



India's Approach to Data Decolonization: Moving Away from the "Data as Resource" Metaphor

10
22

| Mihir Kaulgud



Issue Brief

TABLE OF CONTENTS

1. Abstract	03
2. Introduction	04
3. India's Paradoxical Data Governance Framework	04
4. Data Sovereignty: Whose Sovereignty Is Prioritised?	05
5. Conclusion: A Need To Shift Metaphors And Emphasise Data's Social Aspects	06
6. References	07

Cover Image credits: pexels-cottonbro-4427639

If you have any suggestions, or would like to contribute, please write to us at contact@sprf.in

© Social and Political Research Foundation™

October 2022

ISSUE BRIEF

India's Approach to Data Decolonization: Moving Away from the "Data as Resource" Metaphor

| Mihir Kaulgud

ABSTRACT

This paper critically examines India's data governance framework and its geopolitical dimension. It does so specifically within the context of standard-setting frameworks of the United States, the European Union, and China. These frameworks conceptualise data as a resource and revenue-generating economic asset. This work argues that while the Indian government has been vocal about data sovereignty and combatting data colonialism, it carries over the "data as resource" metaphor into its policy framework. Adopting this metaphor perpetuates the economic dimension of data colonialism, which involves big tech companies profiting by extracting and monopolising people's data. Moreover, this paper argues that to combat data colonialism, the Indian emphasis on data sovereignty must not focus only on the state's sovereignty but on the people who generate data. Focusing on people's sovereignty over their data would entail being attentive to data's social characteristics and moving the foundations of policy away from the "data as resource" metaphor. Innovating data governance from this perspective would entail a more comprehensive data decolonisation, making India a standard-setter in its own right.

Keywords:

Data Governance, Data Colonialism, Data Sovereignty, Data as Resource, Strategic Autonomy

INTRODUCTION

Data governance refers to governmental and inter-governmental regulations about collecting, storing, transferring, and using digital data. It is necessary because “data in digital form are [sic] routinely created, gathered, and shared across the globe to support core societal functions, including healthcare systems, transportation, international commerce, and national security” (Arner et al., 2021). Governments thus consider digital data a valuable asset. Its flow must be regulated in a manner that benefits the citizens and their economies (Ciuriak, 2022).

This paper argues that India’s current stance on data governance carries over a ‘data as resource’ mindset that current standard-setters in the field, such as the US, the EU, and China, have adopted. This mindset perpetuates data colonialism of tech companies over people’s data. The paper then argues that there is a need to shift to being attentive to data’s social characteristics. Moreover, while India emphasises data sovereignty, this paper argues that to combat data colonialism, emphasis on data sovereignty must not focus only on the state’s sovereignty but on the people who generate data. Innovating data governance from this perspective provides India with a window to become a standard setter in its own right.

In terms of global digital governance, there are broadly two camps of governance styles (Basu & Nachiappan, 2020). The first one is that of the United States. It is interested in allowing data flows across borders with little intervention. This occurs specifically to accelerate Silicon Valley innovation and economic benefits. The European Union has modified the US approach by implementing regulations concerning consumer and citizen protection. The second camp consists of countries from BRICS and the Global South, which emphasise data sovereignty and reject data colonialism. In essence, data colonialism is “the idea that foreign technology firms (based largely in the US) are reaping profits from the data generated by citizens from the Global South” (Basu & Nachiappan, 2020). According to Basu et al. (2019), data sovereignty refers to the sovereign right of each country to regulate data that originated within its borders. In parallel with data colonialism, this version of data sovereignty ensures that a country’s data is not used for foreign private companies’ profit but for improving domestic public services and economic growth. Such formulations of data colonialism and data sovereignty, which will be critiqued later in this paper, emphasise the state and cast these issues as questions of geopolitics. China is the most notable state in the second camp. It envisions “an authoritarian internet, where technologies of surveillance and identification help ensure social cohesion and security” (O’Hara & Hall, 2018). The Indian government has also desired to establish India as a pioneer in democratic data regulation. India’s active participation in framing global internet policy in the UN General Assembly and other forums has prompted analysts to note its capacity to shape global policy (Basu & Sherman, 2020).

