

Our data, ourselves: Participation, justice, and alternative futures of data sovereignty in India

Big Data & Society
July–September: 1–14
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20539517251365235
journals.sagepub.com/home/bds



Sagnik Dutta^{1,2,*} and Suruchi Mazumdar^{3,*}

Abstract

Scholars have used the term data colonialism to designate the extractive, asymmetrical relationship between Big Tech corporations and countries in the global South. However, there is scant attention paid to how data colonialism might be reconfigured in the everyday life of postcolonial states. This paper explores how civil society actors and digital rights activists in India challenge data colonialism by articulating new meanings of data sovereignty within the larger context of the postcolonial Indian state's relationship to global Big Tech corporations. Drawing on a series of roundtable conversations with academics and digital rights activists and podcast ethnography, this paper proposes a grassroots activism-based framework of data sovereignty and challenges a state-centric, neocolonial conception of data sovereignty. First, this paper outlines resistance to data colonialism posited by digital rights activists, civil society actors and social movements through commitment to data accountability, transparency and justice. Second, we explore bottom-up negotiations that challenge top-down approaches to data ownership and data sovereignty. Activists highlighted instances of grassroots-led civil society activism in India that emphasise collective mobilisation by citizens to make governments accountable and transparent with regard to their data and thereby articulated a concept of data ownership beyond indigenous conceptions of community. These included initiatives that enhance transparency of public services for citizens, citizen collectives for migrant workers and concerted efforts of workers against state surveillance aided by domestic technological startups. We advance scholarship on data colonialism and data sovereignty by focusing on novel imaginaries of data in a postcolonial context.

Keywords

Data sovereignty, data nationalism, data colonialism, global south, participation, data justice

This article is a part of special theme on Everyday Experiences of Data Colonialism and Data Nationalism. To see a full list of all articles in this special theme, please click here: <https://journals.sagepub.com/page/bds/collections/everydayexperiencesofdatacolonialismanddatanationalism>

Introduction

In 2019, Mukesh Ambani, an industrialist and the chairperson of Reliance Industries, famously said that India's data must be controlled by Indians. In essence, this meant that Indian data should be confined to its national borders (PTI, 2019). Reliance Industries is the parent company of Reliance Jio, the Indian Big Tech giant and a 'national digital champion' (Mihelj, 2023). Ambani supported India's Hindu right-wing Bharatiya Janata Party-led national

¹Department of Culture Studies, Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, The Netherlands

²Jindal Global Law School, O.P. Jindal Global University, Sonapat, India

³Jindal School of Journalism and Communication, O.P. Jindal Global University, Sonapat, India

*Equal contribution by both authors.

Corresponding author:

Sagnik Dutta, Department of Culture Studies, Tilburg School of Humanities and Digital Sciences, Tilburg University, PO Box 90153, Tilburg, 5000 LE, The Netherlands.
Email: S.Dutta_1@tilburguniversity.edu



government's efforts to store data locally (PTI, 2019). In 2018, India's central bank announced rules to restrict financial data within its national borders. While this satisfied the domestic tech start-up sector, it understandably angered global corporations (Sridhar et al., 2022). These instances show how the postcolonial Indian state and a domestic, nationalist corporation evoked discourses of data sovereignty, ostensibly opposed to 'data colonisation' of global corporates (PTI, 2019). State-led, top-down calls for decoloniality were meant to draw attention to the historical structures of inequality and relations of resource extraction between the historical metropole and peripheries in the global South. However, such calls obfuscate power relations within the nation-state. They obscure how nation-states exercise power over their citizens using data and information control, infringement of privacy rights and surveillance.

In this article, we highlight conceptions of data sovereignty formulated by civil society activists who challenge a statist understanding of data localisation and data ownership. Activists emphasised the importance of respecting the agency of data subjects, enhancing knowledge of how their data is utilised, and participation in data governance through demands of transparency and accountability. They challenged a state-led conception of data sovereignty to articulate a grassroots framework, based upon the lived experiences of marginalised communities and activists. The term data sovereignty relies upon a homogeneous, neo-colonial conception of national territory. It implies that data produced within a territory would be bound by its sovereign rules and laws (Mejias, 2023). This may be construed as resistance to historical colonial domination. An influential strand of communication scholarship uses the term data colonialism to draw parallels between the functions of historical colonialism in global economic development and resource extraction from countries of the global South carried out by Big Tech corporations, originating from the global North and China (Coudry and Mejias, 2019). Such conceptions of data colonialism, however, blur new forms of power asymmetries within specific cultural contexts. The concept of data sovereignty pits powerful actors, often construed as countries claiming ownership over data, in opposition to each other. This ignores bottom-up voices and excludes diverse groups such as migrants, refugees and asylum seekers from sovereign spaces.

This paper explores the lived experiences and negotiations of civil society groups and digital rights activists with concepts of data colonialism and data sovereignty in the postcolonial Indian context. Given the Indian state's complex relationship with global Big Tech corporations and domestic tech companies, it is a productive case for advancing a novel understanding of data colonialism and data sovereignty. In May 2020, the Indian state banned the popular short video platform, TikTok, owned by Chinese Big Tech giant Bytedance, following border disputes with China. Discourses of data security and data protection

were evoked to rationalise the ban. In contrast to its contentious relationship with China, the Indian state, however, historically perceived the United States and Silicon Valley giants (in addition to domestic conglomerates) as allies as it positioned itself as a 'digital power' (Athique and Parthasarathi, 2020). Yet, the number of takedown orders of content, issued by the state to tech giants such as the microblogging platform Twitter/ X for violating local laws, shot up in recent years (The Wire Staff, 2024). The Indian state remained committed to supporting a domestic start-up sector in the post-TikTok digital space. Such instances contextualise the Indian state's shifting relationship with global Big Tech giants from the global North and South, in which the discourses of data sovereignty were the focal point.

In this paper, we propose alternative imaginations of data sovereignty that draw upon grassroots activism inspired by the principles of open data and alliance between anti-surveillance and social justice movements. We foreground voices of civil society actors and digital rights activists and explore how such groups attribute new meanings to data sovereignty, within the larger context of the post-colonial Indian state's relationship to global Big Tech corporations. Drawing on secondary sources such as media reports, a series of roundtable conversations with academics and digital rights activists, and podcast ethnography, this paper first outlines the resistance to data colonialism through commitment to data accountability, transparency and justice. Second, it examines bottom-up negotiations that challenge top-down approaches to data ownership and data sovereignty. These include concepts of agency with respect to data, data knowledge, data participation, and data accountability and transparency. These findings contribute to and advance existing scholarship by outlining how novel forms of contestations with data emerge in post-colonial contexts.

The article is divided into sections as follows. First, we discuss the existing scholarship on data colonialism and data sovereignty and situate our contribution. Second, we lay out our research methodology. In the subsequent sections, we discuss myriad aspects of data sovereignty posited by civil society activists and digital rights activists. Finally, we have a discussion and conclusion section where we outline the implications of our findings and indicate directions for future research.

Data colonialism, data sovereignty and new imaginaries of ownership

It has been argued that the elite beneficiaries of data colonialism – Big Tech corporations, originating largely from the United States and China, where the state participates in a hybrid alliance with these corporations – create 'data subjects' using the extractive practices of historical colonialism

and abstract quantifying, computing methods (Birhane, 2023; Calzati, 2022; Couldry and Mejias, 2019, 2024; Kwet, 2019; Pinto, 2018; Posada, 2021; Lehdonvirta, 2022). Data colonialism can be understood in terms of the ‘effects on the bodies, affects and territories of marginalised and multi-ethnic populations’; colonial power based on data colonialism is thus an ‘epistemic order’ based on data (Ricaurte, 2019: 353). A new form of colonialism is being practised by multinational corporations in the United States and other parts of the Global North, based upon economic domination and control over domains of political and social life. Tacheva and Ramasubramaniam (2023) used the term AI empire to gesture towards a new model of colonial control based upon interlocking ideologies of heteropatriarchy, racial capitalism, white supremacy, and modernity/ coloniality. Calzati (2021) reiterated the need to delineate the ‘context-specific motives of datafication, especially in the peripheries’ and analyse the power relations between domination and citizenship within specific regions and contexts (915–916).

