencies in charge of enforcing child support orders to use "automated administrative enforcement" to respond to requests "promptly."[245] California law also tasks the director of Child Support Services with "implementing and managing all aspects of a single statewide automated child support system" that carries out state child support obligations promptly and efficiently.[246]

If these and countless other statutes mandate the automation of specific state functions, general declarations of the efficiency benefits of automation have established automation as official state policy. When enacting campaign disclosure laws, the Kentucky General Assembly found that "computer automation is a necessary and effective means" of processing "vast amounts of data."[247] California has declared that statewide-automated systems are "essential."[248] The federal government

240. Va. Code Ann. § 32.1-273.1 (2023).

241. Cal. Welf. & Inst. Code § 14011.9(a) (2023).

242. Id. § 11265.1(c)(3)(B)(ii).

243. Cal. Educ. Code § 69516 (2023).

244. Colo. Rev. Stat. § 26-2-108(b)(II)(A)–(B) (2023).

245. Ariz. Rev. Stat. Ann. § 25-525(A)–(B) (2023); W. Va. Code Ann. § 48-14-602 (LexisNexis 2023).

246. Cal. Fam. Code § 17308 (2023).

247. Ky. Rev. Stat. Ann. § 121.005(1)(c) (West 2023).

248. Cal. Child Support Automated Sys. Act, Cal. Welf. & Inst. Code § 10080(a)(2) (1999) (repealed 2017).

has also connected automation with increased efficiency in several administrative spaces, including family support.[249]

Many states have also created chief data, information, or innovation offices (CIOs) with the explicit goal of automating state decisionmaking systems to increase efficiency.[250] Vermont created an Agency of Digital Services to provide technological solutions to all parts of state government and avoid costs or save money "as a result of technology optimization."[251] Ohio recently created an Office of Human Services Innovation in its Department of Jobs and Family Services, in part to make statewide policy recommendations for "[s]tandardizing and automating eligibility determination policies and processes for public assistance programs."[252] When creating its CIO position, Puerto Rico stated that the systems the CIO would create "must contribute to a more efficient use" of government resources.[253] In Utah, the state's CIO will approve new funding for automation only if it "will result in greater efficiency in a government process."[254] This is a pattern. Nearly 200 state laws associate automation, CIO missions, and efficiency.[255]

In addition to formalizing automation as a government goal, laws on the books also establish efficiency as government policy, guiding the terms on which agencies use automated tools. At the federal level, the Office of Management and Budget (OMB) and one of its subdivisions, the Office of Information and Regulatory Affairs (OIRA), use technical review and approval processes to implement efficiency mandates like budget controls and narrow versions of cost–benefit analyses over a host of agency actions.[256] As Professor Julie Cohen has demonstrated, OMB/OIRA

---

249. See Computerized Support Enforcement Systems, 63 Fed. Reg. 44795, 44795 (Aug. 21, 1998) (to be codified at 45 C.F.R. pts. 302, 304, 307) ("Full and complete automation is pivotal to improving the performance of the nation's child support program.").

250. See, e.g., Haw. Rev. Stat. Ann. § 27-44 (West 2023) ("The chief data officer shall use the state information assets and analytics to research and recommend processes and tools to improve inter-departmental and intra-departmental decision making and reporting."); Or. Rev. Stat. § 276A.353 (West 2023) ("The Chief Data Officer shall . . . [i]dentify ways to use and share existing data for business intelligence and predictive analytic opportunities.").

251. Vt. Stat. Ann. tit. 3, § 3303 (2023); see also id. at §§ 3301–3305 ("The Agency of Digital Services is created to provide information technology services and solutions in State government.").

252. Ohio Rev. Code Ann. § 5101.061(B)(3) (2023).

253. P.R. Laws Ann. tit. 3, § 9866(f) (2023).

254. Utah Code § 63A-16-903(2)(a)(ii) (2023).

255. Based on a Westlaw advanced search that resulted in 203 hits. State Statute Search Results, Westlaw Precision, https://1.next.westlaw.com/ (select content type "Statutes & Court Rules"; select Advanced Search; select "All States" for jurisdiction; use query: chief +4 data information innovation +4 officer; refine by: efficien! OR reduc! lower cut +4 cost!) (on file with the *Columbia Law Review*) (last visited Sept. 12, 2023).

256. See Cohen, Between Truth and Power, supra note 16, at 194; Eloise Pasachoff, The President's Budget as a Source of Agency Policy Control, 125 Yale L.J. 2182, 2213–23 (2016).

involvement prioritizes efficiency over other values.[257] In particular, OMB/OIRA's integration into the administrative state brings accountants and other professionals focused on "efficient management" to the forefront of agency decisionmaking even when those agencies' missions center public health, equity, or welfare.[258] Those professionals use the logics of accounting and management to make normative decisions about a program's value seem like detached, neutral appraisals of dollars and cents.[259]

This creates a fertile ground for automation. Efficiency mandates to do necessary government work with less funding decouple agency missions from experts trained in the agency's goals and shift power to number crunchers focused on one thing—efficiency—that takes primacy over other agency goals.[260] And automated technologies are universally touted as enhancing administrative efficiency.[261] More specifically, cost–benefit appraisal methods are inherently utilitarian and, therefore, assume that even serious harm, especially to a small minority of the population, could be outweighed by higher levels of economic benefits for others. As a result, cost–benefit analysis implements efficiency mandates in ways that make realizing those benefits through automation more likely.[262]

B. *Efficiency and the Gender Binary*

What do efficiency mandates have to do with binary gender? In addition to falling prey to the same problems as the law of gender data collection and sharing, gender data law privileges the gender binary because it creates a certain type of regulatory automation—namely, one guided by values of efficiency and risk management. This system erases transgender and gender-nonconforming individuals in three ways: The resulting technologies model probabilities that exclude minorities, reflect managerial interests that ignore inclusion, and incorporate coding language that binarizes data inputs.

As we have seen, the law of gender data use mandates and incentivizes automation primarily to verify identity, prevent fraud, and achieve security. In that way, the law envisions automation as a form of governmentality

257. Cohen, Between Truth and Power, supra note 16, at 195.

258. Id. at 194.

259. Id.

260. Id. at 194–95.

261. See, e.g., Citron, Technological Due Process, supra note 17, at 1259.

262. Many state laws explicitly link automation with efficiency mandates. For instance, Texas implemented an automated system to make healthcare eligibility determinations only after a cost–benefit analysis focused almost exclusively on cost savings from automation. Tex. Gov't Code Ann. § 531.191(d) (West 2023). Mississippi's automated child welfare unit can only operate in the most "cost efficient manner" based on a cost–benefit analysis. Miss. Code Ann. § 43-19-31(k) (2023).

aimed at risk management.[263] Algorithmic technologies like the ones experienced by Sasha and Toby are forms of "targeted governance" in which the logics of information, surveillance, and prediction are carried out through data-driven assessment of systemic threats.[264] But assessing risk requires modeling threats,[265] and statistical modeling "depend[s] on assumptions about variables and parameters that are open to contestation."[266] This kind of quantification has been shown to accelerate predictable injustice.[267]

But the problem runs deeper. Modeling for risk requires technologies to rely on probabilities; even systemic threats are potential future harms that may or may not occur.[268] So when technological systems are assessing whether Sasha is a terror threat or Toby is a fraud threat, they are using gender data in a complex probabilistic equation. Policy by probabilities is ostensibly efficient: It captures the realities of most people most of the time. As applied to any given individual, however, what that probability predicts could be off the mark or incorrect. Because transgender and nonbinary individuals make up less than 0.8% of the U.S. population and usually far less in surveys,[269] statistical models designed for efficiency are likely to fail when applied to them, excluding them as "noise."[270] Gender-diverse populations are certainly not the only marginalized groups victimized by technical tools that are trained on data about the general population norm; queer people of color and those at the intersection of

263. Cohen, Between Truth and Power, supra note 16, at 140–57; Currah & Mulqueen, supra note 2, at 576.

264. Mariana Valverde & Michael Mopas, Insecurity and the Dream of Targeted Governance, *in* Global Governmentality: Governing International Spaces 233, 239 (Wendy Larner & William Walters eds., 2004).

265. Calo, Modeling, supra note 90, at 1395.

266. Cohen, Between Truth and Power, supra note 16, at 182.

267. Frank Ackerman & Lisa Heinzerling, Pricing the Priceless: Cost–Benefit Analysis of Environmental Protection, 150 U. Pa. L. Rev. 1553, 1578–79 (2002).

268. Cohen, Between Truth and Power, supra note 16, at 183; see also Calo, Modeling, supra note 90, at 1398–405; Scheuerman et al., supra note 56, at 144:6.

269. See Jody L. Herman, Andrew R. Flores & Kathryn K. O'Neill, How Many Adults and Youth Identify as Transgender in the United States? 1 (2022), https://williamsinstitute.law.ucla.edu/wp-content/uploads/Trans-Pop-Update-Jun-2022.pdf [https://perma.cc/7H4Y-SMTD] (finding that 1.6 million youth and adults in America identify as transgender); Bianca D.M. Wilson & Ilan H. Meyer, Nonbinary LGBTQ Adults in the United States 2 (2021), https://williamsinstitute.law.ucla.edu/wp-content/uploads/Nonbinary-LGBTQ-Adults-Jun-2021.pdf [https://perma.cc/25KS-BDBX] (estimating that 1.2 million adults in America identify as nonbinary). Because the 1.2 million estimate of nonbinary American adults includes transgender nonbinary individuals, and approximately 40% of nonbinary adults identify as transgender, see Wilson & Meyer, supra, at 2–3 & fig.1, the total number of transgender and nonbinary individuals in the United States is likely far less than 2.8 million.