Due to the differing attitudes and priorities of the recognised standard setters, such as the US, the EU, and China, the data regulatory landscape is fragmented (Arner et al., 2021). Despite their diverging approaches, the standard setters view data as a resource. Data is repeatedly framed either as “a natural force to be controlled and as a resource to be consumed” (Puschmann & Burgess, 2014). Data metaphors such as these are vital because they influence public consciousness, policy, governance, technology designs, and business models (Sinha & Basu, 2021; Watson, n.d.). Thinking about data as a resource frames it as something that can be “extracted, refined, valued, bought and sold in different ways” (The Economist, 2017). In other words, data is considered something “that is simply ‘out there’, up for grabs; ready to be mined” (Kovacs & Ranganathan, 2019). Data is also considered an asset that generates revenue, which is an extension of its conceptualisation as a resource (Birch et al., 2021).

INDIA'S PARADOXICAL DATA GOVERNANCE FRAMEWORK

India's emerging policy framework around data governance coalesces around the notion of 'data sovereignty' (Basu & Nachiappan, 2020). While some aspects of India's data governance are based on the European Union's approach, such as India's Personal Data Protection Bill being based on the EU's General Data Protection Regulation [GDPR], its heavy centralisation of power regarding data with the government is reminiscent of the Chinese approach. Mishra (2021b) writes that in India, data is treated as a nationally produced asset over which the government should have control. The underlying assumption in India's policy framework is that the government will equitably distribute the benefits of extracting commercial value from the country's data while protecting it from foreign companies and governments. In the emerging data governance framework, the government has pushed for varying levels of data localisation¹ to combat data colonisation.

The government has access to vast data via exclusive control of the Aadhar biometric data and is effectively exempted from its own rules regarding access to citizens' data. This retained data is meant to "aid state power particularly when it comes to public service delivery and welfare provision." (Basu & Nachiappan, 2020). However, the government is also a market player with concentrated economic power. The government has, at various points, expressed a desire to monetise citizens' data. For example, the Draft India Data Accessibility and Use, released in February 2022 by the Ministry of Electronics and Information Technology [MEITY], proposed data collected, generated, and stored by government agencies available for licensing to private entities (Das, 2022). Indian Railway Catering and Tourism Corporation proposed monetising customer data but scrapped the proposal following criticism over the lack of privacy regulation (Mukopadhyay, 2022). The Ministry of Road Transport made ₹100 crore by licensing vehicle registration data to private companies before removing the scheme in 2019 due to privacy concerns (Saluja, 2021). In Chen's (2021) analysis, the Indian data governance framework has more enablers than safeguards. Enablers are the norms and laws that facilitate the use and reuse of data, whereas safeguards "protect the rights of individuals and entities participating in the data economy". Therefore, the current Indian framework adopts the 'Data as Resource' mindset that is common among the standard-setting data governance models.

Paradoxically, India should both explicitly speak against data colonialism and also adopt the data as a resource mindset. For the Indian government and tech industry, data colonialism is defined by foreign, often Western, big tech companies using data from people in emerging economies to enrich themselves (Basu et al., 2019). Data colonialism is cast as a question of geopolitics wherein the West, yet again, colonises the Global South. Couldry & Mejias (2018) argue for a more expansive definition of data colonialism in terms of economics. They postulate that the hegemony of big tech companies, foreign or not, in using, storing, and profiting from the data extracted from citizens is colonisation. The underlying rationale enabling this hegemony is to view data as a resource that is simply 'out there' to be extracted and exploited. It evokes the terra nullius principle used by colonists, which said that land owned by no one could be legitimately occupied and exploited by any state or company.

Furthermore, colonialism involved reconfiguring social relations in such a way that better enabled colonial extraction. Similarly, with data colonialism, social relations are remade so they can be "continuously tracked, captured, sorted and counted for value as 'data'" (ibid.). For instance, social relations are being re-defined by the increasing ubiquity of Aadhar, UPI, data-tracking devices, and

1 - Keeping data generated by the country in the country's territory.

data-driven content providers. Thus, in this approach, data colonialism is not only a question of geopolitics, where the West colonises the Global South, but also of economics, specifically extraction of people's data and concentration of power in controlling and profiting from people's data.

