The Indian state has tried to project an image of data sovereignty, seeing data in terms of a national asset that needs to be protected against the data colonialism of Western Big Tech through various policy documents and rhetoric. In this paper, I have tried to analyse this claim and tried to unravel how it sees data and defines data colonialism, unravelling the role which the Indian private sector plays in supporting that vision. I have then compared it with the work of the scholars of critical data studies who have critically examined the impact of Big Data from the lens of data colonialism to argue how it fails to challenge the epistemological basis of data colonialism even as it projects to fight it.

There is a particular way in which the Indian state sees data. It is driven by an idea of Data Sovereignty that has its basis in a much older idea of State Sovereignty, where the state has an authority over a territory. According to DB Hollis, this territory should not only be restricted to landmass but also to resources in the bounded space such as human infrastructures, air space, or minerals (or oil) below the surface or in its adjacent sea.

Various government policy documents by the Indian state use the rhetoric of data as an asset that needs to be used for the public good. Data are imagined as a national resource that should only be accessible to Indians, comparing it to a resource like ‘coal’ upon which ‘non-Indians do not have any prima-facie right or claim’. The peculiar thing about this rhetoric is its focus on the Indian private sector when it comes to accessing this ‘national asset’. Here the Indian state seeks to disrupt the ‘free flow of data’ built to protect the ‘Silicon Valley Consensus’ that favours US tech companies. The rationale behind this seems to be that these so called ‘Big Tech’ corporations based outside the country (mainly the US) have had a ‘first mover advantage’ in terms of profiting from the ‘data that would be generated within India’ – leaving the Indian private sector with no opportunity for ‘creating high value digital products’. Therefore, through various laws and policy documents, the Indian state seeks to give the Indian private sector the priority over the monetisation of this data.

This approach towards data decolonisation seems to be at odds with the larger decolonial discourse in critical technology and data studies: in the latter, the very idea of data as an ‘asset’ is critiqued and is seen as the basis of data colonialism. Scholars who form part of this discourse see the trend of ‘datafication’ increasing in various

---

6 Government of India (n 3) 15.
7 Kak (n 4).
8 While the Indian government does not explicitly use the term ‘decolonisation’, its rhetoric of data sovereignty (which I will explicate in the paper later), does seek to oppose the colonisation of data by Big Tech companies, which I will demonstrate both through policy documents and speeches by key government officials.
aspects of life and the consequent need to extract and modify data as a new form of colonialism akin to historical forms of colonialism. Their definition of decolonisation (something which I will elaborate on in the following section) critiques the idea of ‘data as oil’ which, according to them, forms the epistemological basis of data colonialism. But the Indian state has a very different understanding of data colonialism, which I would argue is a diametrically opposite definition of data colonialism, for it is based on control over data as an asset by Indians, and by extension, the private sector based in India. Why does the Indian state define decolonisation in terms of the narrow interests of the Indian industry? Why does the Indian state conflate domestic business interests in data access with national interest or public good?

In order to answer these questions, the first section of the paper will explore the underlying colonial ideology that sees data as an asset or a resource and what it means to decolonise by exploring the literature involving datafication and Big Data from the perspective of data colonialism. The second section will look at the concept of data sovereignty and the way in which the Indian state projects it as well as its decolonial effort through various policy documents and frameworks related to privacy of data, access to data held by companies and state actors and data flows, to see how the Indian state sees data and how it uses the term data colonialism. The third section explores the centrality of the Indian private sector in this understanding of data colonialism through the Aadhaar and the India Stack project. Using speeches by government and private actors, the next section explores the Indian state’s belief in the local private sector as the champion for India’s digital economy. The fourth section explores the centrality of the digital infrastructure for extracting and processing user data. To make it useful is seen as an expression of the ‘coloniality of power’.

Unlike historical colonialism, however, in data colonialism there is not only an exploitation of land, bodies, and natural resources, but an appropriation of all areas of social activity for profit. And unlike the former, data colonialism is not restricted by geography as it works both externally on a global scale, as well as internally. This ‘annexation of life to capital’ is done with the help of digital platforms provided by Big Tech corporations. While railways and maritime trade served as the ‘technical architecture of profit and plunder’ for historical colonialism, today Big Tech corporations monopolise the digital infrastructure for extracting and processing user data.

For instance, Microsoft’s operating software Windows became a monopoly in the operating system market, with their Office software is that this form of ‘data colonialism’ has many features in common with historical colonialism.