Historically, data sovereignty referred to the state’s supreme authority over data pertaining to its population within a territory (Sridhar et al., 2022). It implied that data produced within a certain territory would be bound by the laws and rules of that territory (Mejias, 2023). Since the leak of the National Security Agency’s classified information by Edward Snowden, multiple countries including Brazil, Germany and India have tightened measures, restricting data flow from their respective countries to others (Sridhar et al., 2022). Data localisation, or the restriction of data within national borders, is often legitimised using arguments of sovereignty and data safety and security concerns (Sridhar et al., 2022). Local control of data translates into censorship and control of information and dissent by political elites in totalitarian and authoritarian contexts such as China (Creemers, 2020), Iran, Egypt and Vietnam (Sridhar et al., 2022).

National sovereignty is premised upon a European conception of an ethnically homogeneous nation-state (Mamdani, 2020; Mejias, 2023). European imperial powers transported the conception of the nation-state to colonised societies. Consequently, colonised societies were understood as constituted by nations (Mamdani, 2020). Even in moments of transition from colonialism to postcolonial nationhood, the idea of the nation and nationalism constituted the framework for state construction. This led to conflicts and contestation over national sovereignty in parts of Asia and Africa (Mamdani, 2020). Sovereignty, therefore, connotes a particular relationship between the ruler and the ruled, which endured in postcolonial India (Kapila, 2022: 18). Sovereignty is constituted not through liberal conceptions of freedom, property and rights but through the power of the state to dispossess people (Kapila, 2022: 18). Sovereignty, thus, resides in a relationality constituted through involuntary dispossession and loss. The concept

of data sovereignty can be understood against the backdrop of this exclusionary aspect of India’s state-centric sovereignty. Existing models position states and individuals as the bearers of sovereignty.

Data sovereignty has been variously used to signify national data sovereignty (Irion, 2012); the right of a nation to collect and manage its own data (Rainie et al., 2017: 5–6); indigenous peoples and nations’ right to govern the ‘collection, ownership, and application of data about their peoples, lands, and resources’ (Garrison et al., 2019: 506); regulation of data being governed by the laws of a sovereign nation where the data is located (Hippelainen et al., 2017: 645). Relatedly, the term digital sovereignty is used to designate control over the domain of the digital. Digital sovereignty is constituted by the control of data, software, standards and protocols, hardware, services and infrastructures that constitute the domain of the digital (Floridi, 2019).

There are also attempts to depart from state-centric notions of sovereignty and use the term to reclaim the rights of marginalised groups. In Latin American societies, digital sovereignty is linked to the autonomy of ‘communities and assemblies’, and it asserts the ‘beliefs and struggles of Indigenous people’ (Bravo, 2017, cited in Lehuédé, 2024). This translates to the construction of autonomous digital infrastructure and community ownership of technologies (Lehuédé, 2024). Māori communities in Aotearoa New Zealand evoke the framework of indigenous data sovereignty (Kukutai and Taylor, 2016). Such groups seek to regulate the data that they generate, its usage and the terms of its future community use (Couldry and Mejias, 2024).

Alternative models also include feminist or trans-hacker cooperatives’ control over data. Examples include Brazil-based MariaLab, which introduces feminist principles in technological spaces; Alternative Laboral Trans in Argentina, a non-profit collective founded by trans people; and organisations working on digital rights such as Datysoc in Uruguay (datysoc.org), Hiperderecho in Peru (hiperderecho.org) and InternetLab in Brazil (internetlab.org.br/en) (Couldry and Mejias, 2024). These projects reject data colonialism and advocate for the collection, processing and analysis of data in keeping with community principles (Couldry and Mejias, 2024). Amrute and Murillo (2020) argued that the term data colonialism does not adequately describe the multiple roles that information technologies take across Southern locations and formulate the concept of computing from the South. People from the global South are seen by the authors (Amrute and Murillo, 2020) as less in the shadow of the West; they focus instead on how ordinary citizens live out their relationships with digital technologies, including those which do not affect positive social change.

Community-driven, indigenous responses differ from state-led data sovereignty, which has been subject to critiques of exclusion. Mejias (2023) drew attention to the statist frameworks’ exclusionary function, and contends that

migrants, refugees and asylum seekers become objects of ‘datafied persecution’, being framed as outsiders to sovereign spaces. Developed nations exercise their sovereignty by blocking the entry of these so-called outsiders into the nation, even while facilitating profit generation by corporations that develop surveillance and border control tools (Mejias, 2023). This is tied to a ‘racialised and colonialist notion of nation-building (Mejias, 2023). In postcolonial contexts, coloniality and sovereignty function in complex ways. The Indian state constructs data sovereignty arguments by using narratives of economic self-sufficiency, security and democracy (empowerment of citizens) (Basu, 2024).

In the Indian context, both the state as well as national digital champions advocated for data localisation. Their calls for data colonisation remained couched in the language of decolonisation (Udapa and Dattatreya, 2022). The postcolonial Indian state rhetorically deployed the language and discourse of decolonisation and populist narratives of sovereignty, especially in times of geopolitical conflict, to argue for data localisation (Basu, 2024). But beyond the pale of rhetoric and emotive appeals, the state selectively allows select Big Tech corporations to thrive and even aligns with them. For instance, Meta recently invested \$5.7 billion in Jio Platforms Limited, a part of Reliance Industries Limited, with an eye on the 60 million small businesses across India (Fischer, 2020). Google tied up with the Union Health Ministry in India to launch a digital health ID, which promises access to health services and personal health records on Google Wallet (Sharma, 2024). The state’s approach to data sovereignty is, therefore, performative; it positions itself as a sovereign nation-state that can rightfully assert control over the data of its citizens in moments of geopolitical conflict. The promotion of certain Big Tech corporations that thrive on the large-scale appropriation of Indian citizens’ data alternates with takedown orders sent to global microblogging platform Twitter/ X (The Wire Staff, 2024) and the promotion of a domestic tech start-up sector.

In India, data sovereignty is not merely a domain of contestation between the state, global North- or China-based Big Tech corporations and so-called outsiders such as migrants and refugees. Debates over data positioned digital rights activists, civil society actors and citizens from marginalised communities in conflict with the state and state-led notions of data sovereignty. The country, for instance, witnessed massive Internet shutdowns from 2019 onwards that were meant to contain popular protests against the national government’s policies. The shutdowns were challenged by social movements, citizens and rights groups (Munjal, 2021). Activists and academics saw the narratives of data sovereignty as essentially tied to nationalism and as attempts to control the population (Basu, 2024).

In this article, we explore the meanings of data sovereignty in India by foregrounding the voices of civil society and digital rights activists in relation to the messiness of the

Indian state’s negotiation with big data. We explore how colonial genealogies of sovereignty play out in a post-colonial, global South context. We posit that the Indian case challenges Eurocentric concepts and attempts to conceptualise new frameworks of data sovereignty. This intervention allows us to respond to calls to move beyond ‘data universalism’ and bring to the forefront the heterogeneity and cultural specificities of diverse contexts (Calzati, 2021; Milan and Treré, 2019). In the Indian context, we find the term data sovereignty productive as it aptly captures the range of meanings of sovereignty with respect to data including the state’s appropriation and regulation and the community’s claims on data.