270. Beauchamp, supra note 10, at 2.

several matrices of domination fare worse.[271] But as Os Keyes, a scholar of human-centered design and engineering, has argued, when "an error rate . . . disproportionately falls on one population[,] [it] is not just an error rate: it is discrimination."[272]

Sex and gender data use in the automated state is also decidedly managerial. Managerialism is an ideology and set of practices closely associated with neoliberal governmentality in which values like efficiency, innovation, and data-driven policy take primacy over social values.[273] Efficiency is by no means a bad thing, but a managerial approach to governance relies on narrow, financialized conceptions of costs and benefits to determine efficiencies.[274] That leaves little room for social welfare and gender inclusivity.

For instance, even though scholars talk about interagency MOUs and data-sharing agreements as if they are between governments or government departments, they are really agreements between those departments' *managers*.[275] As noted above, the law of sex and gender data sharing is often not the product of statutory permission but civil servant discretion. Therefore, interagency agreements reflect the goals and orientations of departmental managers or what their departments need to fulfill the jobs of governance. Those goals can undoubtedly overlap with other values, like equity and antisubordination, democracy, or the general welfare. But the extent to which those values are realized through agency action depends on whether they align with managers' goals.[276] And if keeping costs down is state law, efficiency will take center stage in those goals.

The managerial automated state is one that judges its automation on cases closed and dollars saved.[277] Those metrics are designed to elide even significant harm to small populations.[278] That means consigning transgender and gender-nonconforming individuals to repeated moments of everyday vulnerability even as the automated tools responsible for that

271. See Buolamwini & Gebru, supra note 12, at 10 (concluding that, based on an "intersectional demographic and phenotypic analysis, . . . all algorithms perform worse on female and darker subjects when compared to their counterpart male and lighter subjects").

272. Keyes, supra note 79, at 88:13.

273. Cohen, Between Truth and Power, supra note 16, at 171–72.

274. Id.

275. Willard F. Enteman, Managerialism: The Emergence of a New Ideology 154 (1993) (identifying managers of organizations and negotiations among managers as the key instruments of authority in managerialist societies).

276. Id. at 184.

277. Cohen, Between Truth and Power, supra note 16, at 194.

278. Id. at 190–91, 195.

vulnerability are legitimized as effective, "intelligent," and efficient risk-management policymaking.[279]

A third way that the efficiency-focused law of gender data use entrenches the male/female binary centers on database design, coding, and function. If the state wants to put its sex and gender data into databases so the data can be used by data-matching and data-mining systems in the most efficient way possible, coders will choose "Boolean variables" to describe gender instead of a box for an open-ended answer.[280] A Boolean variable is a binary variable with only two options: 0 and 1. As critical information studies scholar Meredith Broussard notes, if the state designs code "for maximum speed and efficiency using a minimum of memory space, you try to give users as few opportunities as possible to screw up the program with bad data entry. A Boolean for gender, rather than a free text entry field, gives you an incremental gain in efficiency."[281] Coding for gender as a Boolean or binary variable is also deeply ingrained in computer science and programming education[282] as well as governments' long history of digitization and automation.[283] At the same time, the practice excludes those who do not identify as either male or female.

C. *Guiding Automation: The Law on the Ground*

While the laws on the books mandate or foster automation to realize efficiency benefits, the law on the ground—including public-sector procurement and the applications of trade secrecy and procedural privacy law in practice—further facilitates the kind of automation that tends to

---

279. Valverde & Mopas, supra note 264, at 239. The problem of regulatory managerialism also explains the insufficiency of the procedural due process proposals in the algorithmic accountability literature. These proposals include audit trails, impact assessments, and humans in the loop of automated decisionmaking systems. See, e.g., Reisman et al., supra note 17, at 3–6 (recommending impact assessments); Citron, Technological Due Process, supra note 17, at 1258, 1305 (fairness standards and audit trails); Froomkin et al., supra note 17, at 38 (requiring humans in the loop); Jones, supra note 17, at 217 (audit trails and requiring humans in the loop); Kaminski, supra note 17, at 1535 (audit trails); Selbst, supra note 24, at 123–25 (impact assessments). Imbued with management values and implemented by compliance professionals, these tools are easily subject to capture. Ari Ezra Waldman, Privacy Law's False Promise, 97 Wash. U. L. Rev. 773, 776 (2020) [hereinafter Waldman, False Promise] (noting that compliance professionals define privacy law's implementation, leading to compliance measures promoting efficiency and risk management rather than the law's stated goals).

280. Meredith Broussard, When Binary Code Won't Accommodate Nonbinary People, Slate (Oct. 23, 2019), https://slate.com/technology/2019/10/gender-binary-nonbinary-code-databases-values.html [https://perma.cc/LB4Q-8KF5].

281. Id.

282. See Natalie Kiesler & Benedikt Pfülb, The Boolean Dilemma: Representing Gender as Data Type, 21 Proc. Koli Calling Int'l Conf. on Computing Educ. Rsch., no. 30, Nov. 2021, at 1, 1.

283. See Hicks, supra note 73, at 29.

flatten gender data into binary male/female options. Procurement, as Professors Deirdre Mulligan and Kenneth Bamberger argue, is both a process and a mindset.[284] As a process, procurement is a pathway through which government agencies send out requests for proposals (RFPs) for new technologies, evaluate them based on a series of defined metrics, and acquire technologies by entering into contracts with for-profit, third-party vendors.[285] It is governed by detailed regulations that promote certain values: low costs, fair bidding, innovation, and healthy competition.[286] As a mindset, procurement positions AI and machine learning as "'the next logical step'" in administrative automation and as "machinery used to support some well-defined function" instead of an exercise in the distribution of power.[287]

Both the process and mindset of technology procurement make it more likely that the technology purchased by the state will embed the gender binary. They do this by immunizing algorithmic technologies from the interrogation necessary to disrupt the status quo—which almost always relies on the gender binary—in three related ways.

First, the process and mindset conceptualize AI and algorithmic technologies as neutral processes that simply help fulfill agencies' missions.[288] In theory, that is why procurement can be done through the neutral language and process of RFPs rather than the political language and process of policy.[289] RFPs are not supposed to make policy; they solicit bids for technologies to implement policy.[290] Under this logic, the technology does what the agency has always done, only more quickly, more cheaply, and supposedly with fewer mistakes. This was precisely the position of the Department of Homeland Security when federal law authorized the creation of new "fusion centers" that pooled national

284. Deirdre K. Mulligan & Kenneth A. Bamberger, Procurement as Policy: Administrative Process for Machine Learning, 34 Berkeley Tech. L.J. 773, 779–80 (2019) (noting that the process of procurement embodies certain bureaucratic values that collectively define a mindset that fails to account for other public values).

285. Id.

286. Id. at 779–80 (citing Steven L. Schooner, Desiderata: Objectives for a System of Government Contract Law, 11 Pub. Procurement L. Rev. 103 (2002)).

287. Id. at 779 (quoting HHS, Solicitation No. 19-233-SOL-00098_BASE 9 (2019), https://sam.gov/api/prod/opps/v3/opportunities/resources/files/39d0a0ce8bfe09391b9fee07833274de/download?&status=archived&token= [https://perma.cc/89H6-YKVS]).

288. Id. at 789.

289. Traditional agency policymaking, at least at the federal level, is governed by the Administrative Procedure Act, which provides two pathways for agency policymaking: rulemaking, which includes a public notice and comment period during which members of the public can provide feedback, and adjudication, in which the agency applies its rules to the entities it regulates. See 5 U.S.C. §§ 551–559 (2018).

290. See Mulligan & Bamberger, supra note 284, at 779–80.

security data.[291] The Department's privacy impact assessment (PIA) stated that fusion centers, which used advanced technology to collect, share, and process large amounts of data related to law enforcement, national security, and terrorism, were simply replicating "many of the interactions the Department was already undertaking."[292] And if technology simply does what an agency has always done, then there is no need to evaluate its underlying assumptions, normative choices, and design. This means that any existing state practice that uses binary sex and gender data will simply be integrated and encoded into a new system without interrogation.

Second, the procurement process and mindset situate agency expertise as dependent on and subordinate to technological expertise, privileging the latter over the former. If agency staff have few technical skills and conceptualize their role as simply using a complex tool that a private-sector expert built, they often assume they are incapable of interrogating the technology even if they wanted to. This presumed ignorance has taken center stage in litigation. In *State v. Loomis*, a due process challenge to Wisconsin's use of an algorithm that took gender into account when determining likelihood of recidivism,[293] no one from the state (even the judges deciding the case) knew how the algorithm worked.[294] The same thing happened in *Estate of Jacobs v. Gillespie*, a challenge to Arkansas's use of an automated system to determine disability benefits.[295] No one from the state saw it as their responsibility to understand how a critical system actually functioned.[296] Without public willingness or desire to interrogate the normative, political, and distributive choices made by algorithmic design, private-sector engineers and managers make those choices. The values and norms of their sociotechnical environment get embedded into automated decisionmaking systems.[297] Therefore, even if an engineer could capture legally relevant variables in design, the technology might still not capture the law's normative goals.[298] It will, instead, reflect the engineers and their

---

291. Fahey, Data Federalism, supra note 16, at 1024–26 (explaining how fusion centers facilitate information exchange between government law enforcement agencies by collocating government personnel and sharing access to information in each other's possession).