Data colonialism mirrors the ‘predatory extractive practices of historical colonialism’. Just as historical colonialism seeks to extract and process raw resources that are lying out there in nature like timber, oil, or coal for it to be made useful, in data colonialism, data are seen as an entity that needs to be extracted and processed for it to be made useful. According to Kovacs and Ranganathan, the idea that data are something that are out there has its origins in cybernetics wherein data are considered to be a layer that is everywhere as well as something that has its own independent existence. Moreover, this independently existing entity has been accorded so much power that it has become the ‘ultimate truth teller’.

It is also based on the doctrine of terra nullius (or no man’s land) which was used as a legal fiction to justify historical colonialism, as it signified that the land is ‘unowned’ and therefore available for occupation due to ‘the absence of prior claims’. Similarly, data are considered to be the new ‘unexplored frontier’ or terra nullius in data colonialism; one that needs to be ‘to be staked out, mapped, and colonized’. Thus, the very logic of data production which sees data as an asset or a resource that needs to be extracted and processed to make it useful is seen as an expression of the ‘coloniality of power’.

1. Data is the new oil: Big Data as a new form of colonialisation

The emergence of Big Data or ‘datafication’ in various spheres of life is increasingly being seen by many scholars as the newest manifestation of colonialism. The argument put forward by these scholars


11 Ricartce (n 10).


13 Coudry and Mejias (n 10) 2.


15 Kovacs and Ranganathan (n 14) 3.


17 Cohen (n 16) 216.


19 Thatcher, Sullivan and Mahmoudi (n 12); Couldy and Mejias (n 10).

20 Couldy and Mejias, ‘Data Colonialism: Rethinking Big Data’s Relation To The Contemporary Subject’ (n 10); Ricartce (n 10); Mann and Daly (n 12).

21 Couldy and Mejias, ‘Data Colonialism: Rethinking Big Data’s Relation To The Contemporary Subject’ (n 10) 6.

becoming an irreplaceable part of any office in the world. As a search engine, Google became a monopoly to such an extent that ‘Googling’ became synonymous with performing a search on the internet. Facebook (which later acquired Instagram and WhatsApp as well) had a similar hold over social media. Along with Twitter (now known as X), it became a source for accessing news around the world for many people, especially the younger generation. These platforms create a different type of ‘social’ in the form of ‘data’ that can be tracked, captured, sorted and counted for capital appropriation, as [it seems] life itself can now be caught in the ‘net of marketization’.45

Historical colonialism established a ‘eurocentric’ worldview that sought to impose ‘a single universalizing narrative of values, beliefs and politics’. Similarly, data colonialism legitimises a ‘regime of knowledge production’ that is considered universal and that excludes alternative forms of ‘data epistemologies that are respectful of populations, cultural diversity, and environments’. Here, a ‘Data Universalism’ based on certain ‘Silicon Valley principles’ is imposed which undercuts alternate forms of understanding data emerging from the Global South. It also homogenises different contexts that have their own particular characteristics and ‘cultural specificity’.39

Therefore, to address data colonialism, the authors suggest ways to think outside of the data driven epistemology. Decolonisation in this sense would mean rejecting the idea that only quantitative methods that are based on Big Data methodologies will offer an accurate picture of reality. It also means conceptualising data outside of its extractive rationality. Milan and Trére call for an epistemological shift when it comes to thinking about data, a call for not just epistemological diversity, but an ‘epistemic justice’ that not only seeks to give voice to the ‘otherness’ but goes further, taking ‘affirmative action against inequality by ‘unleashing novel data imaginaries’. These new forms of data imaginaries could be found in ‘bottom-up data practices’ that form a resistance against datafication based upon oppression. For instance, one can look at how indigenous people define data sovereignty that is different from the way corporations and national governments define it.44 Ricaurte pushes against a sense of ‘ineluctability’ about datafication which ‘normalises’ exploitation of data to such an extent that we cannot think beyond that.

But rather than fighting data colonialism in this manner, many countries in the Global South like India often try to oppose the data colonialism of the West (especially American Big Tech companies) by asserting sovereignty over the data, which they see as being ‘of’ and ‘produced by’ the people within their sovereign territory. As Mihir Kaulgud in his work on data governance in India argues, the policy stance of the Indian state paradoxically speaks against data colonialism and also seeks to perpetuate the data as a resource mindset restricting its meaning to only geopolitics. In the next section, I will explore the concept of data sovereignty in general and how the Indian state manifests its stance on this through various laws and policy documents.

