INTRODUCTION INTRODUCTION

REIMAGINING THE STATE IN A DATA-DRIVEN WORLD

By Jenna Burrell and Ranjit Singh

INTRODUCTION INTRODUCTION

INTRODUCTION INTRODUCTION

INTRODUCTION: REIMAGINING THE STATE IN A DATA-DRIVEN WORLD

By Jenna Burrell and Ranjit Singh

In a society of algorithms,[1] governments can become the biggest customer and the strongest guardrail for data-driven technologies. This duality in a government's relationship with technology companies has come to increasingly shape the emergent nature of the datafied state. On the one hand, the states have come to treat the tech industry as a partner. In the United States, this partnership has emerged under conditions of neoliberalism which has held bipartisan appeal for decades.[2] In the majority world,[3] this partnership has taken shape within the master narrative of modernization and progress — using computing and datafication as symbols of socioeconomic development.[4] On the other hand, the government in its role as a regulator has confronted companies that have monopolized most of the traffic on the internet. The patterns in the growth of the data economy in the last few decades shows how the attention of citizens can be commodified[5] as data and then processed to extract immense value.[6] To contend with these developments, the European Union has taken a more adversarial position toward tech monopolies and passed the most significant regulations to safeguard

1 Jenna Burrell and Marion Fourcade, "The Society of Algorithms," *Annual Review of Sociology* 47, no. 1 (July 2021): 213–37, https://doi.org/10.1146/annurev-soc-090820-020800.

2 Elizabeth Popp Berman, *Thinking like an Economist: How Efficiency Replaced Equality in U.S. Public Policy* (Princeton: Princeton University Press, 2022).

3 Shahidul Alam, "Majority World: Challenging the West's Rhetoric of Democracy," *Amerasia Journal* 34, no. 1 (January 2008): 88–98, https://doi.org/10.17953/amer.34.1.l3176027k4q614v5; Sareeta Amrute, Ranjit Singh, and Rigoberto Lara Guzmán, "A Primer on AI in/from the Majority World: An Empirical Site and a Standpoint" (New York: Data & Society Research Institute, September 14, 2022), https://datasociety.net/library/a-primer-on-ai-in-from-the-majority-world/; Ranjit Singh, Rigoberto Lara Guzmán, and Patrick Davison, eds., *Parables of AI in/from the Majority World* (New York: Data & Society Research Institute, 2022), http://dx.doi.org/10.2139/ssrn.4258527; Paola Ricaurte Quijano, "Ethics for the Majority World: AI and the Question of Violence at Scale," *Media, Culture & Society* 44, no. 4 (May 1, 2022): 726–45, https://doi.org/10.1177/01634437221099612.

4 Linnet Taylor and Dennis Broeders, "In the Name of Development: Power, Profit and the Datafication of the Global South," *Geoforum* 64 (August 2015): 229–37, https://doi.org/10.1016/j.geoforum.2015.07.002.

5 Thomas H. Davenport and John C. Beck, *The Attention Economy: Understanding the New Currency of Business* rev. ed. (Boston: Harvard Business Review Press, 2002); Morten Axel Pedersen, Kristoffer Albris, and Nick Seaver, "The Political Economy of Attention," *Annual Review of Anthropology* 50, no. 1 (2021): 309–25, https://doi.org/10.1146/annurev-anthro-101819-110356.

6 Nick Couldry and Ulises Ali Mejias, "The Decolonial Turn in Data and Technology Research: What Is at Stake and Where Is It Heading?," *Information, Communication & Society* 26, no. 4 (November 2021): 1–17, https://doi.org/10.1080/1369118X.2021.1986102.

competition in the data economy within their borders with uneven global implications.[7] A separate significant intervention has been the efforts of majority world countries[8] and indigenous states[9] in claiming sovereignty over their peoples' data, regardless of where it is held and by whom.