292. DHS, Privacy Impact Assessment for the Department of Homeland Security State, Local, and Regional Fusion Center Initiative 4 (2008), https://www.dhs.gov/xlibrary/assets/privacy/privacy_pia_ia_slrfci.pdf [https://perma.cc/B4W7-EPJY].

293. 881 N.W.2d 749, 753 (Wis. 2016).

294. Mulligan & Bamberger, supra note 284, at 777.

295. First Amended Complaint at 16–17, Est. of Jacobs v. Gillespie, No. 3:16-cv-00119-DPM (E.D. Ark. Nov. 1, 2016).

296. Calo & Citron, supra note 23, at 799 (describing that agency officials "did not know how the system worked").

297. See Bear & Knobe, Normality, supra note 217, at 25.

298. See Noëmi Manders-Huits, What Values in Design? The Challenge of Incorporating Moral Values Into Design, 17 Sci. & Eng'g Ethics 271, 279 (2011) (arguing

managers' traditional goals: efficiency, technical function, and profit.[299] Inclusive and respectful gender data is not one of those goals.

Third, the procurement process and mindset defer to private companies' demands for maximalist intellectual property and trade secrecy protections. To obtain technologies they find both necessary and complex, governments often use procurement contracts that protect the trade secrets of their vendors. For instance, the Alaska Procurement Policies and Procedures Manual requires agencies to treat as confidential anything designated as a trade secret by a third-party vendor in a procurement contract.[300] The Freedom of Information Act and its state equivalents exempt trade secrets, allowing vendors to provide necessary information in response to RFPs without fear of any of it being released to the public.[301] And, as the law and technology scholar Rebecca Wexler has shown, vendors have routinely used trade secrecy claims to protect their sentencing, recidivism, and parole algorithms from being interrogated in court.[302] At present, at least twenty-one states have codified trade secrecy privileges in their evidence rules, further insulating automated technologies from public interrogation.[303] By privileging private technology over the public interest, the procurement process and mindset shield automated technologies from the kind of deep public review that could uncover transgender and nonbinary erasure.

D. *Immunizing Automation: Information Law in Action*

Alongside the procurement process and mindset, agencies and the technology companies that build algorithmic decisionmaking systems leverage information law to foster automation that binarizes gender. Specifically, both the state and technology vendors weaponize privacy impact assessments (PIAs) to prevent anyone from interrogating how algorithmic technologies use gender while prioritizing efficiency and the utilitarianism of cost–benefit analysis.

At the federal level, the E-Government Act of 2002 requires agencies to conduct PIAs for any electronic information system or program that

---

that integrating empirical methods in Value-Sensitive Design is challenging because the values are often abstract and difficult to interpret); Frank Pasquale, Professional Judgment in an Era of Artificial Intelligence and Machine Learning, boundary 2, Feb. 2019, at 73, 74 (arguing that "substituting AI for education and health-care professionals" requires a "corrosive reductionism").

299. Paul Ohm & Jonathan Frankle, Desirable Inefficiency, 70 Fla. L. Rev. 777, 778–79 (2018).

300. Alaska Administrative Manual 81: Procurement 81.195 (2018), http://doa.alaska.gov/dof/manuals/aam/resource/81.pdf [https://perma.cc/R7FY-KADK].

301. Citron, Technological Due Process, supra note 17, at 1293.

302. Rebecca Wexler, Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System, 70 Stan. L. Rev. 1343, 1353–54 (2018).

303. Id. at 1352.

collects information about citizens.[304] Several state laws also require agencies to develop rules for conducting or completing PIAs for any use of technology involving citizen data.[305] PIAs are supposed to describe the information to be collected, its purpose and use, how the information will be secured, when individuals will have opportunities to deny or grant consent, and to what extent the technological system will impact individual privacy.[306] Their goal is to legitimize the use of data-driven technologies by passing them through a form of informal due process, checking them against values like security and privacy.[307] But in reality, both in their design and their application, PIAs do not consider transgender and nonbinary erasure.

Consider, for example, the PIA used by the executive branch of West Virginia.[308] In a "threshold analysis," agencies designate whether the technology being reviewed is major, minor, a support system, or something else.[309] They then have to acknowledge if personally identifiable information (PII) is involved in the system. Gender is included in the list of PII, but there is no opportunity to describe how the technology collects or uses gender data or if those uses are in any way problematic.[310] West Virginia's Data Classification Policy considers gender data "sensitive" but not "restricted,"[311] which means that no additional work or special

304. E-Government Act of 2002, Pub. L. No. 107-347, § 208(b), 116 Stat. 2899, 2922.

305. See, e.g., W. Va. Code Ann. §§ 5A-6B-2, -3(b)(10) (LexisNexis 2023) (defining a PIA as "a procedure or tool for identifying and assessing privacy risks throughout the development life cycle of a program or system"); see also Ga. Code Ann. § 20-2-663(a)(4) (2023); Ohio Rev. Code Ann. § 1347.15(b)(8) (2023). PIAs are also now part of what I have previously called the "second wave" of privacy laws that apply to for-profit, private companies. Ari Ezra Waldman, The New Privacy Law, 55 U.C. Davis L. Rev. Online 19, 21 (2021), https://lawreview.law.ucdavis.edu/online/55/files/55-online-Waldman.pdf [https://perma.cc/TR3C-EFQA]. PIAs are required by the General Data Protection Regulation (GDPR) in the European Union, see Council Regulation 2016/679, art. 5, 2016 O.J. (L 119) 1, 35–36, and the proposed American Data Protection and Privacy Act, see American Data Privacy and Protection Act, H.R. 8152, 117th Cong., § 301(d) (2022).

306. Memorandum from Joshua B. Bolten, Dir., OMB, to Heads of Exec. Dep'ts and Agencies, M-03-18, Implementation Guidance for the E-Government Act of 2002 (Sept. 26, 2003), https://obamawhitehouse.archives.gov/omb/memoranda_m03-22/ [https://perma.cc/7GQZ-7BU9].

307. Selbst, supra note 24, at 123–35 (arguing that for algorithmic impact assessments to be successful, they must take into account the way regulation is filtered through institutional logic).

308. W. Va. Exec. Branch, Privacy Impact Assessment (PIA) Instructions, https://privacy.wv.gov/privacyimpactassessment/Documents/Privacy%20Impact%20Assessment%20v.060523.pdf [https://perma.cc/Y2YF-Z79Z] [hereinafter W. Va. PIA Instructions] (last modified Aug. 25, 2022).

309. Id. at 5.

310. Id. at 5–6.

311. State of W. Va. Off. of Tech., Policy: Data Classification 2–3 (Jan. 6, 2010), https://drive.google.com/file/d/1NNqhRmfaK-SEa0PBuurIrIQGGvMYvvJc/view [https://perma.cc/NX59-JLWF] [hereinafter W. Va., Data Classification Policy] (last

restrictions are necessary to protect it.[312] For instance, if the technology uses only "sensitive" data, the vendor can have free access to those data and store them in jurisdictions with weak privacy laws.[313] The PIA then asks if there is statutory authorization to collect and use citizen data, how it will be used, where the information will be stored, and whether the data can be shared electronically or on paper.[314] Finally, it accounts for controls, asking: "Are there controls in place to ensure that access to PII is restricted to only those individuals who need the PII to perform their official duties?"[315] There are three answer options: "yes," "no," and "NA." "Are there physical controls in place to ensure the files are backed up?"[316] Again, "yes," "no," and "NA" are the only possible—and only required—answers.[317] The PIA concludes by asking whether the agency has an incident response plan and requesting a simple dropdown yes/no answer for whether "additional risk mitigation [is] needed."[318]

The TRACS PIA completed by HUD's Office of Housing follows the same pattern. It notes that the genders of those receiving federal housing assistance will be collected and processed, but there is no space in the PIA design to consider the impacts on diverse gender identities.[319] With PII in the system, the PIA asks for "security control" and provides a check box to indicate that such controls exist.[320] It asks for remote work policies and rules about downloading information, which the Office of Housing answered by listing rules from the Department's handbook.[321] The PIA concludes with questions about security protocols.[322]

This is how PIAs function in the information industry as well. Reduced to checkbox compliance and simple questions, PIAs tend to focus on procedure and security.[323] The capacity of PIAs to have any substantive impact on underlying technologies is also a matter of PIA design. That is, if PIAs do not ask about the scope of gender data, whether the data include transgender, nonbinary, and gender-nonconforming individuals, or how the technology might cause gender erasure, those questions will not be considered. PIAs interrogate only those aspects of technology captured by

---

updated Oct. 21, 2021) (classifying datasets including gender, such as driver history records and personnel records, as sensitive).