We respond to Milan and Treré’s call for a plural, dynamic definition of ‘a global South(s) as a place of and (a proxy for) alterity, resistance and creativity, embracing the dynamism and multiplicity of interpretations while going beyond geopolitical denomination’ (2019: 325). Scholars (Arora, 2024: 326; Milan and Treré, 2019) caution against the instrumentalisation of the term global South by illiberal, authoritarian regimes against so-called Western democratic systems and rights-based data governance, and the romanticising of the category of the South, and argue instead for a critical approach to the biases, incongruences and contradictory aspects of Southern knowledge (2019: 326). Raghunath (2024) argued for a critical Southern standpoint for data governance that challenges the hegemonic powers of the nation-state and Big Tech corporations and focused instead on mediations between the individual and the community rooted in relational autonomy. We foreground how the meaning of data sovereignty is constituted in a situated context such as India through the interventions of civil society and digital rights activists. Our approach pays attention to the power dynamics and recovers the meanings of alternative approaches to data sovereignty. We are attentive to the coloniality of the postcolonial nation-states as well as creative responses to the same by activists who recover new meanings of data sovereignty.

Research methodology

We deployed a combination of research methods for this paper: four online roundtables with ten participants, conducted over four months, podcast ethnography and secondary research with media reports. Digital rights activists, social movement representatives in India, and scholars working on related issues participated in the roundtables that lasted up to two hours and culminated in a podcast series. Each podcast episode lasted for about 30 min. The methodology of online roundtables is based on the premise of creating dialogue and collaboration between researchers and activists. Our shared vision of constructing an alternative notion of data sovereignty through a praxis-based, community-driven, decolonial approach inspired this methodology.

This methodology is inspired by the AI4Dignity project (Udupa et al., 2023) that contributes to a decolonial critique of the role of AI in content moderation by complicating the category of the human in human-machine communication.

Drawing on unique community-based, postcolonial theoretical and methodological insights, the AI4Dignity project created a forum of dialogue/collaboration between ethnographers, researchers and on-ground fact checkers in multiple cultural settings from both the global South and North locations (Udupa et al., 2023). Such a deliberative, participatory approach, which emerged as a critical research methodology in existing scholarship, has been adopted in the past through other collaborative dialogue-based initiatives such as citizen councils (Wong et al., 2023) and Data Justice Lab (Hintz et al., 2022). Critical methods such as citizen councils or dialogue-based initiatives allow researchers to listen to a cross-section of participants; these methods evoke participants' responses to and engagement with complex socio-technical topics (Wong et al., 2023).

Dialogue-based, open participatory methods are effective for understanding public attitudes and community responses towards technology and policy. Inspired by this method, we conducted roundtable conversations with digital rights activists, civil society activists and academics where they deliberated upon the meanings of data sovereignty and data ownership against the backdrop of the Indian state's evolving relationship with tech giants. Such methods can be traced to the focus group tradition that prioritises voice, representation, capacity and an investment into the researchers' relationship *with* and *between* participants (Wong et al., 2023). This methodology is pertinent to our project for its commitment to praxis and the shared vision of a decolonial future.

In the process of putting together the podcast as well as participating in conferences and workshops organised by civil society activists, we immersed ourselves intimately in the everyday activities of activists. We gained insights into shared understanding and culturally grounded meanings of data activism. Thus, we brought to bear upon an ethnographic sensibility to our podcast method. We participated as podcast 'producers and interacted with the primary speakers. In this sense, we positioned ourselves as part of the 'in-group' or the 'community of listeners, guests, content providers, and other podcasters' (Lundström and Lundström, 2021: 291). We captured and examined the interaction between podcast speakers who reproduced a particular universe; the podcast ethnographic method doubled as the window to a wider milieu (Lundström and Lundström, 2021). Our engagement as podcast producers allowed us to overcome the limitations of one-sided listening or non-participant observation and the ethical challenges of ethnographic lurking, which are often said to plague digital ethnographic methods (Lundström and Lundström, 2021).

Past scholarship (Lundström and Lundström, 2021) has highlighted the importance of podcast ethnography as a

critical digital ethnographic method for exploring and engaging with the field. Podcasts are publicly available audio files that capture casual conversations between two or more speakers and retain the spirit of 'everyday social interaction' (Lundström and Lundström, 2021: 296). It has been seen as a useful methodological approach to study controversial subjects such as the practice of white nationalism in Sweden (Lundström and Lundström, 2021) and is said to enable a 'thick description of data' (Geertz, 1973, cited in Lundström and Lundström, 2021). In this paper, we position our podcast series as an ethnographic field site (Lundström and Lundström, 2021).

The specific research objective – how civil society groups, activists and academics negotiated with a state-centric vision and conjured alternative imaginations of data sovereignty – necessitated this methodological approach. Our podcast series, released later as pre-recorded, edited conversations between researchers and the participants, doubled as a medium of participation and for sharing the knowledge with the wider public. The podcast method encouraged the art of listening and cultural literacy on socio-technological issues. It allowed us the opportunity to amplify the voices of digital rights groups. Through this method, we positioned the research participants as central to the study and gained better access to the field and the site of activity of digital rights activists and civil society actors.

We contacted activists and researchers from the Internet Freedom Foundation, a New Delhi-based digital rights group, and the Centre for Internet and Society, a forum for policymakers, activists and academics. These forums have been at the forefront of activism on data justice and privacy rights in India. We attended two workshops conducted by these groups in Delhi on data justice and digital rights over a period of one year. As part of the participant observation method, we focused on the activities and conversations of activists, academics and journalists who participated in the workshops and immersed ourselves in the field. Some of our participants spoke at these forums, which familiarised us with their activism. One of our participants was an activist-cum-scholar who employed social media platforms and digital technologies for advocacy during the anti-farm law mobilisations that took place in India in 2020 and 2021. We also included legal rights activists and scholars who were known for their work on data localisation and Internet shutdowns in India. We remained cognisant of the gender, caste and religious diversity of the participants and tried to ensure that the group was representative. However, activists and legal scholars from Muslim minority groups declined our invitation to participate in the online roundtable because of the sensitivity of the subject. All participants' informed consent was collected via email.

We had three rounds of questions at first, where each participant was invited to introduce themselves, their work and activism, and asked to respond to a prompt. In

the following rounds, we requested participants to comment on each other's responses. Interview questions were shared in advance. The interview questions were conceived keeping in mind the following research aims and objectives: 1. Participants' responses to state-led data sovereignty; and 2. Participants' departure from top-down notions of data sovereignty and ethical conceptualisations in the context of India's digital and data politics. We invited the participants to observe specific instances of the Indian state's changing relationship with big Tech giants such as Twitter/X, Google and TikTok.

The online roundtable sessions lasted for up to two hours; each podcast episode, released on Spotify, lasted for 20 to 30 min. The online roundtable and podcast ethnography methods were complemented by secondary research with media reports. The media reports were selected from electronic databases, based on the key themes emerging from the online roundtables. We treated the media reports, such as newspapers, magazines and digital news reports, as documents, testaments of important political or community actions and different from research outputs (Karppinen and Moe, 2012). The data from the online roundtables and the podcast series were analysed inductively using a grounded theory approach. Themes such as knowledge, accountability and transparency were derived from the data. The themes were derived from the media reports related to the data of the online roundtables and podcast sessions. We aim to represent different meanings attributed by our study participants to the concept of data sovereignty.