In writing the introductory paragraph, we made a slight change in terminology over the first two opening sentences: from government to state. This shift is crucial. State is a conceptual frame used to broadly articulate practices of governing a community of persons living on a definite territory. This element of territoriality has often implied a deep concomitant relationship between the state and the nation,[10] and that nationalism is an essential feature of identity-formation that makes up the state.[11] Of course, borders are one of the many ways of demarcating relations between people and practices of governance can also take on transnational forms,[12] such as the European Union. Yet as Begoña Artexaga succinctly articulates, "The state should … be thought of in ways that are not necessarily totally dislodged from the nation but neither attached to it."[13] Nation-state, however, is not the only analytical frame that can be used to unpack the nature of the state. The formation of any community relies on its peoples' commitment to follow its governance structures. When we become a part of a community, we also become a part of the state that is grounded in its practices of governance. A government is also made up of people who represent this community. It has a defined organization, usually codified in the form of a constitution adopted by the state. As a collective of representatives, the government is obligated to exercise the power of the state in the interest of the community that constitutes it. While the state is a shorthand to encompass a community, a government comprises those who make the rules on how to live within this community's variously constituted borders.

7 Payal Arora, "General Data Protection Regulation — A Global Standard? Privacy Futures, Digital Activism, and Surveillance Cultures in the Global South," *Surveillance & Society* 17, no. 5 (December 2019): 717–25, https://doi.org/10.24908/ss.v17i5.13307.

8 Begoña Aretxaga, "Maddening States," *Annual Review of Anthropology* 32, no. 1 (2003): 393–410, https://doi.org/10.1146/annurev.anthro.32.061002.093341.

9 Tahu Kukutai and John Taylor, *Indigenous Data Sovereignty: Toward an Agenda* (Canberra: Australian National University Press, 2016). See: https://docs.google.com/document/d/1cdj2rbU0XV14E2PJWwpxt-kvVeaxXEa-P8JBX0wBjI0U/edit?usp=sharing for clarity.

10 Begoña Aretxaga, "Maddening States," *Annual Review of Anthropology* 32, no. 1 (2003): 393–410, https://doi.org/10.1146/annurev.anthro.32.061002.093341.

11 Benedict Anderson, Imagined Communities: *Reflections on the Origin and Spread of Nationalism*, rev. ed. (New York: Verso, 2016).

12 Saskia Sassen, "Bordering Capabilities Versus Borders: Implications for National Borders," *Michigan Journal of International Law* 30, no. 3 (2009): 567–97, https://repository.law.umich.edu/mjil/vol30/iss3/2/.

13 Aretxaga, "Maddening States," 398.

This short detour into exploring the state and government conceptually offers a preview into the overall focus of this collection. Words matter. They matter because they often become a key to understanding practices. Keywords for any practice are words that may have broad or generic meanings but take on a certain specificity within the context of that practice. The practices of datafication in organizing the state are no different. By conducting interviews with government officials, being embedded in a government agency, or scrutinizing government documents or datasets, we come to better understand keywords that underlie the practices of infrastructuring the datafied state.

Keywords often cross over from institutional practices and make their way into the discourse of scholarship. Raymond Williams's classic book *Keywords: A Vocabulary of Culture and Society* documented the etymological evolution of words as markers of key shifts in culture and society.[14] Such terms capture our attention via their ambiguity, polysemy, or new frequency of use. Williams's work has spawned many other keywords collections.[15] Likewise, institutions and the public adopt language from scholarship. For example, the word "algorithm" has in the past decade gone from an arcane technical term taught and used by computer scientists to one used in mainstream media and pop culture — and invoked regularly within government as well.[16]

In our efforts to showcase how interpretive flexibility[17] manifests in keywords, this collection differs from traditional collections in significant ways.