From this perspective, India's current approach perpetuates the economic aspect of data colonialism by retaining its underlying rationality and casting it as predominantly a geopolitical issue. To enact its ambition of data decolonisation, the Indian government must address the issue of technology companies having undue power over data collection and processing and making huge profits from it without the people receiving benefits in return. Mishra (2021b) writes that the Indian government is keen to protect domestic private companies and facilitate indigenous data-driven entrepreneurship. The government has close links with private companies like Jio and iSpirit in developing India's digital infrastructure, which have come under scrutiny for cronyism and mixing "public risk and private profit" (Thaker, 2018). Therefore, these relationships seem questionable when the government enters the data business as a market player with concentrated economic power in its hands.

DATA SOVEREIGNTY: WHOSE SOVEREIGNTY IS PRIORITISED?

There is room for government regulation in this regard as a guarantor of public interest. However, this cannot happen while the government itself seeks to facilitate the private sector, whether domestic or foreign, in making a profit from its people's data. The Indian government's outsized role in data governance is that in its championing of data sovereignty, 'sovereignty' is understood as operating first and foremost on the level of the state. For instance, calls for data localisation primarily emphasise rooted in state sovereignty. That is, any data generated within the state's territorial boundary is a national asset subject solely to the state's authority. However, in such a model of data sovereignty, it is assumed that since the government is theoretically the representative of the people, its pervasive role in managing data corresponds to the people having control over their data. Under the current framework, the government overwhelmingly controls citizens' data, but it is unclear to what extent citizens have a say in how the government uses their data. The Personal Data Protection Bill and the Non-Personal Data Framework may not be as successful in empowering individuals to make free economic or political choices regarding their data (Basu & Sherman, 2020).

When the government is framed as an economic and technocratic actor, it is further distanced from being transparent, accountable, and subject to electoral control. Rummens (2017) writes that technocratic governance networks, where public and private stakeholders coordinate 'efficient' market-oriented solutions, are insulated from democratic and parliamentary processes. They thus "lack visibility and appear in the eyes of ordinary citizens as an anonymous force over which they have no electoral control but which nevertheless profoundly shapes their lives".

Following Kovacs & Ranganathan (2019), we can argue that under India's current paradigm of data governance, government and corporations are deemed "the only actors that actually have the power and capacity to engage in the processes of data collection, storage, and analysis, while society is portrayed as somehow naturally benefiting from them, in a vein reminiscent of the 'civilisational' project that historic colonialism claimed to represent." As Sinha & Basu (2021) note, assertions of data sovereignty by national governments should not undermine the root of that sovereignty. They state that "greater control asserted by developing nations against the present (developed) ecosystem that hold the reins of power should not come at the cost of individual or community autonomy".

One of the moves India has made in this avenue is the NITI Aayog developed Data Empowerment and Protection Architecture [DEPA] deployed in the finance sector and piloted in the health sector. It is premised on the proposition that “individuals themselves are the best judges of the ‘right’ uses of their personal data, rather than competing institutional interests.” (NITI Aayog, 2020). DEPA creates the institution of a ‘Consent Manager’, which facilitates interactions between data subjects (who generate data), data fiduciaries (who store the data) and data users (who want to use the data). Besides being revocable and auditable, consent under DEPA is granular, meaning users can be specific about what data is shared, with whom, and for how long. However, DEPA is not framed as a digital public good but as something that will drive private market-led innovation based on public infrastructure (Reddy et al., 2020). For instance, consent managers are private third-party entities that provide their services as a business. It is also drafted amidst a lack of a comprehensive policy framework and a lack of integration of legal framework regarding privacy. Thus, while DEPA considers individual users, it still views data through an economic lens and as something that can drive private economic gain.