Towards activism: Data knowledge, agency, participation and accountability

The Indian state's iterations of data sovereignty can be conceptualised as 'narratives' (Basu, 2024) of localisation and national security. In contrast to state-centric narratives, we underline a decolonial grassroots activism-centric framework of data sovereignty, as highlighted by our respondents, based upon the principles of data knowledge, agency, participation and accountability. This captures the tension between top-down, state-centric and grassroots activism-based notions of data sovereignty.

Data colonisation, data sovereignty and the Indian state: A view from below

This section aims to explore the contemporary Indian state's narrative of data sovereignty and situate it within the discourse of data colonialism. Drawing from the existing scholarship and our roundtable conversations, we suggest that the postcolonial Indian state's responses to data sovereignty must be assessed in the light of 'narratives' (Basu, 2024), discourses, ambiguities, knee-jerk reactions to geopolitical conflicts, and the current administration's imperative

to project the image of a powerful global actor in the Indian public consciousness rather than direct opposition to global tech giants. The Indian state exercises its sovereignty by controlling the data of citizens in neo-colonial ways. In doing so, we draw upon a colonialist understanding of sovereignty that frames the state's treatment of its citizens, especially caste and religious minorities.

Prime Minister of India Narendra Modi, who hails from the right-wing BJP, popularised a discourse of digital sovereignty premised upon a vision of the social and economic empowerment of Indian citizens, economic rationale and discourses of security (Basu, 2024). In June 2020, the Indian state evoked the security narrative and the imagery of 'digital strike' to ban 200 Chinese apps including the popular short video-making platform TikTok after a military confrontation with China in a contentious border region (Basu, 2024). Modi's conflation of national sovereignty and security can be linked to his promotion of a muscular majoritarianism. Such a masculine leadership style purportedly enables him to protect and secure the country from both internal and external enemies (Srivastava, 2015). Modi represents the trend of strong man, authoritarian populist leadership across the world that brings together religious, ethnic majoritarianism and a focus on neoliberal economic policies (Sinha, 2021). The discourse of digital sovereignty doubles as yet another instrument to promote the image of muscular, manly leadership that panders to a populist rhetoric of insecurity.

In the run-up to national elections in April 2024, the Election Commission of India, the country's poll body, issued takedown orders to the microblogging platform X, formerly known as Twitter, asking for the removal of four posts on the grounds of violation of the model code of conduct in relation to the criticism of political parties and workers (The Wire Staff, 2024). One respondent, who is a digital rights activist, observed that through much-publicised confrontations with Twitter/X the Indian state shaped public discourse in India and reasserted the image of a powerful state among elite digital users such as politicians and journalists. This respondent noted that the Indian state's equation with global tech giants oscillated in a continuum, ranging from co-operation (e.g. joint public-private programmes), co-option, confrontation, to coercion (YouTube and Twitter/X's compliance with takedown orders can be seen as instances of coercion). The exercise of control remains messy, disorganised and chaotic. Our respondents observed that the Indian state shied away from clearly articulating guidelines on data localisation or Internet shutdowns because of its geopolitical interest in partnering with the United States in the so-called global democratic block. Thus, state-led data sovereignty in India was different from a centralised form of authoritarian control over data and the digital realm as is practised by the Chinese state. Indian state's knee-jerk regulation of data is challenged by digital rights activists and civil society activists, especially those from marginalised communities, through legal

activism, civil society alliance and collection of sensitive data on takedown orders of social media content by the Indian state and death of workers during the pandemic.

Localising Indian data versus agency

On the question of sovereign authority over Indian data, the state's narrative aligns with that of major Indian private corporations, who feel that Indian data should be localised within Indian territory. In the roundtable conversations, activists posited alternative conceptions of how one might lay claim to data. This points to the limitations in the state's narrative of data sovereignty. The alternative imaginaries of data emphasised by the activists focused on the agency of the people with respect to their data, knowledge of data, freedom to articulate dissent while claiming their data, and accountability and transparency with regard to data. In the activist imaginary, data sovereignty is tied to people's agency and ability to control what happens to personal data. The large-scale collection of data for biometric identification and its subsequent use by multiple private players is understood as a moment of loss of agency over data by the activists. This can be contrasted to the state's understanding of data as a national resource linked to national sovereignty. Another participant argued that with the passage of the new DPDP (Digital Personal Data Protection) Act in 2023, the government creates a new imaginary of people as 'data principals' as opposed to 'data subjects' in the GDPR (General Data Protection Regulation).² Activists, however, felt that this only created an illusion of agency, a 'false promise' without giving people control over their data. One of the activists quipped, 'But do we really have a say in how our data will be utilised?'

A report by the Ministry of Electronics and Information Technology in 2021 emphasised the concept of sovereignty vis-à-vis data of Indian citizens. The report stressed that 'India has rights over data, its people and its organisations' (Ministry of Electronics and Information Technology, 2021; cited by Basu, 2024). In this iteration, the sovereign nation-state is seen as the final arbiter of the data of its citizens and organisations. Prime Minister Narendra Modi lauded the narrative of digital India as aiding the construction of a self-reliant India or *Atmanirbhar Bharat*.

In our roundtable conversations, activists pointed out the pitfalls of a statist conception of sovereignty. It obscures how the state appropriates the voices of citizens when it claims to speak on their behalf and asserts control over their data. A participant underlined how the agency of citizens is elided in conversations of data sovereignty. They stressed how Aadhaar³, a biometric identification of Indian citizens, was originally meant to facilitate the distribution of welfare benefits, but it is now used by private players for providing other services such as air tickets, railway tickets and bank accounts. This is not mandated by the Aadhaar Act. Hence, with Aadhaar, the government created a space

where private data no longer belongs to citizens. A participant concurred, 'Agency no longer lies with the people. That's where data colonialism comes in'. The government merely acts as an enabler rather than an active player in welfare governance by using big data. Moreover, the government allows for the creation of a regime where big data is available to private players, and it makes data subjects vulnerable to their whims. The logic followed by the government is that data should be available across platforms and interoperable, and this would lead to the most efficient form of governance.

Exploring the contesting claims of data sovereignty in an Indian context helps us expand our understanding of both data sovereignty and data colonialism. Data colonialism is not merely practised by multinational corporations situated in the global North. In this instance, a postcolonial state is engaged in a form of data colonialism. Alternative conceptions of data sovereignty emphasise the agency of the people and their desire to exercise control over their data.

Data sovereignty vs data 'knowledge'

One of the activists said that data sovereignty meant that people had 'knowledge' about how their data should be utilised. This can be described as data knowledge, whereby there are limitations on the purposes for which the state could utilise data and the duration for which data collected by the state could be retained. The concept of data knowledge prioritises the agency of data subjects and their right to exercise control over their data. The consent of the subjects of data cannot be presumed continually. One of the activists eloquently pointed out the meaning of data sovereignty and ownership:

I think data ownership would be the knowledge that if my photograph is taken for a driving licence it probably would be used for a driving license and not for a crime investigation or if my photo has been taken for Aadhaar then it will be used only for that purpose. (Interviewee B, 2024)

Another activist concurred that the state could not presume consent in perpetuity for data that it had collected for a particular purpose. They said:

The Data Protection Act gives the state the ability to digitise any information that it obtains if you have ever interacted with the state for a licence, a certificate, or something else, which is pretty much all of us, right? It then assumes that it has the consent to use that data in perpetuity. Once it's taking data from you, it has complete control over it. It's assuming control of data as a national asset. (Interviewee C, 2024)

The state's pervasive and arbitrary use of data becomes a significant concern for marginalised communities historically victimised by the violence of a carceral state. One of the activists, who has worked on state surveillance and

policing in Delhi, highlighted how new technologies of surveillance such as facial recognition technology used by the police rely on data in government databases to arbitrarily arrest people in crime hotspots. The activist emphasised the danger and precarity that this puts data subjects in:

Somebody in an agitated position with, let's say, an instrument in his hand, let's say a rod or, you know, whatever can be trying to protect himself but based on a photograph that you have in the database or the driving license or Aadhaar, you take a CCTV footage. You then identify a person (*using FRT*) and arrest that person. As we have seen a lot of cases in North-East Delhi, where the case has come to trial and the judges quashed all the evidence because it was not strong enough. But then the accused have already spent like 2–3 years in jail. Who accounts for it?. (Interviewee D, 2024)

In this vignette, the activist highlights how new technologies of governance and the state's pervasive use of data exacerbate the precarity of marginalised communities who remain at the receiving end of violence by a carceral state. In this context, the concept of 'data knowledge' acquires significance as it empowers citizens against arbitrary and harmful utilisation of private data by the state. This acquires salience against the backdrop of nationwide protests to a controversial citizenship law that ostensibly discriminated against Muslims and the incidence of communal riots in Delhi. The police used facial recognition technology to target Muslim citizens involved in the citizenship protests. As is highlighted by the activist, in these moments, the ready availability of data that the state 'owns' aids police violence and the targeting of religious minorities. In this context, activist conceptions of data sovereignty become particularly significant for safeguarding the interests of minority citizens. The significance of data usage only for particular purposes for which it is collected by the state acquires salience in a situation where easy harvesting of data by the state can lead to incarceration and undue harassment. While the state posits a meaning of data sovereignty that focuses on the sovereignty of the state, activists conceptualise data sovereignty in relation to the agency of the people, their ability to exercise control over their data and the knowledge of how data is being utilised. It is pertinent to mention that there remains a gap between data knowledge as understood and conceptualised by civil society activists who are the focus of this paper and individuals belonging to marginalised communities who are not always aware of the nuances of data protection and data utilisation. Activist efforts are aimed at bridging this knowledge gap and increasing awareness about the real-life ramifications of data knowledge among ordinary citizens.

The anxious state, data sovereignty and dissent

Activists felt that the Indian state's assertion of data sovereignty is shaped by anxiety over control and territory, even

while it is couched in a language of social and economic empowerment. Activists posited an alternative conception of data sovereignty as a significant concept that allows subjects of data to articulate dissent. Data sovereignty became a meaningful concept for ordinary citizens when they were disgruntled with the state and exercised claims over data to collectively articulate their voices against state action.

The theme of anxiety helps us delineate yet another aspect of how the Indian state's assertion of data sovereignty works in practice. The state posits data sovereignty as a matter of Indians owning and utilising their data for empowerment. But the state's assertion of sovereignty over data comes from a place of anxiety over the articulation of dissent and losing control over citizens.

One of the participants in the roundtable pointed out how taking control over the data of the people comes not from a place of power but that of weakness and anxiety. Activists also posited an alternative conception of data sovereignty that allows so-called data subjects to articulate dissent. Referring to anti-farm law protests that took place across the country in 2020 and 2021, one activist felt that the question of data sovereignty translated into protection and control over data by the government, especially in situations where the latter felt weak and anxious. The state regulated data in such instances by clamping down on social media accounts, especially those on Twitter/X and seen as errant. Data sovereignty, hence, became a site where the anxieties of the government over control and authority were staged.

Activists felt that data sovereignty became a useful concept only in moments where people were unhappy or disgruntled with the state. In such instances, the state could demonstrate its legitimacy and power by satisfying citizens and allaying their concerns. Instead, at present, whenever people express dissent, their voices get banned. In effect, they lose whatever means they had to protest. The postcolonial state's anxiety about legitimacy and sovereignty is moored in colonial genealogies of state sovereignty and control. This anxiety also plays out vis-à-vis data sovereignty in contemporary India. In British India, the exercise of colonial sovereignty was shaped by paranoia and anxiety around threats to the empire (Finden and Dutta, 2024). In postcolonial India, a range of emergency and security legislations were enacted by the state because of its anxiety about territorial sovereignty (McQuade, 2020).

Beyond ownership, towards participation

Grassroots movements for accountability, transparency and localised resistance to data colonialism enhance bottom-up responses to a statist conception of data sovereignty. Our respondents articulated the importance of the framework of justice and participation and the need to claim justice and accountability from the state in relation to public data. Data sovereignty in this sense is synonymous with democratic structures and social movements that historically

strengthened the frameworks of accountability, justice and citizens' control of data. It emerged from our roundtable conversations that community ownership of data could not be translated seamlessly into an Indian context. Therefore, the Indian case presents to us novel ways of thinking beyond the framework of indigenous community ownership of data. In the Indian context, any idea of a pristine community is debatable as religion, tribe and caste-based hierarchies determine social, cultural, political and economic relationships (Azaghu Meena et al., 2023). Our respondents observed that the very notion of appointed leaders taking decisions on behalf of the community countered democratic principles in a society such as India that is marked by deep hierarchical structures. Activists highlighted the instances of grassroots-led civil society initiatives that enhance the transparency of public services, the formation of citizen collectives to help migrant workers during the COVID-19 pandemic, concerted efforts of workers against state surveillance aided by domestic technological startups and organised opposition of sanitation workers against AI-driven facial recognition systems. Such examples, as conceptualised by participants, contribute to bottom-up notions of data sovereignty.

When it comes to a community's interaction with data-driven technologies, the lived experiences of members must be considered (Bokil et al., 2021, cited in de Souza and Bhardwaj, 2024). It has been argued that collective interest is pertinent when the impact of data-driven systems is varied and unequal (Milan and Treré, 2019). A report published in 2020 by a committee of experts, constituted by the Indian government, articulated a framework of community rights as a form of governance for non-personal data from a legal perspective (de Souza and Bhardwaj, 2024).

Our respondents reiterated the importance of social movements and democratic structures that enable accountability and transparency in relation to data. The grassroots activism-based alternative framework of data sovereignty, proposed in this paper, aligns with data justice principles that reiterate the importance of open data (Johnson, 2014) and increased collaboration between anti-surveillance and social justice activism (Dencik et al., 2016). The grassroots activism-based, bottom-up imaginaries of data sovereignty, as highlighted by our respondents, draw from India's social movement traditions that historically championed the principles of open data, public accountability-based governance frameworks and anti-surveillance movements. Data sovereignty, as understood by civil society activists, is not inherently in opposition to the principles of transparency and accountability. Activists reconfigure the concept of data sovereignty to emphasise the importance of public accountability in governance of their data and not merely as a state-oriented conception of sovereignty, understood in terms of territorial restriction and localisation of data.