312. Id. at 3.
313. Id.
314. W. Va. PIA Instructions, supra note 308, at 8–9.
315. Id. at 10.
316. Id. at 11.
317. Id.
318. Id. at 11.
319. TRACS PIA, supra note 194, at 8.
320. Id. at 9.
321. Id. at 11.
322. Id. at 15–16.
323. Waldman, Industry Unbound, supra note 24, at 132–33.

their questions; civil servants can answer only with the options they are provided.

Asking more probing questions on PIAs will not solve the problem. PIAs are necessarily cursory. They are often reduced to simple charts with "yes" or "no" answer options so they can be completed by nonexperts.[324] As a result, they become tools for legitimizing otherwise data-extractive technologies without any deep interrogation of their impact on even those facets of technology design covered by the PIA.[325] For government agencies that have already decided they want to purchase a particular automated technology, PIAs like the ones used by HUD or West Virginia become window-dressing procedures, a form of performative compliance, that offer the gloss and patina of accountability without any of the work. They are, in short, formalities. And yet, they retain power backed by the formal law; a PIA is a necessary precondition of using new automated systems. Just like their corporate counterparts, state providers of PIAs legitimize quests for automation.

V. LESSONS FOR THE AUTOMATED STATE

Law plays a critical role in creating an automated state that prioritizes efficiency and, therefore, binarizes sex and gender data. This conclusion reinforces the notion, now well established in the law and political economy literature, that economic and distributional systems are creatures of law.[326] In addition to buttressing some of what we already know about the law, this Article's case study of sex and gender data offers several additional insights into the automated administrative state in general, insights that challenge and add nuance to the conventional wisdom about the state's use of algorithmic tools. This Part explores four of those lessons.

First, despite the popular view that automation erodes discretion, this Article demonstrates discretion's persistence. Second, contrary to the conventional account about the primacy of engineering expertise in the automated state, this Article shows how much the state and engineers rely on stereotypes and perceptions of common sense when designing technology and doing their jobs. Third, challenging the view that automation occurs in a regulatory void, this Article shows how automation is a product of neoliberal approaches to law. Finally, contributing to scholarship focusing on technology's subordinating capacities, this Article shows how the law of automation creates a state that is simultaneously awash in gender data but devoid of gender-diverse data, subjecting transgender, nonbinary, and gender-nonconforming individuals to all the

---

324. Id. at 133.

325. Waldman, False Promise, supra note 279, at 785.

326. Regarding the law and political economy literature, see Jedediah Britton-Purdy, David Singh Grewal, Amy Kapczynski & K. Sabeel Rahman, Building a Law-and-Political-Economy Framework: Beyond the Twentieth-Century Synthesis, 129 Yale L.J. 1784, 1791–94 (2020).

harms of the data-driven state without any of the benefits. With these lessons, this Part concludes by returning to privacy law principles of data minimization and antisubordination for a new framework to govern sex and gender data: The state should collect, share, and use only as much gender data as is necessary to contribute to the liberation of gender-diverse populations.

A. *Persistent Discretion*

Many law and technology scholars have argued that automating state apparatuses takes away opportunities for civil servants to exercise discretion, a key rationale for the administrative state in the first place and a critical tool for individualized care for those in need of government assistance.[327] Although discretion in the administrative state looks different today than it once did, the law of sex and gender data collection, sharing, and use demonstrates the continued strength and persistence of street-level bureaucratic discretion in the automated state.

Automated decisionmaking does disrupt some of the traditional functions of street-level bureaucrats. For instance, instead of having a social worker visit disabled residents in person to determine how much in-home care they needed, Arkansas turned to an algorithm (with disastrous results).[328] But frontline worker discretion is critical to data pathways in the automated state. Required by law to collect sex and gender data, civil servants decide how to collect it. And they sometimes change the law while doing so: Whether out of ignorance or intent, frontline workers sometimes decide to ask for gender on voter registration forms even though the law requires sex.[329] In addition, because some state laws merely permit rather than explicitly require interagency data sharing, street-level bureaucrats also decide how, when, with whom, and under what terms to share sex and gender data. Within frameworks constructed by law, civil servants also have significant discretion when procuring new technologies from third-party vendors. And civil servants squeeze and stretch the formal procedural requirement of PIAs to push their procurement decisions over the finish line. There appears to be far more discretion in the automated state than scholars have realized.

Much scholarship elides street-level bureaucrats' persistent and significant discretion in the automated state because it is focused

---

327. See Lipsky, supra note 126, at 10–22; Calo & Citron, supra note 23, at 799; Metzger, supra note 134, at 1900; Mulligan & Bamberger, supra note 284, at 778.

328. Ark. Dep't of Hum. Servs. v. Ledgerwood, 530 S.W.3d 336, 339–40 (Ark. 2017); see also Leslie Newell Peacock, Legal Aid Sues DHS Again Over Algorithm Denial of Benefits to Disabled: Update With DHS Comment, Ark. Times (Jan. 27, 2017), https://arktimes.com/arkansas-blog/2017/01/27/legal-aid-sues-dhs-again-over-algorithm-denial-of-benefits-to-disabled-update-with-dhs-comment [https://perma.cc/U2U7-UHDW].

329. See supra note 139.

elsewhere—namely, on the algorithmic system itself.[330] That focus yields essential insight. Expanding the scope of scholarly attention to the prerequisite stages of automation can yield even more.[331] Algorithms need data, and those data can effectively train algorithmic systems only when aggregated and pooled in large quantities. Sometimes, states purchase data from brokers.[332] Large amounts of sex and gender data are collected through forms and aggregated through interagency agreements and interstate compacts, all of which are drafted and negotiated by street-level bureaucrats. Civil servants even have some discretion to affect the designs of the technologies they buy from private, for-profit companies depending on the nature of the procurement contracts. At the automation stage, civil servants exercise their power and discretion to immunize algorithmic technologies from public interrogation. Automation may muddle our traditional conceptions of agency expertise, but it does so while adding new opportunities for frontline workers to exercise power, discretion, and knowledge.

History shows that the persistence of such discretion poses risks for transgender, nonbinary, and gender-nonconforming individuals. Dean Spade has written extensively about the administrative state's hostility to transgender people.[333] Political scientist Paisley Currah points to state agencies' inconsistent and irrational practices for changing gender designations on official documents as evidence of systemic transphobia in government.[334] And technology historian Mar Hicks has shown how bureaucrats took advantage of newly computerized welfare allocation systems in post–World War II Britain to erase transgender identities: They used their discretion to deny gender designation change requests while programming transgender citizens' files into the computer as "aberrant" instead of simply changing M to F or F to M.[335] This history is reason enough for gender-diverse communities to doubt the promises of an automated state, whether infused with discretion or not.

B. *Persistent Stereotypes*

In addition to showing that discretion persists, this Article's case study of the state's use of sex and gender data complicates the extant narrative about agency expertise in the information age. Scholars argue that

---

330. See supra note 17.

331. See David Lehr & Paul Ohm, Playing With the Data: What Legal Scholars Should Learn About Machine Learning, 51 U.C. Davis L. Rev. 653, 655–58 (2017) (making a similar recommendation, but focusing only on machine learning rather than the law's role in mandating, fostering, and incentivizing data collection, sharing, and use).

332. See supra note 144.

333. See Spade, Normal Life, supra note 19, at 9–11; Spade, Documenting Gender, supra note 36, at 737–39.

334. Currah, supra note 67, at 7–9, 28.

335. Hicks, supra note 73, at 27.

automation shifts expertise in state agencies from frontline workers hired because of their substantive knowledge of agency work to engineers and programmers who design the algorithms that make policy.[336] That is undoubtedly true to an extent, but the reality is more complicated. When it comes to the collection, sharing, and use of sex and gender data, expertise takes a back seat to stereotypes and perceptions of common sense.

Popular understandings of sex and gender affect data pathways from the beginning. Statutes, sharing agreements, and procurement contracts capturing sex and gender data are often imprecise; they refer only to "sex" or "gender" without specifying how that information should be collected or used. This could be explained by the limits of language, the need to build majorities and coalitions when passing laws, or the inherent complexity in governing the modern state.[337] But interviews with civil servants responsible for designing forms and negotiating data-sharing and procurement contracts make clear that many civil servants simply presume that sex and gender are obvious and matters of common sense.[338] Vague statutes are also often interpreted according to common sense or ordinary meaning.[339] Unfortunately, although views are changing, most people think that sex and gender are binary and static.[340]

When they conceptualize sex and gender as "common sense" categories, the laws on the books and on the ground codify, rely on, and entrench stereotypes. For instance, as legal historian Anna Lvovsky demonstrates, anti-vice police and state liquor board agents claimed they could use "common sense" to identify gay people and, thereby, shut down bars for "'becom[ing] disorderly'" or knowingly "'permitt[ing] . . . degenerates and undesirable people to congregate.'"[341] To do so, they

336. Citron, Technological Due Process, supra note 17, at 1296–98.

337. See, e.g., Calo & Citron, supra note 23, at 813–14; Joseph A. Grundfest & A.C. Pritchard, Statutes With Multiple Personality Disorders: The Value of Ambiguity in Statutory Design and Interpretation, 54 Stan. L. Rev. 627, 640–41 (2002) (describing several reasons for ambiguity, including language, politics, and discretion delegated to administrative agencies and courts); Victoria F. Nourse & Jane S. Schacter, The Politics of Legislative Drafting: A Congressional Case Study, 77 N.Y.U. L. Rev. 575, 594–96 (2002) (documenting "deliberate ambiguity" in statutes).