2. Sovereignty over data: Looking at data from the lens of the Indian state

At the Economic Times Startup Awards 2019, the then telecom, law, and IT (Information Technology) minister Ravi Shankar Prasad put it thus: ‘Let me make it very clear: Narendra Modi will never compromise on data sovereignty of India’. This ideal of Data Sovereignty and the need to fight Big Tech is not unique to the Indian state. It has its origins in the Peace of Westphalia Treaties, wherein sovereignty is linked to a particular territory. According to Polatin-Reuben and Wright, it has become a ‘catch-all term’ to describe the data approach of vari-

26 Couldry and Mejias, ‘Data Colonialism: Rethinking Big Data’s Relation to the Contemporary Subject’ (n 10) 6.
27 Couldry and Mejias The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism (n 10) 69.
28 Ricaurte (n 10) 1.
29 Milan and Trére (n 12) 324.
30 Ricaurte (n 10) 324.
31 Couldry and Mejias, ‘Data Colonialism: Rethinking Big Data’s Relation to the Contemporary Subject’ (n 10); Ricaurte (n 10).
32 Milan and Trére (n 12) 325.
33 Ricaurte (n 10) 328.
ous countries wherein their governments want ‘to subject data flows to national jurisdictions’. For instance, China (which already did not allow tech giants like Twitter (now known as X), Google, and Facebook (now Meta) to operate in the country) has recently started cracking down on its own homegrown Big Tech. It also sees data as a resource for economic development, as can been seen from a series of Big Data White Papers published by a think tank that was quite influential under the Ministry of Industry and Information Technology of the Chinese government, as well as the Big Data Strategy (2017) and the 15th Five Year Plan (2016–2020) of the Communist Party of China (CPC). While the European Union (EU) has always foregrounded data governance in terms of fundamental rights, especially right to privacy, lately it has also started making more of economic rationale in its policymaking, focussing on increasing the competitiveness of the data economy of the European region. It has therefore started putting regulatory pressure on Big Tech companies like Amazon, Apple, Google, and Meta, on matters of data privacy, online safety, free speech, and anti-competitive practices. In December, 2020, it also proposed a Digital Markets Act, whereby these companies will be charged hefty fines if they continue with their monopolistic practices. They have also proposed the Digital Services Act, whereby companies must pay a fine if they do not comply with the content moderation rules laid down in the Act. The uniqueness of the Indian state, however, comes from the trust that it seems to project in domestic industries to compete with the Data Colonialism of western companies like Google.

The rhetoric of data sovereignty and the need to fight data colonialism has been articulated in various government policy documents. All of them see data as an asset which can be used for public good. While addressing an Indian-American audience at NRG stadium in Houston along with then President of the United States Donald Trump, PM Modi talked about how

Data is the new gold... The focus of Industry 4.0 is totally on data. It is India where data prices are the cheapest in the world and 1 GB data is available for 25-30 cents... Low-cost data is becoming the new identity of digital India, and is also redefining governance in the country.

The Economic Survey of India 2018–19 even goes so far as to say how ‘the private sector may be granted access to select databases for commercial use. Consistent with the notion of data as a public good, there is no reason to preclude commercial use of this data for profit’. Even the Department for the Promotion of Industry and Internal Trade’s Draft E-commerce policy talks about how ‘Indiain and its citizens have a sovereign right to their data. This right cannot be extended to non-Indians (the same way that non-Indians do not have any prima-facie right or claim to, say, an Indian coal mine).’ Since it is like an ‘Indian coal mine,’ it suggests that the Indian government should impose restrictions on any cross-border flow of data. One of the reasons behind this move is that even though Indian citizens constitute the largest proportion of Facebook users (thereby generating the most amount of data), their data is stored in data centres in Europe and North America.

This has been seen as a new form of colonialism, a digital or data colonialism wherein the data of developing countries is being used to fuel the growth of Big Tech companies based in the developed world. In his work on ‘digital colonialism’, Michael Kwet specifically talked about the ‘imperial control’ of the United States over the global South through the monopolistic practices of Big Tech firms like Uber, Google and Meta and other educational technology platforms.

Thus, in order to protect a ‘resource’ like data from being exploited by these American Big Tech firms, the Indian state believes it needs to establish sovereignty over this resource through various policy documents and legislations like the data localisation policy. Data Sovereignty is seen as a mechanism through which one can fight against the new form of data colonialism that seeks to control this all-important ‘resource’ of the modern times just as historical colonialism exploited other resources like wood, coal, oil, gold, etc.