Data accountability and transparency against neo-colonial datafication

Since the early 2010s, India has witnessed various forms of data-centric activism led by social movements that focused on ideals of social justice, democracy, equality and citizens' control over data. This can be traced back to the campaigning that led to the passage of the Right to Information Act in 2005. In many non-Western contexts such as societies across Latin America, digital sovereignty refers to the rights of local communities and assemblies to regulate the data that they generate and autonomous digital infrastructures (Lehuedé, 2024). However, in India, social movements and rights groups claimed accountability and transparency from the state, especially in digital public infrastructure and welfare delivery systems and aimed to make publicly available information legible to citizens. One of our respondents, a legal rights activist, highlighted the importance of citizen-led initiatives and organisations such as DataKind Bengaluru and LibTech that initiated the struggle to establish infrastructures of accountability by engaging with open government data (OGD). This respondent observed:

DataKind... aggregate publicly available information from sources and make them legible to citizens... citizens are trying to work with it to make more accountability, transparency happen.... (Interviewee A, 2024)

The notions of sovereignty overlap with data justice in such imaginations. Taylor (2017) connected open data to the larger questions of information justice that hold relevance for data as a tool of governance and governmentality. Hanbal et al. (2023), for instance, observed that the availability of datasets in India's livelihood programmes such as the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA) contributed to the claims of accountability and the questioning of government performance. LibTech, a research project of The Program on Liberation Technology at Stanford University, for instance, worked towards implementing rights-based legislations such as the NREGA, National Food Security Act (NFSA) and the Forest Rights Act that were introduced between 2004 and 2013 and 'were imagined as attempts to improve participatory democracy and the bargaining power of the marginalised' (Narayanan, 2024). Collectives such as LibTech sought to enable 'citizens' participation in governance' through 'datafication' (Hanbal et al., 2023). OGD (Open Government Data) enthusiasts assume that open government data leads to open government principles, with the data being made available to citizens; this affords the opportunity to voice their concerns (Hanbal et al., 2023). This respondent also mentioned 'Covid tracker initiatives' by citizen collectives in India. Citizen-led digital media networks such as Stranded Workers Action Network (SWAN) relied on publicly available information (such as media reports) and

information passed by word of mouth/ region-centric volunteer networks to create a database of migrant workers in need of food, ration and financial help. Such initiatives contrasted with the absence of government data on job losses and the deaths of migrant workers during the pandemic.

Marginalised groups challenged contemporary forms of social control through data-driven technologies such as Aadhaar, India's largest biometric information-based digital identity system, that had a colonial genealogy. Considering the colonial lineage of biometric identification systems such as fingerprint scanning (Cole, 2009; Sengoopta, 2003), the challenge to Aadhaar remains pertinent. One of our respondents, who is a legal scholar and activist, reiterated the role of grassroots organisations such as the Safai Karmachari Andolan or the sanitation workers' movement in the opposition to Aadhaar. Bezwada Wilson, who was born into a family of manual scavengers and led the manual scavenging eradication movement, was among the petitioners who challenged the constitutional validity of Aadhaar, often referred to as the 'pioneering national ID platform' (Rao and Nair, 2019).

The genealogy of technologies such as fingerprint scanners and biometric matching algorithms can be traced to the racial profiling practices of colonial subjects in British India as well as to 'slave branding' (branding enslaved people with a hot iron to mark or signify Black bodies as commodities) (Glouftsiou and Casaglia, 2022). A fingerprint registration scheme, for instance, was introduced by the colonial administration in British Bengal in the 19th century to identify and eliminate suspected practices of corruption and fraud – natives were suspected of collecting pensions by impersonating deceased pensioners (Glouftsiou and Casaglia, 2022). The technological advancement and application of fingerprinting in the present day is the outcome of the new ability to compute and compare biometric data (Rao, 2019). The colonial apprehension of suspected fraud, 'duplicate', 'fake identities' by the poor plagued post-colonial governments in India. Rao and Nair (2019) observed that government welfare delivery systems continue to be riddled by concerns of 'ghost', 'duplicate' and 'fake identities', which are said to prevent the poor from reaping the benefits of development programmes.

The biometric identity system's adoption and roll-out in 2019 remained a contested terrain as diverse pro-democracy groups, citizen collectives, social movement organisations, groups representing tribal women such as Adivasi Women's Network and digital rights groups – such as Rethink Aadhaar, Article 21 Trust, the Internet Freedom Foundation, the Bachao Project (a collective of Muslim women experiencing online safety violations) and the Free Software Movement of India – opposed the national government's proposed move to link Aadhaar with voter identification cards in 2021 (The Wire Staff, 2021).

Grassroots responses to new forms of data colonialism, led by the contemporary postcolonial state and the domestic

tech start-up sector, remained relevant in the face of public concern over data breaches and privacy violations. One respondent observed that systems such as GPS-tracking of sanitation workers, by local governments in the northern city of Chandigarh, lacked procedural safeguards on data security and failed to articulate suitable penalties for data breaches. The system had access to workers' Aadhaar numbers apart from location details. Data security and consent were important to our respondents. Another respondent, who worked with the CIS, observed that instances of data breaches became frequent in India during COVID, especially when it came to medical data.

Khabar Lahariya, a news network led by Dalit women in multiple Indian languages from the northern and central states, recentred the claims of citizens' data breach and privacy concerns in bottom-up narratives of data sovereignty, as pointed out by our respondents. In a news report published in November 2020, *Khabar Lahariya*, for instance, drew attention to how Indian private corporations such as Chennai-based start-up Garuda Aerospace that manufactured drones or unmanned aerial vehicles (UAV) worked with state governments and the local police to step up 'sanitation and surveillance' in upcoming, 'smart cities' such as Varanasi, Rourkela, Raipur and Kanpur at the time of the pandemic (Harekal et al., 2020).

Sanitation workers in Bangalore and Madurai opposed and demanded the removal of AI-driven facial detection systems and fingerprint-based biometric sensors that were used to track workers' in-time and out-time at work (Azaghu Meena et al., 2022). In 2020, the Safai Karmachari Andolan, a network championing the rights of sanitation workers, organised strikes in the Indian city of Chandigarh to protest GPS-enabled watches that the latter were forced to wear (Chaturvedi, 2020) by the municipal corporation, a form of local government in India, to improve work efficiency. Sanitation workers opposed technology-enabled tracking systems for reasons of inconvenience rather than the violation of the basic right to privacy. Wilson, the national convenor of Safai Karmachari Andolan, as quoted in an article published by *MediaNama*, an information portal and resource hub on technology and policy in India, said that supervisors resumed the casteist practice of 'bonded labour' and 'modern day slavery' (Chaturvedi, 2020) with GPS-enabled watches that enabled minute automated monitoring of workers' movement.

The grassroots activism-based model of data sovereignty thus entails resistance to new forms of colonialism. States and the domestic start-up sector are the perpetrators of this new mode of extraction. Such an activism-based framework also aligns with Taylor's (2017) capabilities-based approach of data justice that privileges context and group-specific negotiation with data in everyday lived experiences. While some groups and societies may identify the benefits, others may challenge datafication and surveillance as oppressive structures (Taylor, 2017).

Discussion and conclusion

In India, the state-led efforts of data localisation and national data sovereignty often fulfilled an economic rationale. While national data sovereignty in India should be seen in terms of narratives (Basu, 2024) and discourses, the state was committed to boosting the local start-up sector and catering to the interests of domestic corporates and national digital champions such as Reliance. The national digital technological infrastructure was supported by Big Tech firms such as Facebook and other Silicon Valley giants. This article moves beyond a statist, neo-colonial conception of data sovereignty and explores ways in which novel conceptions of data sovereignty are articulated by civil society actors and digital rights activists through a grassroots activism-based framework that encapsulates the principles of data agency, knowledge, transparency and accountability. Such grassroots activism-based alternative imaginations of data sovereignty, as proposed in this paper, draw from the principles of open data (Johnson, 2014) and increased collaboration between anti-surveillance and social justice activism (Dencik et al., 2016). Being strongly entrenched in conceptions of data justice, such activism-based alternative imaginations depart from India's state-led 'narratives' of data sovereignty (Basu, 2024). The latter remains entrenched in 'narratives' (Basu, 2024) of localisation and selective cooperation and conflict with global Big Tech corporations.