338. Waldman, Opening, supra note 133 (manuscript at 21) (demonstrating the salient role of supposedly "common-sense" assumptions about sex and gender in how civil servants involved in form design do their work).

339. Smith v. United States, 508 U.S. 223, 241–44 (1993) (Scalia, J., dissenting).

340. Kim Parker, Juliana Menasce Horowitz & Anna Brown, Pew Rsch. Ctr., Americans' Complex Views on Gender Identity and Transgender Issues 4 (2022), https://www.pewresearch.org/social-trends/wp-content/uploads/sites/3/2022/06/PSDT_06.28.22_GenderID_fullreport.pdf [https://perma.cc/6DM8-FREQ].

341. Anna Lvovsky, Vice Patrol 29–41 (2021) (first quoting N.Y. Alcohol & Bev. Law § 106(6) (McKinney 2021); then quoting Record on Review at 7, Gloria Bar & Grill v. Bruckman, 259 A.D. 706 (N.Y. App. Div. 1940)); see also, e.g., Nan Alamilla Boyd, Wide Open Town 109–11 (2003); Chauncey, supra note 37, at 8–9; John D'Emilio, Sexual Politics,

relied on queer stereotypes and then arrested any man who did not meet police expectations of masculinity.[342] This same idea, that sex categorizations are common sense and that individuals obviously fit into one or the other, is still being used by those seeking to restrict the rights of transgender people to use public restrooms that accord with their gender identities.[343] Therefore, statutes and agreements that leave the words "sex" and "gender" unspecified allow supposedly "commonsense" perceptions—namely, stereotypes—to dominate how the law is implemented in practice.

C. *Persistent Legal Intervention*

Some scholars have suggested that automation and its harms have arisen in a regulatory or legal void.[344] But, as this Article shows, the law has not been hands-off. This Article's case study of sex and gender data pathways suggests that the law creates a particular kind of neoliberal state—namely, one premised on the pathologies of risk-based governance and data maximalism. This puts gender-diverse populations at risk.

The neoliberal state is thoroughly infused with market-oriented thinking: a belief that the market is the best way to advance social welfare and that only market-based options are workable.[345] Unlike the classical liberal state, neoliberal governance can be interventionist, leveraging law to enhance efficiency in institutions, minimize transaction costs, make decisions based on cost–benefit analysis, and use ever-growing information databases to deliver so-called "smart" forms of governance.[346] This type of governance relies on mass quantification, datafying as much about a population as possible and using those data to model potential future outcomes about who or what poses risks.[347]

Sexual Communities 14–15 (1983); Lillian Faderman & Stuart Timmons, Gay L.A. 28–30 (2006).

342. See Lvovsky, supra note 341, at 42 (noting that agents built their cases on the confidence "that they could spot queer men, immediately and infallibly, on the basis of the telltale mannerisms of the fairy"). For a more robust discussion of queer stereotypes that law enforcement officers and investigators relied on, see generally id. at 36–41.

343. See, e.g., Petition for a Writ of Certiorari at 14, Gloucester Cnty. Sch. Bd. v. Grimm, No. 20-1163 (U.S. filed Feb. 19, 2021), 2021 WL 723101 (suggesting that a public school should be free to make "commonsense" distinctions between male and female use of public bathrooms).

344. Calo & Citron, supra note 23.

345. David Singh Grewal & Jedediah Purdy, Introduction: Law and Neoliberalism, 77 Law & Contemp. Probs. 1, 13–14 (2014); see also Jamie Peck & Adam Tickell, Conceptualizing Neoliberalism, Thinking Thatcherism, *in* Contesting Neoliberalism: Urban Frontiers 26, 33 (Helga Leitner, Jamie Peck & Eric S. Sheppard eds., 2007).

346. See Britton-Purdy et al., supra note 326, at 1796–800 ("Planning was essential if politics was to serve the goal of efficiency.").

347. See Cohen, Between Truth and Power, supra note 16, at 183; K. Sabeel Rahman & Hollie Russon Gilman, Civic Power 124 (2019).

That poses two problems for gender-diverse populations. First, the technologies used to model risk are not neutral; rather, their "assumptions about variables and parameters are open to contestation."[348] So, too, are the decisions to weigh a particular problem as more or less of a threat and to accept a certain amount of harm as too small enough or too unlikely to require remediation.[349] If—and that is a big *if*—they account for small populations like transgender, nonbinary, and gender-nonconforming individuals, these models may accept even extreme and likely harm as insufficiently weighty.

Second, data maximalism is uniquely dangerous to those whose data are not always consistent. Under the logics of neoliberal governance, more is better because more data means better trained algorithms, better predictions, and better security at a fraction of the cost of overinclusive or "dumb" surveillance.[350] Data maximalism means "a utopian governance dream—a 'smart', specific, side-effects-free, information-driven utopia."[351] In other words, more data are supposed to allow the government to use the resources of the neoliberal state—concerned not with social welfare but with risk management—in as efficient, targeted a manner as possible.

Sex and gender data are used by the state in automated forms of "targeted governance" that identify and evaluate the presence and magnitude of risk factors in people, spaces, and activities.[352] More information is supposed to help the state do that better.[353] For example, more data are supposed to help the state distinguish between two or more people with similar names.[354] Sex and gender are not the only types of data that can do that. But that doesn't matter. Once the state commits to the neoliberal goal of targeted or smart governance, surveillance and data collection become pathologies. Collecting more data is always better.

But the state's use of gender data poses difficult-to-resolve data dilemmas for transgender, nonbinary, and gender-nonconforming individuals such that more is not always better. On the one hand, traditional approaches to collecting sexual-orientation and gender-

348. Cohen, Between Truth and Power, supra note 16, at 182.

349. Id.

350. Paul Ohm & Nathaniel Kim, Legacy Switches: A Proposal to Protect Privacy, Security, Competition, and the Environment From the Internet of Things, 84 Ohio St. L.J. 101, 144–45 (2023) (proposing a designed-in capacity for users to switch from "smart" technologies, which extract data, to "dumb" technologies, which are not targeted or algorithmically determined).

351. Valverde & Mopas, supra note 264, at 239.

352. Id. at 245.

353. Id. at 246 (explaining how believers in "targeted governance" are "highly optimistic" that continuing to collect good data will increase efficiency).

354. Citron, Technological Due Process, supra note 17, at 1274–75 (discussing how the No Fly List system erroneously captures innocent people with names similar to those of people the government is actually seeking to prevent from flying).

identity (SOGI) data erase the identities of millions of people, harming nonbinary people, LGBTQ+ elders, bisexuals, and many other marginalized groups within the queer community.[355] Therefore, more and more accurate data could improve LGBTQ+ access to healthcare,[356] help identify discrimination,[357] and highlight injustice,[358] thereby informing needed policy changes. Still, data are power, and the state has a long history of weaponizing demographic data in service of white supremacy, cisnormativity, and heteropatriarchy.[359] There is virtue in the state

355. See, e.g., Sonia K. Katyal, The Numerus Clausus of Sex, 84 U. Chi. L. Rev. 389, 406 (2017); Nancy J. Knauer, "Gen Silent": Advocating for LGBT Elders, 19 Elder L.J. 289, 342 (2012); Nancy C. Marcus, Bridging Bisexual Erasure in LGBT-Rights Discourse and Litigation, 22 Mich. J. Gender & L. 291, 295 (2015); Cara E. Trombadore, Police Officer Sexual Misconduct: An Urgent Call to Action in a Context Disproportionately Threatening Women of Color, 32 Harv. J. Racial & Ethnic Just. 153, 168 (2016); Kenji Yoshino, The Epistemic Contract of Bisexual Erasure, 52 Stan. L. Rev. 353, 459 (2000).

356. See, e.g., Kellan E. Baker, Carl G. Streed, Jr. & Laura E. Durso, Ensuring that LGBTQI+ People Count—Collecting Data on Sexual Orientation, Gender Identity, and Intersex Status, 384 New Eng. J. Med. 1184, 1186 (2021); Alex S. Keuroghlian, Electronic Health Records as an Equity Tool for LGBTQIA+ People, 27 Nature Med. 2071, 2071 (2021); Carl G. Streed, Jr., Chris Grasso, Sari L. Reisner & Kenneth H. Mayer, Sexual Orientation and Gender Identity Data Collection: Clinical and Public Health Importance, 110 Am. J. Pub. Health 991, 991 (2020); Shaun Turney, Murillo M. Carvalho, Maya E. Sousa, Caroline Birrer, Tábata E.F. Cordeiro, Luisa M. Diele-Viegas, Juliana Hipólito, Lilian P. Sales, Rejane Santos-Silva & Lucy Souza, Support Transgender Scientists Post–COVID-19, 369 Science 1171, 1172 (2020).