First, keyword collections often present a single definition for each keyword with the author drawing from a vast array of scholarship to illustrate diversity in its meaning. However, the end result is a definition meant to be, more or less, comprehensive. In this collection, for several keywords we

14 Raymond Williams, *Keywords: A Vocabulary of Culture and Society* rev. ed. (New York: Oxford University Press, 1983).

15 Andrea Cornwall, "Buzzwords and Fuzzwords: Deconstructing Development Discourse," *Development in Practice* 17, no. 4/5 (August 2007): 471–84, https://www.jstor.org/stable/25548244; Craig Jeffrey and John Harriss, *Keywords for Modern India* (Oxford: Oxford University Press, 2014); Stephanie Nohelani Teves, Andrea Smith, and Michelle Raheja, eds., *Native Studies Keywords* (Tucson: University of Arizona Press, 2015); Benjamin Peters, ed., *Digital Keywords: A Vocabulary of Information Society and Culture* (Princeton: Princeton University Press, 2016); *The Keywords Feminist Editorial Collective, ed., Keywords for Gender and Sexuality Studies* (New York: NYU Press, 2021); AI Now Institute, "A New AI Lexicon: Responses and Challenges to the Critical AI Discourse," 2021, https://ainowinstitute.org/series/new-ai-lexicon; Nanna Bonde Thylstrup et al., eds., *Uncertain Archives: Critical Keywords for Big Data* (Cambridge: The MIT Press, 2021).

16 Tarleton Gillespie, "The Relevance of Algorithms," in *Media Technologies: Essays on Communication, Materiality, and Society*, eds. Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot (Cambridge: MIT Press, 2013), 167–93, http://ieeexplore.ieee.org/document/6733906.

17 Trevor J. Pinch and Wiebe E. Bijker, "The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other," *Social Studies of Science* 14, no. 3 (August 1984): 399–441, https://doi.org/10.1177/030631284014003004.

invited multiple teams of authors to define the same word, including public interest, bureaucracy, and counterdata. We did this to showcase how authors foreground different themes and examples to define the same keyword. We do this to disrupt the assumption that these words could or should have one single, conclusive definition.

Second, in the same spirit, we invited contributors to indicate which of the three viewpoints they planned to take in defining their keyword.

- *High-level viewpoints:* for conceptual clarification and etymological histories. This can include terms with ambiguous, multiple, or shifting definitions that are important for understanding the datafied state.
- *Viewpoints from within the datafied state:* for terms with currency and value within government.
- *Viewpoints from outside of the datafied state:* for terms that represent a critique of government generally, the datafication of government specifically, or that are used to argue for alternatives.

The first viewpoint is the more traditional one, representing the way Williams's *Keywords* collection[18] defined the genre. By explicitly including the second and third viewpoints, however, we hoped to open a door for those who find literature reviews to be a strange or inaccessible idea. Those who know what they know from being on the ground: working within government or from an outside standpoint resisting it.

Third, we encouraged academic contributors to collaborate with a first-time or non-traditional coauthor[19] who brings a distinct viewpoint, lived experience, or deeper grounding in the keyword and to think mindfully about

18 Williams, *Keywords*.

19 Mariolga Reyes Cruz, "What If I Just Cite Graciela? Working Toward Decolonizing Knowledge Through a Critical Ethnography," *Qualitative Inquiry* 14, no. 4 (June 2008): 651–58, https://doi.org/10.1177/1077800408314346.

citational justice[20] in their writing. We believed that doing so would help uplift more voices in the space of scholarship around the datafied state and expand our shared community.

Finally, this collection represents views that span the globe to highlight that the datafied state does not have a singular form. While data-driven systems as a distinct form of authority, discourse, and action have the capacity to shape the political culture of a nation-state, the state often has its own repertoire of norms, institutions, and traditions that push back.[21] The interplay between the two manifests in different meanings of a keyword in different geographies. Attending to this difference is crucial for the global project of mapping the ongoing datafication of the state.

In the following sections, we dive deeper, engaging in some definitional work to situate our readers and synthesizing the contributions in this collection to guide readers. We expand on our definition of the state by asking, "What does it mean for the state to be datafied?" We point to recurring terms like "data" and "public" that appear in multiple contributions and terms that surreptitiously found their way into several entries — "surveillance" being the most notable. We conclude with reflections on who this collection is for, what life we hope for it to have as we release it into the world, and the possible trajectory of future efforts.

What Does It Mean for a State to Become Datafied?