CONCLUSION: A NEED TO SHIFT METAPHORS AND EMPHASISE DATA'S SOCIAL ASPECTS

The underlying issue in India's framework remains that data is seen only as a resource. This view facilitates exploitation by private corporations, on the one hand, or propels the government to treat it purely as a national asset. However, the ‘data as resource’ metaphor ignores data's social characteristics. As Svensson & Guillen (2020) elaborate, data is never simply ‘given’. Highly networked societies generate data. It has to be captured, quantified, and processed. All of these practices give the data a particular shape, suggesting that “data is deeply cultural and infused with societal norms and values. Data does not naturally appear as it is collected and manipulated by people, shaped by human decisions, interpretations and filters” (ibid., p. 72). Personal data, information about a natural person that is identifiable, and non-personal data, such as weather information or anonymised health records, can both be characterised as social.

Prior to data becoming a corporate or national asset, it is first and foremost generated by people, by bodies situated within a social world. Data as a resource is a disembodied construction of data, which abstracts it away into an immaterial realm outside of an individual's grasp (Kovacs & Ranganathan, 2019). Sara Watson (n.d.) also notes that industrial metaphors, such as data being a resource, “share an experiential perspective of a bodiless conglomerate technocratic actor, seeing like a Google, as it were.” On the other hand, embodied metaphors connect data to the social world, to the bodies that generate it and that it has consequences for. Some examples of embodied metaphors are ‘data as fingerprints’ or ‘data as DNA.’

Another aspect of data governance that Indian policymakers must consider is that data links people to one another. Data is collected, processed, and used in a relational manner. It classifies, sorts, and categorises individuals in relation to one another (Viljoen, 2020). One type of data relation is between the data subject whose data is collected and the data collector/processor. Another is between data subjects, whereby data codifies relations based on group identity, social patterns, and preferences. Business and governance models will perpetuate social inequalities by viewing data in purely economic terms, only as a resource stripped of social or embodied dimensions. Datafication,

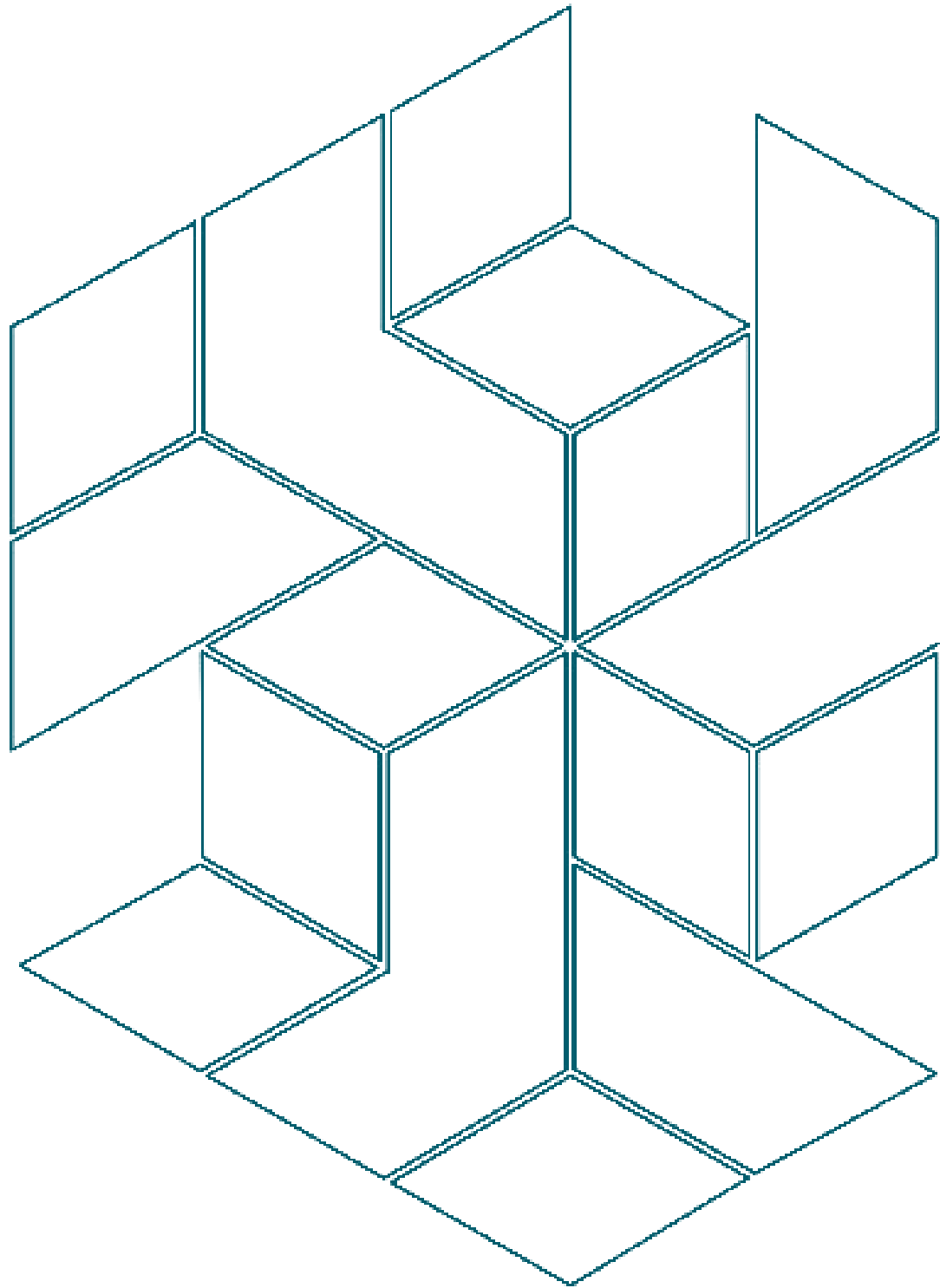
which is the production of data or how aspects of life are transformed into data, can create data relations that amplify and mirror social inequalities. Moreover, ignoring the social dimension of data also means foregoing the possibility of socially beneficial data production.