357. See, e.g., Gender Identity in U.S. Surveillance Grp., Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys, at xiv (Jody L. Herman ed. 2014), https://williamsinstitute.law.ucla.edu/wp-content/uploads/Survey-Measures-Trans-GenIUSS-Sep-2014.pdf [https://perma.cc/Q6AT-RAAJ]; Madeline B. Deutsch, JoAnne Keatley, Jae Sevelius & Starley B. Shade, Collection of Gender Identity Data Using Electronic Medical Records: Survey of Current End-User Practices, 25 J. Assoc. Nurses AIDS Care 657, 662 (2014); Sari L. Reisner, Kerith J. Conron, Scout, Kellan Baker, Jody L. Herman, Emilia Lombardi, Emily A. Greytak, Allison M. Gill & Alicia K. Matthews, "Counting" Transgender and Gender-Nonconforming Adults in Health Research: Recommendations from the Gender Identity in US Surveillance Group, 2 Transgender Stud. Q. 34, 37–38 (2015); Charlotte Chuck Tate, Cris P. Youssef & Jay N. Bettergarcia, Integrating the Study of Transgender Spectrum and Cisgender Experiences of Self-Categorization From a Personality Perspective, 18 Rev. Gen. Psych. 302, 303 (2014).

358. See, e.g., Leonore F. Carpenter & R. Barrett Marshall, Walking While Trans: Profiling of Transgender Women by Law Enforcement, and the Problem of Proof, 24 Wm. & Mary J. Women & L. 5, 23–30 (2017) (arguing that more accurate data would assist in proving patterns and practices of systemic profiling); Jordan Blair Woods, LGBT Identity and Crime, 105 Calif. L. Rev. 667, 675–76, 710, 724 (2017) (stating that the lack of available data makes it difficult to identify LGBT inequalities in the criminal system).

359. See, e.g., Ruha Benjamin, Race After Technology: Abolitionist Tools for the New Jim Code 36 (2019) (arguing that race-neutral technologies, laws, and policies perpetrate white supremacy); Catherine D'Ignazio & Lauren F. Klein, Data Feminism 14–17 (2020) (arguing that data historically have been used by those in power to consolidate their control); María Lugones, Heterosexualism and the Colonial / Modern Gender System, Hypatia, Winter 2007, at 186, 196 (arguing that gender differentials were a tool of colonization); Lauren E. Bridges, Digital Failure: Unbecoming the "Good" Data Subject

sometimes knowing less.[360] This is why many transgender and nonbinary individuals refuse to disclose or are uncomfortable disclosing gender identity data, even in trans-specific studies, out of concern for their privacy.[361] And because gendered classifications cannot be extricated from racial ones, transgender and nonbinary persons of color feel these harms most acutely.[362]

Scholars and advocates have long debated how to navigate this dilemma with respect to racial categories on the U.S. census and SOGI data in government surveys and in healthcare contexts.[363] Some think the state should get out of the business of collecting and using SOGI data altogether.[364] Indeed, despite how technology companies frame their algorithms' strengths, many algorithms do not need that much data to achieve their results. Several algorithmic systems that claim to make accurate predictions because they use hundreds or thousands of data inputs fare no better than standard linear regressions that use two or four.[365]

Banning certain types of data collection, sharing, and use has been central to some social movements. For instance, the movement to "ban the box" seeks, at a minimum, to remove the box to check on employment application forms if job applicants have been convicted of felonies.[366] The policy intends to stop discrimination at its source by eliminating, or at least

---

Through Entropic, Fugitive, and Queer Data, Big Data & Soc'y, Feb. 11, 2021, at 1, 14 (arguing that society has historically used data to compare others to the white, heterosexual male).

360. Daniel J. Solove, The Virtues of Knowing Less: Justifying Privacy Protections Against Disclosure, 53 Duke L.J. 967, 988–98 (2003).

361. See, e.g., Hale M. Thompson, Patient Perspectives on Gender Identity Data Collection in Electronic Health Records: An Analysis of Disclosure, Privacy, and Access to Care, 1 Transgender Health 205, 210 (2016).

362. Currah, supra note 67, at 18, 21 (noting that the gender binary is inherently a function of race and colonization).

363. Several of the many excellent explorations of the U.S. Census's collection of data on race include the sources cited supra note 36. For a discussion of how the Census undercounts members of the LGBTQ+ community, see Kyle C. Velte, Straightwashing the Census, 61 B.C. L. Rev. 69, 72–73 (2020).

364. See, e.g., Clarke, supra note 2, at 942; Katri, supra note 35, at 644, 712–14; Wipfler, supra note 35, at 529–30.

365. See, e.g., Dressel & Farid, supra note 61, at 2–3 (finding that the COMPAS risk assessment software, which incorporates 137 different data points, performed no better than a linear regression relying on two independent variables); Matthew Salganik, Ian Lundberg, Alexander T. Kindel & Sara McLanahan, Measuring the Predictability of Life Outcomes With a Scientific Mass Collaboration, 117 Proc. Nat'l Acad. Scis. 8398, 8400 (2020) (demonstrating that machine-learning methods using thousands of data points poorly predicted life outcomes and were only somewhat better than regressions using four predictor variables).

366. See Johnathan J. Smith, Banning the Box but Keeping the Discrimination?: Disparate Impact and Employers' Overreliance on Criminal Background Checks, 49 Harv. C.R.-C.L. L. Rev. 197, 200 (2014).

delaying, a data point that allows employers to screen out candidates without looking at their credentials.[367] To achieve their goal, advocates built a movement with formerly incarcerated persons and successfully lobbied city and state governments across the country to remove the criminal history box from public employment forms.[368] Similarly, some advocates have called for eliminating gender designations on birth certificates, passports, and other official documents.[369] They argue that the risks are too high and that alternative technologies exist to verify identities.[370]

But these abolitionist responses may not achieve their goals and could have unintended effects. Even if algorithms exclude certain datapoints, machine learning may still be able to identify patterns by proxy.[371] Furthermore, at least a couple of studies suggest that the current iteration of "ban the box" laws have unintended consequences; employers may be discriminating even more on the basis of race.[372] And, as Professor Jessica Clarke has shown, the relevance of sex, gender, assigned gender at birth, and gender identity varies.[373] There are powerful reasons to want "each context of sex or gender regulation [to] consider[] the relative merits of various strategies for achieving nonbinary gender rights, including third-gender recognition, the elimination of sex classifications, or integration into binary sex or gender categories."[374]

---

367. See Michelle Natividad Rodriguez & Anastasia Christman, Nat'l Emp. L. Proj., Fair Chance—Ban the Box Toolkit: Opening Job Opportunities for People With Records 4 (2015), https://s27147.pcdn.co/wp-content/uploads/NELP-Fair-Chance-Ban-the-Box-Toolkit.pdf [https://perma.cc/5983-UDR2]; see also Jessica S. Henry & James B. Jacobs, Ban the Box to Promote Ex-Offender Employment, 6 Criminology & Pub. Pol'y 755, 757 (2007) (describing how, "in addition to promoting employment discrimination against ex-offenders, the question deters ex-offenders from even applying for city jobs").

368. See Smith, supra note 366, at 211–15.

369. See, e.g., Clarke, supra note 2, at 947 (passports); Katri, supra note 35, at 644, 710–14 (birth certificates and other official documentation); Wipfler, supra note 35, at 529–30 (birth certificates).

370. See, e.g., Clarke, supra note 2, at 981–83; Katri, supra note 35, at 644, 710–14; Wipfler, supra note 35, at 529–30.

371. See, e.g., Talia Gillis, The Input Fallacy, 106 Minn. L. Rev. 1175, 1180–81 (2022).

372. See Stephen Raphael, The Intended and Unintended Consequences of Ban the Box, 4 Ann. Rev. Criminology 191, 205 (2021); see also Angela Hanks, Ctr. for Am. Progress, Ban the Box and Beyond 14 (2017), https://www.americanprogress.org/wp-content/uploads/sites/2/2017/07/FairChanceHiring-report.pdf [https://perma.cc/9S32-94VR] (arguing that "ban the box" should be "just one element of a multi-pronged strategy to remove barriers to employment that people with criminal records face").

373. See Clarke, supra note 2, at 990.

374. Id.

D. *Persistent Subordination*

The automated administrative state's approach to sex and gender data is both over- and underinclusive, harming gender-diverse populations from both sides. On the one hand, the state collects sex and gender data in a myriad of contexts. As a result, many transgender people who hold inconsistent gender designations on official documents avoid participating in daily life, from obtaining healthcare and practicing licensed professions to traveling and attending school.[375] Transgender and nonbinary people vote at lower rates than the broader LGBTQ+ community and the population at large in part because strict voter identification laws transform the voting booth into gender dysphoric triggers.[376] Knowing that the state uses sex and gender data to determine identity and maintain security, many gender-diverse populations are forced to the margins of society as they avoid the risk of harm.

On the other hand, the law, civil servants, and technology designers make decisions that exclude those who do not fit neatly in binary gender categories.[377] The law of gender data collection triggers a form design process riddled with incentives to maintain the status quo and integrates biased perceptions that sex and gender are matters of common sense, elevating the gender binary.[378] The law of gender data sharing normalizes the gender binary, conflates sex and gender, and makes all state agencies dependent on databases that look the same.[379] The law of gender data use prioritizes efficiency and immunizes algorithmic systems from interrogation, which leaves the gender binary intact.[380] To be sure, some transgender individuals can respond honestly to questions with binary answer options. But without any way of identifying who among those who check "male" are transgender men and who among those who check

375. See, e.g., Judson Adams, Halle Edwards, Rachel Guy, Maya Springhawk Robnett, Rachel Scholz-Bright & Breanna Weber, Transgender Rights and Issues, 21 Geo. J. Gender & L. 479, 532 (2020); Currah & Mulqueen, supra note 2, at 565.