The government has also reiterated this stance of Data Sovereignty at global forums such as the G-20 Summit held in Osaka in 2019, where India did not sign the ‘Osaka Track’, an international declaration on data flows pushed by Japan and endorsed by the US and European Union for free flow of data across borders. Accessibility of data to...
Indian actors (companies, individuals, and the government) remains the primary strategic interest of the Indian state.\textsuperscript{57}

Interestingly, when it comes to the laws regulating data in the country, the government has taken a backseat when it comes to the mandates restricting cross-border data flows. When the \textit{Personal Data Protection Bill (2019)} was first introduced in the Parliament, it mandated the storage of ‘sensitive personal data’ within India.\textsuperscript{58} In the bill, the government could notify any data as ‘critical personal data’ which must be stored and processed only in India.\textsuperscript{59} When the bill was referred to in the Joint Parliamentary Committee (JPC) of the Parliament, this localisation mandate was defended on four fronts:\textsuperscript{60}:

1. In order to achieve national security objectives, wherein a hostile country can use data for surveillance; to manipulate consumer opinion or behaviour; as well as to enforce law, as law enforcement agencies need timely access to data.

2. To provide better informational privacy using data protection regulations within the country.

3. To boost employment, as data centres and other associated industries can generate massive employment opportunities.

4. In order to gain bargaining power with other countries due to a strong internet presence and massive amounts of data at its disposal, bolstering the digital economy.

But this bill was withdrawn by the government in August 2022 and the draft of a new bill, \textit{Digital Personal Data Protection Bill (2022)} was put up for public consultation in November of the same year and is due to be tabled in the Parliament later in 2023.\textsuperscript{61} This new bill did away with specific data localisation mandates that restricted cross-border data flows of sensitive personal data and critical personal data.\textsuperscript{62} It allowed the data fiduciary to transfer data outside India ‘in accordance with such terms and conditions as may be specified’ by the central government.\textsuperscript{63} Some reports have talked about how this was done after backlash from the industry and how it is a move to simplify cross-border data flows ‘to promote ease of doing business’.\textsuperscript{64} Interestingly, the process of determining the countries to which the data will be transferred is not clear, with some reports suggesting that it would be based on trust and ‘reciprocity’ with India.\textsuperscript{65} When the draft of this new bill, \textit{Digital Personal Data Protection Bill (2023)} was tabled before the parliament in August 2023 after this consultation, instead of providing a list of countries where data can be transferred, they would notify a blocklist of countries to which they can ‘restrict the transfer of personal data by a Data Fiduciary for processing’.\textsuperscript{66}

Despite taking a step back in terms of the localisation mandates of the data protection bill; based on some of the other policy documents and public proclamations by its own ministers it seems like the Indian state does seek to play an active role in helping domestic startups and companies with the infrastructure to overcome the ‘first mover’s advantage’ to give it the possibility to create ‘high value digital products’.\textsuperscript{67} This is something which I will touch upon in the next section. The presumption here is that these companies being from India represent national interest, and should therefore be treated differently. Data colonialism is seen purely as a new manifestation of historical colonialism. Whereas in the latter it was the East India Company that represented British colonialism, in the former, it is the American Big Tech firms like Amazon that represent the new form of data colonialism. Data colonialism as an idea is not challenged here. In fact, its exploitative nature is being embraced and even nurtured as a tool to ironically fight it. The only thing that is being challenged in the name of data colonialism is its origin in the western countries like the United States. Yet it is not clear how companies based in India would necessarily serve the local interest and use the same business models for which the Big Tech companies have been criticised.\textsuperscript{68} In fact, a company that is the symbol of India’s fight against data colonialism, Reliance (partly owned by Big Tech like Meta), will likely dominate the digital commerce space as Amazon has done with online marketplaces.\textsuperscript{69} This is something that I will discuss in a later section on the specific form of techno-nationalism in India.

In the next section I will look at the centrality of the Indian private sector in projecting the image of decolonisation by the Indian state.

---

\textsuperscript{57} Basu (n 56).
\textsuperscript{58} Personal Data Protection Bill (2019).
\textsuperscript{59} Government of India (n 3) 15.
\textsuperscript{63} Digital Personal Data Protection Bill (2022).
3. The role of the private in creating public goods: The centrality of the private sector in the decolonisation effort of the Indian state

The Indian state seeks to project a uniqueness about the way it approached data governance, differentiating itself from the laissez-faire approach of the United States which led to a monopoly of the Big Tech companies as well as the authoritarian approach of the Chinese state, which is considered undemocratic. Certain key private actors play an outsized role in projecting this policy and the rhetoric around it. One of the key actors who plays this role is Nandan Nilekani, one of the principal architects of the Aadhaar project. He talks about the uniqueness of this Indian approach by highlighting how it has created ‘open digital platforms from scratch and tailored them to the Indian context’ and how ‘instead of leaving them in the hands of a few private technology companies (as in the case of the West), the Indian government has built these systems as public goods’.