A state doesn't exist without the community of people it circumscribes, hence counting people has always been a constitutive element of making up

20 The Citational Justice Collective et al., "Citational Justice and the Politics of Knowledge Production," *Interactions* 29, no. 5 (August 2022): 78–82, https://doi.org/10.1145/3556549; Angela Okune, "Self-Review of Citational Practice" (Zenodo, May 21, 2019), https://doi.org/10.5281/zenodo.3066861.

21 For examples of efforts to map socio-technical change through the mutually constitutive relationship between the social shaping of technology and the technical building of nation-states, see Ranjit Singh, "Give Me a Database and I Will Raise the Nation-State," *South Asia: Journal of South Asian Studies* 42, no. 3 (May 2019): 501–18, https://doi.org/10.1080/00856401.2019.1602810; Wiebe E. Bijker, "Dikes and Dams, Thick with Politics," *Isis* 98, no. 1 (2007): 109–23, https://www.journals.uchicago.edu/doi/abs/10.1086/512835?journalCode=isis; Sheila Jasanoff and Sang-Hyun Kim, *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power* (Chicago: University of Chicago Press, 2015); Margaret Jack and Seyram Avle, "A Feminist Geopolitics of Technology," *Global Perspectives* 2, no. 1 (June 2021): 24398, https://doi.org/10.1525/gp.2021.24398.

the state. Building on these practices of counting, the development of the field of statistics by the beginning of the 19th century transformed the conception of the nation-state. German thinkers and statesmen of the time "brought to full consciousness the idea that the nation-state is essentially characterized by its statistics."[22] In fact, statistics as a keyword was initially employed to describe a "science dealing with the facts of a state."[23] Its importance was framed by articulating its relationship with history — "History is ongoing statistics, statistics is stationary history."[24]

The modern state has always been datafied; it is constructed through numbers and data.

Datafication of the state tends to take on an ominous form when thinking through sociological definitions of the state. Max Weber's definition is a case in point. In "Politics as a Vocation," he defined the state as, "a human community that (successfully) claims the *monopoly of the legitimate use of physical force* within a given territory."[25] The state "is considered the sole source of the 'right' to use violence."[26] Yet Weber also obliquely mentions that force is not the only means of the state, simply that it is its defining feature. Another of Weber's major contributions to our understanding of the state is his definition of bureaucracy as an impersonal system of rules that exists independently of any particular government worker who may be tasked with enacting it. This impersonality is a move toward fairer treatment of citizens,[27] although rule-following is rarely straightforward and often involves arbitrary forms of judgment.[28] While Jennifer Raso and Victoria Adelmant's contribution on bureaucracy notes the derogatory sense underlying the term's use in everyday speech today, Weber's work lacks such connotations. On the contrary, he articulated this impersonality and the calculability of

22 Ian Hacking, *The Taming of Chance* (Cambridge: Cambridge University Press, 1990), 18.

23 Stuart Woolf, "Statistics and the Modern State," *Comparative Studies in Society and History* 31, no. 3 (July 1989): 590, https://doi.org/10.1017/S0010417500016054.

24 Hacking, *The Taming of Chance*, 24.

25 Max Weber, *From Max Weber: Essays in Sociology*, ed. Hans Heinrich Gerth and C. Wright Mills (New York: Oxford University Press, 1946), 78, emphasis in original.

26 From Max Weber, 78.

27 "When fully developed, bureaucracy also stands, in a specific sense, under the principle of sine ira ac studio. Its specific nature, which is welcomed by capitalism, develops the more perfectly the more the bureaucracy is 'dehumanized,' the more completely it succeeds in eliminating from official business love, hatred, and all purely personal, irrational, and emotional elements which escape calculation. This is the specific nature of bureaucracy and it is appraised as its special virtue." Weber, 215–16, emphasis in original.

28 Akhil Gupta, *Red Tape: Bureaucracy, Structural Violence, and Poverty in India* (Durham: Duke University Press, 2012); Matthew S. Hull, *Government of Paper: The Materiality of Bureaucracy in Urban Pakistan* (Berkeley: University of California Press, 2012).

rules that underlie a bureaucracy's operation as its "purely technical superiority over any other form of organization. The fully developed bureaucratic mechanism compares with other organizations exactly as does the machine with the non-mechanical modes of production."[29] We can take either of Weber's definitions and consider how datafication redirects or otherwise alters the state.