The resistance to data-driven, colonialist technologies such as Aadhaar, for instance, was led by marginalised caste groups such as the Safai Karmachari Andolan or the anti-manual scavenging movement. Caste-oppressed and women's groups often remained the subject of and opposed the national digital surveillance infrastructure, initiated by the local start-up and domestic corporate sectors. Such domestic digital infrastructure, also supported by global corporations, bolstered the Indian state's claims to national sovereignty. The instances of grassroots opposition by backward caste communities to novel forms of data colonialism and national data sovereignty coincided with citizens' calls for accountability and transparency through open data movements. Activists also underlined the importance of agency vis-à-vis the private data of individual citizens and groups in contrast with the state's emphasis on national sovereignty. Data sovereignty also becomes a significant concept that enables citizens to articulate dissent. Citizens' claims to data sovereignty contrast with the state's anxious, neo-colonial attempts to regulate data. Finally, digital rights activists and civil society actors in India organised collectively to ensure accountability and transparency in the use of government data. These collectives symbolise new notions of community mobilised by digital rights activists as they negotiate the government's large-scale use of data.


The article, thereby, contributes to the scholarship of data colonialism and digital/data sovereignty in various

ways. First, this article draws attention to how data colonialism and data sovereignty might be conceptualised in a postcolonial context such as India. It highlights the practices of novel forms of data colonialism and sovereignty, practised by the postcolonial Indian state, Big Tech giants and domestic start-up sectors, beyond the bipolar power centres of the United States and China (Couldry and Mejias, 2019). Second, this article outlined the myriad ways in which alternative conceptions of data sovereignty are articulated collectively by digital rights activists and civil society actors against the backdrop of a statist, neo-colonial understanding of data sovereignty posited by the postcolonial Indian state. This model of collective claim-making with data is different from claims of data sovereignty based on indigeneity (Lehuedé, 2024). It rests on notions of collective citizenship that emphasise transparency and accountability.

The article complicates the conception of the global South and highlights the peculiarities within specific cultural contexts, as it sheds light on alternative imaginations of data sovereignty that are different from Chinese, Latin American and European models. Unlike China's state-led conceptions, our respondents emphasise questions of agency, dissent, transparency and accountability. Such decolonised articulations directly challenge state-oriented conceptions of data sovereignty. Alternative models of data sovereignty, based on principles of resistance, data knowledge, accountability and transparency, also question notions of indigenous sovereignty and community-owned data, as prevalent in Latin American societies (Lehuedé, 2024). Therefore, we reclaim a space of postcolonial agency in discussions on data sovereignty, data justice and data colonialism, a conversation that often leaves out grounded experiences of diverse postcolonial subjects.

ORCID iDs

Sagnik Dutta  <https://orcid.org/0000-0002-8719-8296>

Suruchi Mazumdar  <https://orcid.org/0000-0003-0995-5455>

Ethical consideration

Ethics approval was obtained from JGU Research and Ethics Review Board.

Consent to participate

Informed consent was obtained verbally and over email from all interviewees.

Consent for publication

Informed consent for publication was obtained by email.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. We would like to thank our students Khushi Sukhija, Manushree Sarkar, Moli Shah, Sheikha Mariyam Sam, Aryan Gupte and Tanishi Ranjan for their research assistance with the podcast series.
2. The question of cross-border transfer of data of Indian nationals came up in discussions on earlier iterations of the DPDP Act. In 2018, a ten member committee of experts headed by Justice BN Srikrishna came up with a report highlighting key concerns around data protection in India. The report argued for personal data to be stored on servers within India and transfers within the country to be subject to certain safeguards. A 2019-version of the Bill allowed for the transfer of certain categories of data only if the country provided adequate level of protection. The 2022-draft bill which eventually became the Act adopted a different approach where the Central government could notify where any personal data might be transferred. For more details see: <https://prsindia.org/billtrack/digital-personal-data-protection-bill-2023>
3. Aadhaar is a unique 12-digit identity number assigned to Indian citizens and linked to a centralised database of biometric and biographical information.

References

- Amrute S and Murillo LFR (2020) Introduction: Computing in/ from the South. *Catalyst: Feminism, Theory, Technoscience* 6(2). <https://doi.org/10.28968/cftt.v6i2.34594>.
- Arora P (2024) Creative data justice: A decolonial and indigenous framework to assess creativity and artificial intelligence. *Information, Communication & Society*: 1–17. doi: 10.1080/1369118X.2024.2420041.
- Athique A and Parthasarathi V (2020) *Platform Capitalism in India*. Cham: Palgrave Macmillan.
- Azhagu Meena SP, Veeraraghavan R, Kapania S, et al. (2023) Inheriting discrimination: Datafication encounters of marginalized workers. In: *ICTD '22: Proceedings of the 2022 international conference on information and communication technologies and development*, (19), pp.1–11.
- Basu A (2024) India's digital sovereignty narrative. (unpublished – on file with the author).
- Birhane A (2023) The algorithmic colonization of Africa. In: Cave S and Dihal K (eds) *Imagining AI: How the World Sees Intelligent Machines*. Oxford: Oxford University Press, pp. 247–260.
- Bravo L (2017) A seed sprouts when it is sown in fertile soil. In: *Technological Sovereignty*. Vol. 2. Barcelona: Descontrol, pp. 109–122. Available at: <https://sobtec.gitbooks.io/sobtec2/> (accessed 23 February 2023).
- Calzati S (2021) Decolonising “data colonialism” propositions for investigating the realpolitik of today's networked ecology. *Television & New Media* 22(8): 914–929.
- Calzati S (2022) Decolonising “data colonialism”: Propositions for investigating the realpolitik of today's networked ecology. *Television & New Media* 22(8): 914–929.
- Chaturvedi A (2020) “Bonded Labour”: sanitation workers in Chandigarh protest against being forced to wear GPS tracking devices. Available at: <https://www.medianama.com/2020/10/223-chandigarh-smart-watches-sanitation-workers-2/>. (accessed 31 August, 2024).
- Cole SA (2009) *Suspect Identities: A History of Fingerprinting and Criminal Identification*. Cambridge, MA: Harvard University Press.
- Couldry N and Mejias U (2019) Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media* 20(4): 336–349.
- Couldry N and Mejias U (2024) *Data Grab: The New Colonialism of Big Tech and How to Fight Back*. Chicago: The University of Chicago Press.
- Creemers R (2020) China's conception of cyber sovereignty: Rhetoric and realisation. In: Broeders D and van den Berg B (eds) *Governing Cyberspace: Behaviour, Power and Diplomacy*. Lanham, Maryland, USA: Rowman & Littlefield, pp.107–144.
- de Souza S and Bhardwaj K (2024) India's conception of community data and addressing concerns for access to justice. *DISO* 3(16). <https://doi.org/10.1007/s44206-024-00102-5>.
- Finden A and Dutta S (2024) Counterterrorism, political anxiety and legitimacy in postcolonial India and Egypt. *Critical Studies on Terrorism* 17(2): 176–200.
- Fischer D (2020) Facebook invests \$5.7 billion in India's Jio platforms. Available at: <https://about.fb.com/news/2020/04/facebook-invests-in-jio/> (Accessed 22nd February 2025).
- Floridi L (2019) Translating principles into practices of digital ethics: Five risks of being unethical. *Philosophy & Technology* 32(2): 185–193.
- Garrison NA, Hudson M, Ballantyne LL, et al. (2019) Genomic research through an indigenous lens: Understanding the expectations. *Annual Review of Genomics and Human Genetics* 20(1): 495–517.
- Glouftsiou G and Casaglia A (2022) Epidermal politics: Control, violence and dissent at the biometric border. *EPC: Politics and Space* 41(3): 567–582.
- Hanbal R, Prakash A and Srinivasan J (2023) Seeing data like a state: A case of open government data in India's livelihoods program. *Information Polity* 28(3): 1–17.
- Harekal K and Devi M and Khabar Lahariya Bureau (2020) Deep impact: COVID-19, surveillance technology and marginalised identities. Available at: <https://khabarlahariya.org/deep-impact-covid-19-surveillance-technology-and-marginalised-identities/> (accessed 16 August 2024).
- Hintz A, Dencik L, Redden J, et al. (2022) Civic participation in the datafied society: Towards democratic auditing? Data Justice Lab. Available at: https://datajusticelab.org/wp-content/uploads/2022/08/CivicParticipation_DataJusticeLab_Report2022.pdf (accessed 16 August 2024).