The principal public good being referred to by Nilekani is this Aadhaar project. It was aimed at providing a basic identity to Indian citizens and making India’s welfare system more efficient. But, a digital infrastructure or a digital public good was built around it with the help of certain private actors, making it a huge part of India’s answer to data colonialism. An effort has been made to portray this as a unique path taken by India, which is distinct from the western and Chinese approach, where powerful entities and players from the Indian private sector play an outsized role in its formation and deployment – even as it portrays itself to be a public good.

A closer look at the Aadhaar project reveals how a few private companies and individuals played a huge role in its formation. The Unique Identification Authority of India (UIDAI), the authority responsible for giving Unique Identification Numbers (UID) to residents in India, was initially headed by Nandan Nilekani, who has held various positions in the company Infosys, including the post of CEO and co-chairman. He was the chairperson, and according Usha Ramanathan, a human rights activist and an expert on law and poverty, was also given the status, protocol and privileges of a minister, without having to meet the constitutional requirement that a minister has to be a member of Parliament... This device, by which he was given the rank of cabinet minister without the constraints of the position, was used to facilitate lateral induction of corporate leadership into the government.

Thus, many people from the corporate space came over to build the entire system. This included ‘former Infosys employee, Pramod Varma, [who] became Aadhaar’s chief architect; Viral Shah, a PhD in Computer Science and co-inventor of Julia programming language; Shankar Maruwada, CEO of Marketics; Sanjay Swamy, CEO of mCheck among others’. These were the domain and technology experts that were necessary to build such a system which required the necessary expertise to deal with big data analytics.

Apart from this, initially even its status as a public authority was very much questionable. Although UIDAI was initially constituted under the Planning Commission in 2009 via a gazette notification, it was until recently only a dabbious public... quasi-government authority operating without a legal framework. In 2016, it was only after the passing of the Aadhaar Act (2016) that it became a statutory authority and a government department under the Ministry of Electronics and Information Technology. Moreover, even the deployment of the Aadhaar project is dependent on the private entities as for the enrolment process, UIDAI appoints state government and private registrars and they, in turn, appointed private enrolment agencies to enrol people in the programme.

Similarly, another public good - India Stack, ‘a collection of systems and standards’ is going further by giving a number of services to everyone with an Aadhaar number. These services include storage for electronic documents (Digi-Locker), electronic signatures (e-Sign), Unified Payments Interface (instant real-time payment service) and e-KYC (know your customer). It is a platform that allows various third parties (businesses, nonprofits, and state organisations) to create services using Easy-GST and ClearTax, which help out with fulfilling tax obligations and submitting information to the tax authorities on time and in the required formats; or organisation like FRS Labs or BioEnable, which provide document verification. India Stack is created by an organisation iSPIRT, which is a non-profit organisation that supports India’s digital start-ups. It is an intermediary for private businesses who want to start their services using the India Stack platform. Many of its ‘volunteers’ (as iSPIRT describes them on its website) were with the UIDAI – like Pramod Verma, its chief system-architect and technology advisor since it started, Sanjay Jain (UIDAI’s chief product manager for two years), and Vivek Raghavan (who is also currently the Chief Product Manager and Biometric Architect at UIDAI). This ‘revolving door’ policy has been criticised for its lack of transparency by international organisations and academics.


77 Hicks (n 59).

78 Hicks (n 59).

watchdog agencies like the Privacy International, which, in its report on financial technology in India, said that:

Having India Stack as a product produced by a group of ‘volunteers’ – rather than, say, within the UIDAI – has certain advantages from their point of view: they do not have to operate transparently, there is no requirement for them to be subject to right to information legislation or procurement rules. Thus, this important initiative – potentially as important as anything coming from government ministries – is not subject to that degree of oversight.80

Nilekani also had a hand in the appointment of the head of the National Payments Corporation of India (NPCI), a non-profit created by the Reserve Bank of India and the Indian Banks’ Association, which owns the Unified Payments Interface (UPI), its instant payment system.81