Datafication can be interpreted simultaneously as (1) an amplification[30] of the state as a force and (2) as an investment in the technical superiority and impersonality of bureaucracy in organizing the state.

Approaching datafication as amplifying the state as a force raises a critical question, who is this force directed toward and against? Perhaps force by the state is desirable if it is directed at some unelected power, like the tech industry, and is wielded on behalf of a public or a marginalized group who lacks power. Yet, it is often the case that a datafied state is one in which the state and private firms link up together in deeper alignment. Neoliberal policies have often enacted this sort of model, particularly in the United States. Firms today provide services to the state that help to expand its reach and ability to oversee all people within its territorial boundaries, at its borders, and even those beyond its borders.[31] This is an alignment in which tech firms function as capture corporations (as Burcu Baykurt argues in her contribution). The force of the state is not one opposed to private tech, but interlinks with firms to exert force and control over the populace. This alignment is also a way of excising parts of the state, reducing costs and gaining efficiencies, but also making the state less able to uphold public values. Ludmila Costhek Abílio and Carolina Cruz note this in their own contribution on bureaucracy, showing how platform companies have taken on certain functions and services that traditionally belonged to the state.

29 Weber, *From Max Weber*, 214.

30 Philip E. Agre, "Real-Time Politics: The Internet and the Political Process," *The Information Society* 18, no. 5 (October 2022): 311–31, https://doi.org/10.1080/01972240290075174; Kentaro Toyama, *Geek Heresy: Rescuing Social Change from the Cult of Technology* (New York: PublicAffairs, 2015), 17–37.

31 John Cheney-Lippold, "Jus Algoritmi: How the National Security Agency Remade Citizenship," *International Journal of Communication* 10 (2016): 1721–42, https://ijoc.org/index.php/ijoc/article/view/4480.

In some cases, the populace may actively join this alignment between the state and private sector tech firms as well. Youngrim Kim, writing on data publics, argues that notions of patriotism and duty in Korea led citizens to join together with government agencies and tech companies to realize a national COVID-19 surveillance infrastructure. She argues it is a western conceit to assume citizens, through their membership in civil society groups, are primarily an opposing force that seeks to resist or reform the state.

If, alternatively, we understand datafication as a way of investing in the more logical, impersonal enactment of rules within bureaucracies, then the replacement of a bureaucrat with a computer certainly gives the appearance of greater impersonality, though it also serves to conceal the human traces and judgment calls underlying an automated façade. It also risks enacting rules in ways that are inflexible to the point of being nonsensical or even cruel. Amina Abdu and Abigail Jacobs's contribution on public administration makes this point that datafication became a way that public agencies (composed of unelected civil servants) sought to solidify their legitimacy;[32] however, new criticism of tech is calling this legitimacy into question. Considering both these possibilities together, it is not clear whether the datafied state always acts in the public interest. Anne L. Washington and Joanne Cheung argue that the public interest must be grounded in engaging with edge cases and those in the margins.

While the datafied state may rhetorically aspire to uphold the public interest, whether it does so in practice is an empirical question.

Opening up the possibilities of empirical investigation draws our attention to the everyday lived experience of the datafied state. In this respect, anthropological approaches to defining the state trace a different genealogy of the field. Anthropology has offered alternatives to the state (as a western