India can become an innovator and standard setter in the global data governance landscape by creating a robust decolonial framework, which departs from the 'data as resource' metaphor and reframes data metaphors to emphasise the social and embodied aspects of data. If India manages to create a decolonial data governance framework, it can become a template for other emerging economies concerned about data colonisation, data sovereignty, and better public welfare delivery. Carving out a niche for itself will also fulfil India's geopolitical ambitions and desire to cultivate strategic autonomy.

REFERENCES

- Arner, D. W., Castellano, G., & Selga, E. (2021). The Transnational Data Governance Problem (University of Hong Kong Faculty of Law Research Paper No. 2021/039). Berkeley Technology Law Journal, Forthcoming. <http://dx.doi.org/10.2139/ssrn.3912487>
- Basu, A., & Nachiappan, K. (2020). India and the Global Battle For Data Governance (Seminar Paper). India & Digital Worldmaking. <https://www.india-seminar.com/2020/731.htm>
- Basu, A., & Sherman, J. (2020, January 23). Key Global Takeaways From India's Revised Personal Data Protection Bill. Lawfare. <https://www.lawfareblog.com/key-global-takeaways-indias-revised-personal-data-protection-bill>
- Ciuriak, D. (2022). Unfree Flow with No Trust: The Implications of Geoeconomics and Geopolitics for Data and Digital Trade. Centre for International Governance Innovation. <http://dx.doi.org/10.2139/ssrn.3963074>
- Das, A. (2022, February 23). Explained: India plans to monetise public data. What is the draft data policy all about?. CNBCTV18 <https://www.cnbctv18.com/technology/explained-india-plans-to-monetise-public-data-what-is-the-draft-data-policy-all-about-12594022.htm>
- O'Hara, K., & Hall, W. (2018). Four internets: the geopolitics of digital governance (CIGI Papers, 206). The Centre for International Governance Innovation, University of Southampton Institutional Repository. <https://eprints.soton.ac.uk/427838/>
- Chen, R. (2021). Mapping Data Governance Legal Frameworks Around the World: Findings from the Global Data Regulation Diagnostic (Policy Research Working Paper No. 9615). World Bank. <https://openknowledge.worldbank.org/handle/10986/35410>
- Viljoen, S. (2021). A relational theory of data governance. Yale Law Journal, (131), 573–654. https://www.yalelawjournal.org/pdf/131.2_Viljoen_1n12myx5.pdf
- Mishra, N. (2021b). Data Governance and Digital Trade in India: Losing Sight of the Forest for the Trees? (ANU College of Law Research Paper No. 21.16). Forthcoming Chapter in Data Sovereignty along the Digital Silk Road. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3835497
- NITI Aayog. (2020). Data Empowerment And Protection Architecture Draft for Discussion. <https://www.niti.gov.in/sites/default/files/2020-09/DEPA-Book.pdf>
- Rummens, S. (2017). Populism as a threat to liberal democracy. In P. Ochoa Espejo, P. Ostiguy, C. Rovira Kaltwasser, & P. Taggart (eds.) The Oxford Handbook of Populism (pp. 554–570). Oxford University Press.
- Kovacs, A., & Ranganathan, N. (2019). Data sovereignty, of whom? Limits and suitability of sovereignty frameworks for data in India (Working Paper 03). Data Governance Network. https://datagovernance.org/files/research/IDP_-_Data_sovereignty_-_Paper_3.pdf