This huge intermingling of state and private actors is portrayed as a public good, where the state is projected as the provider of a digital infrastructure upon which private players can create new services. It is shown in contrast to the power of the Big Tech giants like Google and Meta who monopolise the digital infrastructure.82 But as we saw earlier, private companies and individuals still play a huge role in the very foundation and functioning of these public goods. These private stakeholders are also responsible for building governance frameworks around data which are endorsed and implemented by the state, like the Data Empowerment And Protection Architecture, 2020 paper by NITI Aayog, which acknowledges the involvement of iSPIRT.83

Jacqueline Hicks has called this development in India ‘Digital ID Capitalism’, thus separating it from other western forms of digital capitalism, which is led by the private sector.84 Here, the state not only regulates data-driven enterprises, but creates the infrastructure for the domestic data-driven industries. She argues that ‘Digital ID Capitalism’ is a form of post-colonial capitalism that occurs in the Global South. Here, the state justifies its support of the domestic industries as it believes that the West continues to be the beneficiary of its colonial past.85 Consequently, in the case of the Aadhaar as well as India Stack project, the Indian private actors are seen by the state as crucial players in the building of digital public goods. They are considered necessary to oppose the colonisation by the western Big Tech by creating alternative systems that reduce the influence of services that they provide. In this manner, the rhetoric of data sovereignty of the Indian state cannot be decoupled from the domestic private industry interests. In the following section, I will talk about some of the implications of this rhetoric, and how the private sector in India plays up this rhetoric and tries to benefit from it.

4. Techno-nationalism: Indian state’s belief in the local private sector as the champion for fighting the data colonialism of the west

On 3 October 2021, Panchjanya, a Rashtriya Swayamsevak Sangha (RSS)86 linked weekly featured the following cover story with the face of Amazon Chief Executive Officer Jeff Bezos:

Whatever the East India Company did in the 18th century to capture India, the same is visible in the activities of Amazon.87

Along similar lines, while speaking at the Republic Summit, Mukesh Ambani, Chairman of one of the biggest conglomerates in India, Reliance Industries (which operates jio infocomm) that has 54.1% market share in broadband services, echoed the same rhetoric as the Indian state:

In this world, data is the new oil and data is the new wealth. India’s data must be owned by the people and especially not by global corporations. Data colonisation is as bad as the older colonisation and digital freedom is as important as the freedom we won.88

Why is the head of one of the most powerful corporations in India and the richest man in the country talking about data colonialism? Why is he furthering the rhetoric of data sovereignty being propagated by the Indian state? Answers to these questions can be found as a new wave of techno-nationalism has taken hold over not just the state, but the private sector as well. It can also be seen in parallel with the rise of ‘corporate nationalism’ where the ‘foreign’ or ‘domestic’ identity of a company along with a sentiment of ‘domestic self-reliance or sharing the fruits of capital locally’ plays a dominant role.89 This can be seen in the case of the digital dispute between Reliance and Amazon over the sale of assets by the Future Group to Reliance (their main rival), which Amazon says is illegal, considering it had an agreement with the Future Group in 2019 that gave them ‘special rights over Future’s retail assets’.90 Future Group, while arguing that it has not done anything wrong, accused Amazon of not just illegally exerting control, but also called it ‘the East India Company of the 21st century’ and ‘Big Brother in America’.91 Another

82 Nilekani (n 70).
83 Kak and Sacks (n 41).
84 Hicks (n 71).
85 Hicks (n 71).
91 Singh and Khemka (n 89); Kalra (n 90).
instance of this is the case of WhatsApp (owned by Meta), which, despite claiming to have all the approvals, faced many issues in starting its payment service: many filed petitions saying that allowing a ‘foreign entity’ to start its payment service would endanger India’s financial data.93 Similarly, BharatPe, a fintech company which offers a UPI based payments app, was accused of tagging its competitors as ‘Chinese’ or ‘American’ by circulating pamphlets highlighting the significant ownership of these companies by American and Chinese firms.94 The ‘foreign-ness’ of a company seems to overshadow every other form of legal assessment.95

Pratap Bhanu Mehta called this a form of techno-nationalism, where technology becomes a tool of national power, and big companies like Reliance are believed to be carriers of this power.95 He argues that a company like Reliance becomes a symbol, a national icon (versus a demonetized Rs 500 and Rs 1,000 notes post-midnight.98 Interest-in the financial history of independent India!' His image was then put app, congratulated the PM by saying ‘Paytm congratulates Honoura-

In the next section I will critically examine how the colonial mentality of extractivism seeps into the decolonisation rhetoric of the Indian state.