32 See also, Theodore Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton: Princeton University Press, 1995).

phenomenon), including social structures that operate at a smaller scale to govern populations through the normative and institutional forms that define culture, such as religion, family, civil society, and the economy. In anthropology, this notion of the state as a western phenomenon dissolved with the firmer integration of political economy into the study of culture.[33] The state was traditionally approached in the field "as a given — a distinct, fixed and unitary entity that defines the terrain in which other institutions function."[34] However, more recently, attempts have been made to "bring together the ideological and material aspects of state construction, and understand how 'the state' comes into being, how 'it' is differentiated from other institutional forms, and what effects this construction has on the operation and diffusion of power throughout society."[35] Attending to the cultural constitution of the datafied state involves following ongoing and emergent cultural struggles that are waged in two interrelated aspects of state-formation: first, in the sphere of representation, and second, in the domain of everyday practices of bureaucracies. Writing on automation, Georgia van Toorn, Chris O'Neill, Maitreya Shah, and Mark Andrejevic illustrate both these aspects in the ongoing investments in automation by governments across the world. They showcase how automation represents speed, efficiency, and precision, and enacts a cascading logic[36] that fundamentally reorganizes everyday bureaucratic work.

Exploring these ongoing forms of reorganization further, bureaucratic procedures routinely rely on tools to manage state-citizen relations. Citizen data is one such tool. Every tool has its affordances and limits; they offer a perspective embedded in their very construction. Mardiya Siba Yahaya and Bonnita Nyamwire show in their contribution how citizen data collected during the process of issuing biometrics-based digital IDs not only represents bodies of citizens made available for scrutiny at a distance, but

33 George E. Marcus and Michael M. J. Fischer, *Anthropology as Cultural Critique: An Experimental Moment in the Human Sciences*, 2nd ed. (Chicago: University of Chicago Press, 1999).

34 Aradhana Sharma and Akhil Gupta, eds., *The Anthropology of the State: A Reader* (Malden, MA: Blackwell Publishing, 2006), 8.

35 Sharma and Gupta, 8.

36 Mark Andrejevic, *Automated Media* (New York: Routledge, 2019).

also the state itself as a collective moving away from corruption and toward progress and development. The everyday experiences of living with data mutually shape the meaning of the state for bureaucrats and citizens alike.[37] As a public interest technology practitioner involved in negotiation over how and when to use digital tools in delivery of government services, Maria Filippelli offers us a window into these experiences.

Clues to the material and discursive formation of the datafied state lie in mundane bureaucratic procedures that must accommodate data-driven technologies.

Questioning and Resisting the Datafied State

The word "public" (as in "public sector") often serves as a proxy for the state. In some instances, it means ownership by the state. When the datafied state is a topic of research, data-driven systems owned and operated by the government and the infrastructures they are imbricated within become suitable case studies.[38] A number of terms in this collection include this word ("counterpublics"; "public administration"; "data publics"; "public interest"). However, the word "public" takes on a multitude of meanings in these contributions, moving beyond government ownership. In some other instances, public(s) emerge as manifestations of "amorphous and unarticulated" collectives of people who organize themselves in the face of problems and/or issues that affect them to express their concerns.[39] By acting upon such problems, John Dewey argued,[40] the public manifests its capacity to hold the state accountable. Finally, the use of the word "public" is also a part of exploring the principle of openness, as in the public disclosure of data. In her contribution, Malavika Raghavan highlights how despite diverse definitions

37 Bidisha Chaudhuri, "Programmed Welfare: An Ethnographic Account of Algorithmic Practices in the Public Distribution System in India," *New Media & Society* 24, no. 4 (April 2022): 887–902, https://doi.org/10.1177/14614448221079034.

38 Lina Denick, Joanna Redden, Arne Hintz, Harry Warne, "The 'Golden View': Data-Driven Governance in the Scoring Society," *Internet Policy Review* 8, no. 2 (June 30, 2019), https://policyreview.info/articles/analysis/golden-view-data-driven-governance-scoring-society.

39 John Dewey, *The Public and Its Problems* (New York: H. Holt and Company, 1927), 131.

40 Dewey, *The Public and Its Problems*.

of "open data," certain common features hold. Open data is concerned with publicly accessible datasets and the negotiation over the formats in which they are made available.

"Public" melds the institutions of government with the people subject to that government.

The state, in theory, represents the interests of all people within its boundaries. This framing inevitably brings us to the question — how unified are these interests? What about when these interests are at odds? What differences are being glossed over when "the public" is referred to with the singular "the"? Many of the contributions in this collection pick up on this theme. Washington and Cheung's public interest and Matthew Bui and Bianca Wylie's counterpublic both show how the public is rhetorically deployed to serve the interests of those in power. Both essays engage the notion of the public with necessary skepticism, investigating cases of urban public space management to illustrate who is implicitly included and excluded by the term.