- Hippelainen L, Oliver I, Lal S (2017) Towards dependably detecting geolocation of cloud servers. In: Yan Z, et al. (eds) *Network and System Security*. Cham: Springer, pp.643–656.
- Interviewee A (2024). Interview by Sagnik Dutta and Suruchi Mazumdar (Zoom), 18th April.
- Interviewee B (2024) Interview by Sagnik Dutta and Suruchi Mazumdar (Zoom), 18th April.
- Interviewee C (2024) Interview by Sagnik Dutta and Suruchi Mazumdar (Zoom), 18th April.
- Interviewee D (2024) Interview by Sagnik Dutta and Suruchi Mazumdar (Zoom), 18th April.
- Irion K (2012) Government cloud computing and national data sovereignty. *Policy & Internet* 4(3–4): 40–71.
- Kapila K (2022) *Nullius: The Anthropology of Ownership, Sovereignty, and the Law in India*. HAU books. Chicago: The University of Chicago Press.
- Karppinen K and Moe H (2012) What we talk about when we talk about document analysis. In: Just N and Puppis M (eds) *Trends in Communication Policy Research: New Theories, Methods and Subjects*. Bristol: Intellect, pp.3–20.
- Kukutai T and Taylor J (2016) Data sovereignty for indigenous peoples: Current practice and future needs. In: Kukutai T and Taylor J (eds) *Indigenous Data Sovereignty: Toward an Agenda*. Canberra: Australian National University Press, pp.1–24.
- Kwet M (2019) Digital colonialism: US empire and the new imperialism in the global south. *Race & Class* 60(4): 3–26.
- Lehdonvirta V (2022) *Cloud Empires: How Digital Platforms are Overtaking the State and How We can Regain Control*. Cambridge, MA: MIT Press.
- Lehuedé S (2024) An alternative planetary future? Digital sovereignty frameworks and the decolonial option. *Big Data & Society* 11(1): 20539517231221778.
- Lundström M and Lundström T (2021) Podcast ethnography. *International Journal of Social Research Methodology* 24(3): 289–299.
- Mamdani M (2020) *Neither Settler Nor Native*. Cambridge, MA, London: Harvard University Press.
- McQuade J (2020) *A Genealogy of Terrorism: Colonial Law and the Origins of an Idea*. Cambridge: Cambridge University Press.
- Mejias UA (2023) Sovereignty and its outsiders: Data sovereignty, racism and immigration control. *Weizenbaum Journal of the Digital Society* 3(2): 1–9.
- Mihelj S (2023) Platform nations. *Nations And Nationalism* 29(1): 10–24.
- Milan S and Treré E (2019) Big data from the south(s): Beyond data universalism. *Television & New Media* 20(4): 319–335.
- Munjal D (2021) In India, are internet shutdowns in accordance with law? Not always. Available at: <https://www.newslaundry.com/2021/10/29/in-india-are-internet-shutdowns-in-accordance-with-law-not-always> (accessed 30 August, 2024).
- Narayanan R (2024) What would an X-Ray of india's compassion reveal? The Wire. Available at: <https://thewire.in/politics/what-would-an-x-ray-of-indias-compassion-reveal>
- Pinto RA (2018) Digital sovereignty or digital colonialism? *SUR* 27: 15–26.
- Posada J (2021) The coloniality of data work in Latin America. In: Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21), May 19–21, 2021, p. 2, Virtual Event, USA. NewYork, NY, USA: ACM.
- PTI (2019) Mukesh Ambani urges PM to take steps against data colonization. *The Economic Times*, 18 January 2019. Available at: <https://economictimes.indiatimes.com/tech/ites/mukesh-ambani-urges-pm-to-take-steps-against-data-colonisation/articleshow/67585615.cms?from=mdr> (accessed 16 August, 2024).
- Raghunath P (2024) Critical data governance: A southern standpoint to the study and practice of data. *Technology and Regulation* 2024: 37–46.
- Rainie SC, Schultz JL, Briggs E, et al. (2017) Data as a strategic resource: Self-determination, governance, and the data challenge for indigenous nations in the United States. *International Indigenous Policy Journal* 8(2): 1–29.
- Rao U (2019) Population meets database: Aligning personal, documentary and digital identity in aadhaar-enabled India. *South Asia: Journal of South Asian Studies* 42(3): 537–553.
- Rao U and Nair V (2019) Aadhaar: Governing with biometrics. *South Asia: Journal of South Asian Studies* 42(3): 469–481.
- Ricaurte P (2019) Data epistemologies, the coloniality of power, and resistance. *Television & New Media* 20(4): 350–365.
- Sengoopta C (2003) *Imprint of the Raj: How Fingerprinting was Born in Colonial India*. London: Macmillan.
- Sharma NC (2024) Google to revolutionise Indian healthcare with ABHA ID on Wallet and AI innovations, says top exec Bakul Patel. *Business Today*. Available at: <https://www.businesstoday.in/industry/pharma/story/google-to-revolutionise-indian-healthcare-with-abha-id-on-wallet-and-ai-innovations-says-top-exec-bakul-patel-448800-2024-10-04> (Accessed 22 February 2025).
- Sinha S (2021) Strong leaders', authoritarian populism and Indian developmentalism: The Modi moment in historical context. *Geoforum* 124: 320–333.
- Sridhar V, Potluri SR and Rao S (2022) Data localization and its effects on cross border digital trade. In: Sridhar V (ed) *Data-Centric Living Algorithms, Digitization and Regulation*. London and New York: Routledge, pp.262–282.
- Srivastava S (2015) Modi-masculinity: Media, manhood, and “traditions” in a time of consumerism. *Television & New Media* 16(4): 331–338.
- Tacheva J and Ramasubramanian S (2023) AI Empire: Unraveling the interlocking systems of oppression in generative AI's global order. *Big Data & Society* 10(2): 20539517231219241.
- Taylor L (2017) What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society* 4(2): 205395171773633.
- The Wire Staff (2021) ‘Dangerous move’: Over 500 individuals, orgs decry EC proposal to link Aadhaar, Voter ID. Available at: <https://thewire.in/rights/dangerous-move-over-500-individuals-orgs-decry-ec-proposal-to-link-aadhaar-voter-id> (accessed 16 August 2024).
- The Wire Staff (2024) X says it disagrees with ECI orders to take down political posts. Available at: <https://thewire.in/government/>

- x-twitter-election-commission-takedown-orders-disagrees (accessed 16 August, 2024).
- Udupa S and Dattatreya EG (2022) *Digital Unsettling: Decoloniality and Dispossession in the Age of Social Media*. New York: New York University Press.
- Udupa S, Maronikolakis A and Wisiolek A (2023) Ethical scaling for content moderation: Extreme speech and the (in) significance of artificial intelligence. *Big Data & Society* 10(1). <https://doi.org/10.1177/20539517231172424>.
- Wong YN, Jones R, Das R, et al. (2023) Conditional trust: Citizens' council on data-driven media personalisation and public expectations of transparency and accountability. *Big Data & Society* 10(2). <https://doi.org/10.1177/20539517231184892>.