5. A data-centric regime: How the decolonisation of the Indian state remains within the psychological limits of data colonialism

While historical colonialism has arguably (defined strictly in terms of territorial occupation) largely ceased to exist, its political, social, cultural, and economic legacy continues to haunt the colonised even today with data colonialism being its newest iteration.94 One of those legacies is that the state machinery in non-western formerly colonised countries continues to embrace those same colonial practices. The reason for this persistence is what Anibal Quijano called ‘a coloni-
sation of the imagination of the dominated’.95 Therefore, in order to understand its lasting impact today, it makes sense to look at how colonialism captures this ‘imagination of the dominated’.

Is there a reason why the exploitative nature of colonialism persists even after the coloniser has left? One answer can be found in the work of one of the most renowned critics of colonialism, Ashis Nandy, an Indian political psychologist who defines colonialism not simply as a political or an economic phenomenon, but also as a ‘psychological state’, which ‘may not always begin with the establish-
ment of alien rule in a society and end with the departure of the alien rulers from the colony’.96 He takes the example of India to show how colonialism began operating 75 years before it became dominant and continues to remain a dominant ideology in many sectors even 35 years (at that time when his book was published) after the British left India.97

92 Singh and Khemka (n 89).
94 Singh and Khemka (n 89).
96 Mehta (n 95).
99 Block (n 97).
100 Block (n 69).
101 Block (n 69).
104 Couldry and Mejias ‘The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism’ (n 10).
105 Quijano, ‘COLONIALITY AND MODERNITY/RATIONALITY’ (n 18) 169; Ricaurte (n 10).
107 Nandy (n 106).
If one wants to understand how colonialism seeps into the mindset of the Indian state even when it is overtly trying to project a resistance to data colonisation by the western Big Tech, one should look at the particular way in which Nandy describes how dissent is managed in historical colonialism. According to Nandy, while there is a system of psychological rewards and punishments to ingrain colonial norms in the psyche of colonial subjects, it can be easily identified and challenged. What is more dangerous is the culture that is created wherein ‘the ruled are constantly tempted to fight their rulers within the psychological limits set by the latter’. Similarly in the case of the Indian state, this psychological limit to the way it imagines data governance is a ‘data driven rationality’, which, according to Ricaurte, is based on three assumptions: (1) data reflects reality, (2) data analysis generates the most valuable and accurate knowledge, and (3) the results of data processing can be used to make better decisions about the world.

The Indian state accepts this rationality as it believes that India has to harness data as a resource and control it by keeping it within its territory. This can be clearly seen from the localisation mandate in the Report of the Joint Parliamentary Committee (JPC) on the PDP Bill, 2019 where the report mentions how

Data is core to the future of our economy and is unlike any other resource. Data is now treated as an asset, deriving implicit value generated from insights, patterns and distribution of data and its amalgamation with other data. It is available nationally and internationally, providing an impetus to the economy and innovation.

As we can see, data has been given such an importance that the very future of the Indian economy is being portrayed as being dependent on the ability to derive value from it. Therefore, even when it tries to take on the data colonialism of the West, it fights within those epistemological limits of the ‘data-centric regime’ and does not reimagine a resistance outside of it. According to Ricaurte, the state in a non-western country also embraces this ‘dominant data epistemology’ and plays a systemic role in data colonisation by

1. developing legal frameworks
2. designing public policy,
3. using artificial intelligence systems for public administration,
4. hiring technological services,
5. acquiring products for public administration and surveillance purposes,
6. implementing public policies and digital agendas, and
7. facilitating education and the development of labour forces.

In order to pursue this agenda, the state is dependent on the private sector to supply them with the technology, as can be seen in the case of Aadhaar and India Stack mentioned in the earlier section. In fact, the importance of private sector over here is based on an ‘extractive’ and ‘practical rationality’, which deems only them as capable enough to process and appropriate data. At the same time, the society is considered a ‘natural beneficiary’ of this appropriation of data, just as the ‘civilizational project’ was supposed to benefit humanity in historical colonialism. This gives them immense power to shape this digital agenda.

Moreover, there is little to no participation or even consideration of the indigenous communities in this agenda. Therefore, the ambit of data sovereignty becomes restricted to a very small powerful elite including the state and other powerful private sector players.

6. Conclusion
While examining the data sovereignty approaches of different BRICS countries (Brazil, Russia, India, China and South Africa) in 2014 (the year Narendra Modi became India’s Prime Minister), Polatin-Reuben and Wright described India as having ‘a weak approach to data sovereignty’, arguing how India is seemingly more interested in ‘protecting the data of its Western investors’ rather than data that is ‘nationally generated’, citing the signing of a Cyber Pact with the UK, which aims to protect British data that is stored in Indian Data Centres. This suggests that there has been an intensification for the calls for a stronger approach to data sovereignty with the new government, which, through its policy documents and general rhetoric, is more willing to assert its sovereignty over data. This trend (as mentioned before as well) is not something particular to India, as many countries are now looking to assert sovereignty over the data generated in their own territories.