"The public" is a term that holds power by implying a consensus that often does not exist.

Facing a state aligned with some publics over others, citizens respond using a diverse set of new and old tactics. The two contributions to this collection on counterdata illustrate these tactics in distinct ways. While Seyi Olojo takes the route of a broader historical review and makes a definitional intervention, Vanessa Massaro, Darakhshan J. Mir, Terrell Mosley, and Nathan C. Ryan re-examine how recidivism is measured in the context of the US criminal legal system. Practices of using data to counter policies and practices of the state have a long history. Olojo points to the work of

Ida B. Wells who collected statistics on lynching in the late 19th century to show the unjust and pervasive targeting of Black men by this form of extra-judicial violence. Similarly, sociologist W. E. B. DuBois created data visualizations to represent Black Life and to challenge monolithic and racist representations of Blackness.[41]

If data can be used to construct the state, it can also be used to deconstruct it.[42]

It should neither be surprising nor revelatory that many civil society groups and social movements have made acquiring, analyzing, and presenting data a part of their practices of seeking justice, policy change, or simply greater visibility.[43] In part, they seek the legitimacy that quantification and data have achieved in the modern state. They often leverage, as Raghavan also shows, the bounty that recent open data efforts have offered in accessing government data for the sake of transparency. However, gaps in data collection are pervasive. Alessandra Jungs de Almeida, Lauren Klein, and Catherine D'Ignazio, in their contribution on missing data, move beyond the reinterpretation of official data collected by the state to call attention to under-resourced efforts to painstakingly collect data to fill gaps created by state neglect or intentional silence on critical social problems. Data is neither raw[44] nor always available; it must be produced to become a resource for building as well as resisting the state. Data has politics[45] that are at play in the infrastructural processes of data collection, circulation, curation, and interpretation.[46]

It is against the backdrop of these processes of managing citizen data that contests are fought over whose interests the state aligns with and who gets left behind.

Finally, Stephanie Russo Carroll, Marisa Duarte, and Max Liboiron take these contests over data as a point of departure in their contribution on

41 Whitney Battle-Baptiste and Britt Rusert, eds. *W. E. B. Du Bois's Data Portraits: Visualizing Black America* (New York: Princeton Architectural Press, 2018).

42 For a more detailed account of how 'deconstruction' is used in STS studies of how technoscience is enacted in legal contexts and court proceedings through cross-examination of evidence, refer to the work of Sheila Jasanoff, *Science at the Bar: Law, Science, and Technology in America* (Cambridge: Harvard University Press, 1995), 211–15.

43 Stefan Baack, "Datafication and Empowerment: How the Open Data Movement Re-Articulates Notions of Democracy, Participation, and Journalism," *Big Data & Society* 2, no. 2 (December 2015): 1–11, https://doi.org/10.1177/2053951715594634; Stefania Milan and Lonneke van der Velden, "The Alternative Epistemologies of Data Activism," *Digital Culture & Society 2*, no. 2 (December 2016): 57–74, https://doi.org/10.14361/dcs-2016-0205; Emily Edwards et al., "Shaheen Bagh: Making Sense of (Re) Emerging 'Subaltern' Feminist Political Subjectivities in Hashtag Publics through Critical, Feminist Interventions," *New Media & Society*, December 7, 2021, 1–22, https://doi.org/10.1177/14614448211059121; Lauren Kogen, "From Statistics to Stories: Indices and Indicators as Communication Tools for Social Change," *The International Journal of Press/Politics*, (April 2022) https://doi.org/10.1177/19401612221094246.

44 Lisa Gitelman, *"Raw Data" Is an Oxymoron* (Cambridge: MIT Press, 2013).

45 Evelyn S Ruppert, Engin Isin, and Didier Bigo, "Data Politics," *Big Data & Society* 4, no. 2 (July 2017): 1–7, https://doi.org/10.1177/2053951717717749.