376. See How Voter ID Laws Disenfranchise Transgender Americans, Democracy Docket (June 29, 2021), https://www.democracydocket.com/analysis/how-voter-id-laws-disenfranchise-transgender-americans/ [https://perma.cc/AF38-HLC2] ("27% [of transgender eligible voters] live in states with voter ID laws, but lack qualifying identification that reflects their name and gender." (citing Kathryn O'Neill & Jody L. Herman, The Potential Impact of Voter Identification Laws on Transgender Voters in the 2020 General Election 2 (2020), https://www.democracydocket.com/wp-content/uploads/2021/06/Trans-Voter-ID-Feb-2020.pdf [https://perma.cc/SD6Q-UB44])).

377. Albert & Delano, supra note 77, at 539–40.

378. See supra section II.C.

379. See supra section III.D.

380. See supra section IV.B–.D.

"female" are transgender women, transgender individuals remain hidden within the data, unable to benefit from granular insights.[381]

Some argue that substantive due process and equal protection law can effectively solve these problems. Substantive due process is supposed to guarantee fundamental rights essential to a democratic society;[382] equal protection requires that similarly situated individuals be treated similarly unless there is a valid justification otherwise.[383] Legal scholar Margaret Hu has argued that the use of data-matching systems and AI to classify certain individuals as risks of fraud, terrorism, or general criminality may constitute a violation of the presumption of innocence.[384] Several scholars argue that a state violates the equal protection clause when its algorithmic decisionmaking systems disproportionately harm certain marginalized populations.[385]

But antidiscrimination protections are hanging on by mere threads. Courts have chipped away at their efficacy in general.[386] It is particularly difficult to demonstrate discriminatory intent in the design and use of automated systems, when algorithms often operate as black boxes and when using proxy variables closely associated with protected identities can achieve discriminatory goals just as well.[387] Besides, our goal should be to do what we can to stop these problems from happening in the first place.

---

381. Albert & Delano, supra note 77, at 540–41 (referring to this phenomenon as "category-based erasure").

382. See, e.g., Obergefell v. Hodges, 576 U.S. 644, 663–64 (2015) (holding that "[t]he identification and protection of fundamental rights" is part of the Court's constitutional duties); Eisenstadt v. Baird, 405 U.S. 438, 453 (1972) (holding that withholding contraceptives from unmarried individuals "conflicts with fundamental human rights"); Griswold v. Connecticut, 381 U.S. 479, 484–86 (1965) ("The present case, then, concerns a relationship lying within the zone of privacy created by several fundamental constitutional guarantees.").

383. See, e.g., City of Cleburne v. Cleburne Living Ctr., Inc., 473 U.S. 432, 439 (1985) ("The Equal Protection Clause of the Fourteenth Amendment commands that no State shall 'deny to any person within its jurisdiction the equal protection of the laws,' which is essentially a direction that all persons similarly situated should be treated alike." (quoting U.S. Const. amend. XIV, § 1) (citing Plyler v. Doe, 457 U.S. 202, 216 (1982))).

384. Margaret Hu, Big Data Blacklisting, 67 Fla. L. Rev. 1735, 1759, 1776 (2015).

385. E.g., Barocas & Selbst, Big Data's Disparate Impact, supra note 12, at 673–76.

386. See, e.g., Cristina Isabel Ceballos, David Freeman Engstrom & Daniel E. Ho, Disparate Limbo: How Administrative Law Erased Antidiscrimination, 131 Yale L.J. 370, 375–84 (2021) ("When agencies act in ways that have significantly different effects along racial or ethnic lines, a claim to that effect is cognizable under neither administrative law nor antidiscrimination law.").

387. See, e.g., Pasquale, Black Box Society, supra note 59, at 40–41; Barocas & Selbst, Big Data's Disparate Impact, supra note 12, at 712–13.

VI. PRIVACY LAW PRINCIPLES AND NON-REFORMIST REFORMS

So far, this Article has demonstrated how law creates an automated state aimed at efficiency and, as a result, binarizes gender and erases and harms gender-diverse populations. This Part considers the normative question of the role of the state: Given the law's role in transgender and nonbinary erasure, should the state ever collect, share, and use gender data at all? If so, can the state to do so in a way that serves the interests of gender-diverse populations in an automated state rather than the disciplinary and surveillant goals of the government? I confess to being uncertain. State power has long been used to force legibility on state subjects. Even state-sponsored schemes to improve the human condition through legibility often fail inside a structure designed to do the opposite.[388] And yet, some legibility seems necessary to provide effective healthcare, enforce antidiscrimination law, and consciously account for historic marginalization and erasure. Therefore, this Part offers a tentative middle ground based on privacy principles: As advocates strive for the abolition of gender data as a classificatory, securitizing, and identification tool, we can also engage with policymakers and local, state, and federal street-level bureaucracy to find a better balance between legibility and privacy in an age of automation.

A. *Which Kind of Privacy*

Legal philosopher Anita Allen argues that historically, "Women have had too much of the wrong kinds of privacy."[389] Patriarchal forces pretextually leverage privacy to entrench traditional gender roles; "enforce isolation" in the home to cut off opportunities for growth, education, and flourishing;[390] and, in one not-uncommon but extreme case, permit a husband to abuse his wife behind the "curtain [of] domestic privacy."[391]

Gender-diverse populations suffer the same imbalance. This Article has shown that transgender, nonbinary, and gender-nonconforming

388. Scott, Seeing Like a State, supra note 21, at 309–10 ("Any large social process or event will inevitably be far more complex than the schemata we can devise, prospectively or retrospectively, to map it."); see also Eric A. Stanley, Atmospheres of Violence 118 (2021) (arguing that state efforts toward LGBTQ+ inclusion and recognition are forms of harm and that queer communities should resist state legibility generally in favor of abolitionist approaches to human flourishing).

389. Anita Allen, Uneasy Access: Privacy for Women in a Free Society 37 (1988).

390. Id. at 52. For more on the use of privacy as pretext to enforce traditional gender and heteronormative dynamics, see generally Ruth Colker, Public Restrooms: Flipping the Default Rules, 78 Ohio St. L.J. 145, 164 (2017) ("The privacy justification is actually a pretext for the articulation of gender stereotypes about the inappropriateness of men being exposed to women's private, bodily functions."); Susan Hazeldean, Privacy as Pretext, 104 Cornell L. Rev. 1719 (2019).

391. State v. Rhodes, 61 N.C. 453, 459 (1868).

individuals are erased or hidden from much public health surveillance. In these cases, they have too much of the wrong kind of privacy. At the same time, they are made legible as potential fraudsters by automated systems created by laws focusing on security, classification, categorization, and identification. Here, gender-diverse populations have too little of the right kind of privacy.

Managing state gender-data collection means reversing this imbalance. Gender-diverse populations deserve legibility or privacy when each serves human flourishing, equity, and full democratic participation. Finding that balance is precisely what queer data scientist Kevin Guyan seeks to do with his call for advocates, scholars, and representatives of affected communities to help build the state's "gender competence."[392] In other words, policymakers, street-level bureaucrats, and coders building algorithmic technologies for the state do not understand the power, limits, history, and dangers of collecting, sharing, and using gender data. They write and implement laws that collect sex and gender data without knowing why and assuming that doing so is uncontroversial common sense. They disseminate sex and gender data as if they are fungible with other pieces of information. And they use that data in algorithmic systems as if doing so has no special consequences. Our job is to teach them otherwise, growing popular consciousness along the way. Engaging with these civil servants and policymakers requires advocates to embrace the nitty-gritty of government work, but it offers opportunities for direct impact.

Those responsible for the law on the books and on the ground must have an "understanding that historical and social factors mean that equality of opportunity is a fiction, an awareness of power differences between and within LGBTQ communities, and attention to the intersection of LGBTQ identities with other identity characteristics."[393] They need to be willing "to assume a contrarian role in data discussions" that decenter traditional pathways and hierarchies of power.[394]

B. *Principles for Gender Legibility*

To achieve that goal, this Article suggests three principles, derived from privacy scholarship, to govern state gender-data practices: necessity, antisubordination, and inclusivity. A necessity principle asks whether sex or gender data are necessary to achieve a government goal, and if so, which goal. For example, as argued above, gender is an ineffective metric for security and identification; genders (and sexes) can change. Only cisgender people retain the sexes and genders they are assigned at birth; everyone else is at risk when gender is presumed static. Plus, there are so many other effective means of verifying identity, from using static traits to

392. Guyan, supra note 41, at 155.

393. Id. at 156.

394. Id.

personal histories. Therefore, using sex or gender data simply to ensure applicants for government assistance or voters or licensed professionals are who they say they are violates the necessity principle.

That said, the state has often argued that sex or gender data are necessary for some purpose it considers legitimate. Before marriage equality, for instance, sex was considered necessary for determining the validity of marriages.[395] Therefore, we need an antisubordination principle to clarify which government goals merit the use of sex or gender data—namely, those goals, like antidiscrimination and health equity, that disrupt traditional hierarchies of power and benefit gender-diverse populations. Transgender and nonbinary scholars have long argued that deficits in gender-affirming healthcare stem from, among other things, the marginalization of gender diversity in health studies, the subsequent erasure of populations not identifying as men or women from public reports and policymaking, and the ultimate neglect of gender diversity in medical and public health degree-granting programs.[396] In these contexts, taking gender into account may improve the lives of people traditionally erased.