- Mukopadhyay, S. (2022, August 27). IRCTC withdraws tender to monetise customer data over privacy concerns. Mint <https://www.livemint.com/news/india/irctc-withdraws-tender-to-monetise-customer-data-over-privacy-concerns-11661557821781.html>
- Watson, S. (n.d.). Data is the New “_____”. DIS Magazine. <http://dismagazine.com/discussion/73298/sara-m-watson-metaphors-of-big-data/>
- Sinha, A., & Basu, A., (2021, August 13). Why Metaphors for Data Matter. Bot Populi. <https://botpopuli.net/why-metaphors-for-data-matter/>
- Puschmann, C., & Burgess, J. (2014). Big data, big questions: Metaphors of big data. *International Journal of Communication*, (8), 1690–1709. <https://ijoc.org/index.php/ijoc/article/view/2169/1162>
- Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data’s relation to the contemporary subject. *Television & New Media*, 20(4), 336–349.
- Reddy, S., Bedi, P., Sinha, A., & Mohandas, S. (2020). The Centre for Internet and Society’s comments and recommendations to the: Data Empowerment and Protection Architecture. The Centre for Internet and Society. <https://cis-india.org/depacomments1>
- Birch, K., Cochrane, D. T., & Ward, C. (2021). Data as asset? The measurement, governance, and valuation of digital personal data by Big Tech. *Big Data & Society*, 8(1). <https://doi.org/10.1177/20539517211017308>
- Basu, A., Hickok, E., & Chawla, A. S. (2019, March 19). The Localisation Gambit. Centre for Internet and Society. <https://cis-india.org/internet-governance/blog/the-localisation-gambit-unpacking-policy-moves-for-the-sovereign-control-of-data-in-india>
- The Economist. (2017, May 6). Data is giving rise to a new economy. <https://www.economist.com/briefing/2017/05/06/data-is-giving-rise-to-a-new-economy>
- Thaker, A. (2018, May 1). The New Oil: Aadhar’s mixing of public risk and private profit. The Caravan. <https://caravanmagazine.in/reportage/aadhaar-mixing-public-risk-private-profit>
- Saluja, N. (2021, February 11). Govt earned over Rs 100 crores by sharing vehicle registration data with private entities: Gadkari. *Economic Times*. <https://m.economictimes.com/industry/auto/auto-news/govt-earned-over-rs-100-crores-by-sharing-vehicle-registration-data-with-private-entities-gadkari/articleshow/80862992.cms>
- Svensson, J., & Poveda Guillen, O. (2020). What is Data and What Can it be Used For?: Key Questions in the Age of Burgeoning Data-essentialism. *Journal of Digital Social Research*, 2(3), 65–83.



SPRF.IN