46 Ranjit Singh and Steven J. Jackson, "Seeing Like an Infrastructure: Low-Resolution Citizens and the Aadhaar Identification Project," *Proceedings of the ACM on Human-Computer Interaction* 5, no. CSCW2 (October 2021): 1–26, https://doi.org/10.1145/3476056.

Indigenous data sovereignty to break free from certain core assumptions about the state. They show how assumptions around territoriality are grounded specifically in settler colonialism, a particular practice of state building by seizing land. The result is the dispossession of Indigenous people from land and lives, cultural artifacts, as well as knowledge. Data about Indigenous people, collected in the course of research or government demographics, likewise has often been misanalyzed and misinterpreted to uphold power; it is often used against the interests of Indigenous groups from whom it was collected. They observe that "open" data is a permissive framework that always benefits settlers and their systems of government over Indigenous communities. Their contribution includes guidelines to establish the terms of collaboration around data between Indigenous and non-Indigenous groups. They ask the more fundamental question about whether and how data can be collected, who holds it, who owns it, and who has access to it, arguing that Indigenous groups globally are not another minoritized constituency who use data to make appeals to an overarching state.

Indigenous states exist with their own sovereign systems of governing relations informed by specific intellectual, ceremonial, and ancestral traditions — including relations embedded in and through data; a sovereignty that they must continually reassert and defend against duplicitous, treaty-breaking settler governments.

Conclusion: The Search for Keywords of the Datafied State

In writing up the previous two sections on defining the datafied state, we have covered all contributions to this collection. Yet, as must be obvious to

our readers, there are many keywords that have been left behind — abolition, surveillance, procurement, border control, to name a few. We have had contributors whose lives interrupted the completion of their contributions; busy bureaucrats, public servants, and experts with deep experience in public administration couldn't be brought into the collection as successfully as we had hoped. However, we see these setbacks as normal natural challenges of taking on the task of putting together a collection of keywords for the datafied state. Our intention in working toward building this community was never to be comprehensive, rather it was to invite a broader conversation on the shifting nature of the state as it appropriates ever more complex data-driven systems. Furthermore, readers will find that many of the contributions in the collection can be read from the lens of a different keyword. Many of the contributions, for example, name surveillance[47] as a foundational concern in the transformation of state-citizen relations through data. Similarly, it is hard to separate conversations on public interest and public private partnerships from discussions on procurement.[48] Yet, our effort is partial; this collection is a product of the community that we could gather around our shared research interest, while being physically located within the United States. We hope that our readers see it as a resource for gathering their own communities to engage with the ongoing emergent challenges of contending with the datafied state and as an invitation to explore which keywords matter most to them.

47 Torin Monahan and David Murakami Wood, eds., *Surveillance Studies: A Reader* (New York: Oxford University Press, 2018); Virginia Eubanks, *Automating Inequality : How High-Tech Tools Profile, Police, and Punish the Poor* (New York: St. Martin's Press, 2017); Michele Gilman and Rebecca Green, "The Surveillance Gap: The Harms of Extreme Privacy and Data Marginalization," *NYU Review of Law and Social Change* 42, no. 2 (2018): 55, https://socialchangenyu.com/review/the-surveillance-gap-the-harms-of-extreme-privacy-and-data-marginalization/; Chaz Arnett, "Race, Surveillance, Resistance," *Ohio State Law Journal* 81, no. 6 (2020): 1103–42, https://kb.osu.edu/items/70ad60c0-d30e-4e7b-b740-a7d03c0095a9.

48 Mona Sloane, Rumman Chowhury, John C. Havens, Tomo Lazovich, and Luis Rincon Alba, "Procurement as Policy: Administrative Process for Machine Learning," *Berkeley Technology Law Journal* 34, no. 3 (2019): 773–852, https://doi.org/10.15779/Z38RN30793; Mona Sloane et al., "AI and Procurement: A Primer," PDF (New York: New York University, 2021), https://doi.org/10.17609/BXZF-DF18.