And an inclusivity principle will ensure that when the state does need to collect, share, and use sex or gender data, it does so in ways that respect gender-nonconforming individuals. Here, transgender and nonbinary scholars have provided recommendations for how to ask for gender data in certain contexts, including providing two-step questions (asking for assigned sex at birth and gender, for example), opportunities to opt out, and spaces to self-identify.[397] This is not simply a matter of adding more boxes to gender questions on forms;[398] as we have seen, gender binaries can be entrenched in data-sharing agreements, interstate compacts, and automation mandates. Inclusivity also means writing gender diversity into law, redesigning algorithms and technologies procured from private vendors, updating legacy computer systems, and rethinking the role of gender data in the automated state from the ground up.

Although ambitious, this framework is well within the tools available under current legal discourse on privacy. Privacy law and theory are important places for inspiration here because privacy law is supposed to allow individuals to disclose certain information in certain contexts and

---

395. See, e.g., Littleton v. Prange, 9 S.W.3d 223, 231 (Tex. Ct. App. 1999) (voiding a marriage between a woman who was assigned male at birth and a cisgender man as a same-sex marriage); see also Frontiero v. Richardson, 411 U.S. 677, 690–91 (1973) (holding unconstitutional a federal law that required different qualification criteria for male and female military spousal dependency).

396. See supra notes 356–357.

397. See supra notes 354–355.

398. See Bivens, supra note 40, at 893.

withhold that information in other contexts.[399] Privacy scholars are also used to dealing with data dilemmas such as data in exchange for access and disclosure in exchange for seamless commerce.

One way privacy law tries to navigate these dilemmas while fostering prosocial behavior is through the principle of data minimization. Data minimization is the principle that organizations should collect only as much data as is absolutely necessary to achieve a stated purpose.[400] It is at the core of modern approaches to consumer privacy law, both in the United States and in the European Union.[401] In the context of an information economy in which data is used to manipulate consumers, data minimization could, if enforced effectively, starve data-extractive organizations of dangerous weapons.[402] Therefore, the principle of data minimization (or necessity) seems like a perfect antidote to the automated state's pathology of gender data maximalism.

That said, data minimization is half a loaf. It may try to stanch the flow of data, but it permits unrestricted data collection if its purpose is clearly defined, previously disclosed, and legitimate. States could easily meet that requirement, justifying gender data as necessary for verifying identity or securing spaces. Instead of relying on data minimization alone, policymakers and civil servants should also approach data collection, sharing, and use through an antisubordination lens. Privacy values do that, as well.

Over the last fifty years, much privacy scholarship has shifted from an individualistic conception of privacy to one that recognizes the inextricable connection between data, privacy, and hierarchies of power.[403] Specifically, critical privacy scholars see privacy as an antidote to

---

399. Ari Ezra Waldman, Privacy as Trust 69 (2018) [hereinafter Waldman, Trust] (discussing the privacy interests that relate to the disclosure of information); Julie Cohen, What Privacy Is For, 126 Harv. L. Rev. 1904, 1910–12 (2013) (discussing the role of privacy in society and for self-making).

400. Woodrow Hartzog & Neil Richards, Legislating Data Loyalty, 97 Notre Dame L. Rev. Reflection 356, 365 (2022), https://ndlawreview.org/wp-content/uploads/2022/07/Hartzog-and-Richards_97-Notre-Dame-L.-Rev.-Reflection-356-C.pdf [https://perma.cc/HEQ2-HH2H] [hereinafter Hartzog & Richards, Legislating].

401. Id. at 365–66; see also Regulation 2016/679, supra note 305, at art. 5(1)(c).

402. See Hartzog & Richards, Legislating, supra note 400, at 365–66.

403. Compare Alan F. Westin, Privacy and Freedom 7 (1967) (defining privacy with respect to autonomy and choice), with Neil Richards, Why Privacy Matters 39 (2022) ("'Privacy' is fundamentally about *power* . . . . Struggles over 'privacy' are in reality struggles over the *rules* that constrain the power that *human information* confers."); see also Julie E. Cohen, Turning Privacy Inside Out, 20 Theoretical Inquiries L. 1, 22 (2019) ("[C]ommon relationships in contemporary commercial and civic life . . . are about power, and privacy theory should acknowledge that fact . . . ."); Daniel J. Solove, Privacy and Power: Computer Databases and Metaphors for Information Privacy, 53 Stan. L. Rev. 1393, 1398 (2001) (arguing that the problem with information databases is that they make "people feel powerless and vulnerable, without any meaningful form of participation in the collection and use of their information").

manipulation and domination. Civil rights scholar Khiara Bridges noted this link early on; she recognized that privacy is a right of the privileged because those dependent on government services, like low-income pregnant persons of color, have no choice but to disclose personal information, accept surveillance, and submit to invasive inspections in exchange for critical medical, financial, and social support.[404]

Many other scholars have followed Professor Bridges's lead. Because of the centrality of privacy for sexually minoritized populations—including women, transgender people, and gay people, among others—law and technology scholar Danielle Citron has argued that the law should provide special protection for sexual privacy.[405] Multifaceted rules from criminal law to tort law would ensure that intimate information available to others could only be used to benefit, rather than harm, the most vulnerable.[406] In other words, Professor Citron wants privacy law to take sex into account. Professor Scott Skinner-Thompson has called for privacy law to take account of intersectional identity and provide additional protections for those subordinated by institutional marginalization.[407] Similarly, privacy law scholars Neil Richards and Woodrow Hartzog have argued that technology companies that collect and process data should not be allowed to benefit from that data if it means harming their users.[408] Like fiduciaries who are entrusted with their clients' personal information to pursue their clients' interests, state automation could be similarly informed by fiduciary values that ensure that data-driven tools will only help, not hurt, the most marginalized.[409]

These same principles can guide political and bureaucratic approaches to sex and gender data. The automated state collects, shares, and uses gender data in service of a commitment to efficient targeted governance that covers most people most of the time. That commitment takes us down a dangerous path: one in which the state collects a lot of sex and gender data while saddling transgender, nonbinary, and gender-

404. Bridges, Poverty, supra note 30, at 8–10.

405. Danielle Keats Citron, The Fight for Privacy: Protecting Dignity, Identity, and Love in the Digital Age, at xvii, xviii (2022); Danielle Keats Citron, Sexual Privacy, 128 Yale L.J. 1870, 1881–82 (2019) [hereinafter Citron, Sexual Privacy].

406. Citron, Sexual Privacy, supra note 405, at 1928–35.

407. Skinner-Thompson, supra note 42, at 6.

408. See Neil Richards & Woodrow Hartzog, A Duty of Loyalty for Privacy Law, 99 Wash. U. L. Rev. 961, 964–65 (2021) [hereinafter Richards & Hartzog, Loyalty]. This argument has a history. See, e.g., Daniel Solove, The Digital Person 103 (2004) (positing that businesses that are collecting personal information from users should "stand in a fiduciary relationship" with those users); Waldman, Trust, supra note 399, at 79–92; Jack M. Balkin, Information Fiduciaries and the First Amendment, 49 U.C. Davis L. Rev. 1183, 1186 (2016) ("[M]any online service providers and cloud companies who collect, analyze, use, sell, and distribute personal information should be seen as information fiduciaries toward their customers and end-users."); Neil Richards & Woodrow Hartzog, Taking Trust Seriously in Privacy Law, 19 Stan. Tech. L. Rev. 431, 457 (2016).

409. See Richards & Hartzog, Loyalty, supra note 408, at 966–67.

nonconforming individuals with all the dangers but none of the benefits of data-driven governance. This Article seeks a new path: one in which the state collects, shares, and uses only so much inclusive sex and gender data as is necessary to benefit, protect, and support gender-diverse populations. Achieving these goals will not be easy. Nor will they be realized tomorrow. But we can start tomorrow.

CONCLUSION

This Article begins a critical conversation about how law creates, fosters, and incentivizes a particular kind of automated governance that excludes and harms transgender, nonbinary, and gender-nonconforming individuals. The law both on the books and on the ground tends to binarize sex and gender data from collection to use. This not only harms those who exist outside of the gender binary the most but also endangers anyone subordinated by the reification of strict gender norms.

This narrative has been obscured because it is more than just statutes and court cases that are responsible for binary gender data in algorithmic systems. The on-the-ground policymaking of street-level bureaucrats, binding data contracts between state agencies, efficiency mandates, policy by procurement, and data protection compliance are all part of a larger puzzle that reveals institutionalized hostility to anyone outside the gender binary. Gender data in the automated state is, therefore, a case study in the risks posed by law: how it allocates power, how it forces legibility, and how it excludes.

But we are not without hope. In revealing the full picture of the law's role in creating an automated state that excludes gender minorities, this Article gives space for experts and members of affected communities who have long recommended inclusive approaches to gender data collection and those who argue that gender data collection is unnecessary in certain contexts. Their work, cited throughout this Article, can bring data minimization and antisubordination principles into practice. The automated state is not going away; together, we can guide it on a new, more inclusive path.