Through this paper, I have tried to analyse this approach and this rhetoric of data sovereignty as well as the sense of data colonisation that the Indian state is trying to portray through various laws and policy frameworks. I have tried to show the centrality of the private sector in this imagination of data sovereignty, both in terms of its formulation and implementation. It also takes part in furthering the rhetoric of data colonisation, which serves its interests and allows it to shape policy around access to data. I have further argued that the rhetoric of decolonisation followed by the Indian state falls within the same epistemological ambit of Big Data that has led to the rise of Western Big Tech that it wants to take on.

While looking at Big Data or Datafication through the lens of colonialism is useful given many similarities between these two processes, we must also critically engage with it, as we risk reviving old binaries such as identity politics vs. anti-capitalist struggles, structure vs. agency, etc. at best, and at worst, simple reversals like tradition is good and modernity is bad. Moreover, there is also a risk of creating its own ‘blind spots’ with terms like ‘Global South’ and ‘colonial’, where we miss the different narratives and experiences of various countries with their own issues. There can be other conceptual blind spots in the colonial perspective like race, gender, caste, etc., which can be overcome through intersectionality by engaging with other lenses like feminism or critical

108 Nandy (n 106) 3.
109 Ricaurte (n 10) 2.
110 Joint Committee On The Personal Data Protection Bill (n 60) 8.
111 Ricaurte (n 10) 3.
112 Ricaurte (n 10) 7.
113 Ricaurte (n 10) 8.
114 Couldry and Mejias, ‘Data Colonialism: Rethinking Big Data’s Relation to the Contemporary Subject’ (n 10) 5.
115 Couldry and Mejias, ‘Data Colonialism: Rethinking Big Data’s Relation to the Contemporary Subject’ (n 10); It also complements the earlier point on techno-nationalism in India where the interest of a private corporation (in this case Reliance) was seen as similar to the interest of the whole nation.
116 Ricaurte (n 10).
117 Kovacs and Ranganathan (n 14).
119 Polatin-Reuben and Wright (n 39).
120 Kiran Asher, ‘Latin American Decolonial Thought, Or Making The Subaltern Speak’ (2013) 3 Geography Compass https://doi.org/10.1011/gec3.12102 accessed on 12 August 2023; Milan and Treré (n 12).
121 Kak (n 4) 308.
race theory. This can help us in understanding how intersecting fields of race, class or gender also shape the political economy of Datafication or Big Data.124

Nevertheless, the decolonial approach towards data is useful in bringing to the fore an extractive rationality underpinning datafication, and identifying alternative visions of data that do not conform to this data-driven rationality. This is not easy, given the dominance of this narrative of data-driven rationality. As Nandy said in his work, colonialism restricts even the resistance of the colonised.125 Therefore, a radical reimagining of the concept of data sovereignty must be promoted.

We also need to ask who gets to define data sovereignty in the face of data colonialism? Is it the state, the citizens or the whole community? Is everyone in the community being given an equal opportunity to define sovereignty, including the marginalised groups within the society? As we have seen in this paper, since colonialism is an ideology which is no longer bound by binaries of Global South vs Global North, we should probably take inspiration from postcolonial theory that data sovereignty can potentially challenge the very ideological basis of data colonialism. For instance, Tairuru, Burch and Finlay-Smits talk about principles of Maori Data Sovereignty where, unlike the western colonial understanding of data as a commodity that needs to be extracted for enrichment or profit, data ‘refers to something treasured by a specific group of people which, therefore, requires guardianship and protection’.126 They refer to data as a taonga or treasure and data contributors therefore as the guardians of data; where no one entity can own big data. They focus on ‘where data originates, who gathers it, and how the wider community can benefit’, thereby protecting the priorities of data holders i.e., people supplying data.127 Interestingly, the Draft National E-Commerce Policy (2019) by the government mentioned before in the article does talk about the ‘Maori Data Sovereignty Network’ as an example of ‘Taking back control over data’. While it talks about how ‘Maori data should be subject to Maori governance and that Maori organisations should be able to access this data to support their development aspirations’, it fails to understand how the Maori data sovereignty principles are fundamentally against the kind of colonial resource commercialisation and extraction being pushed by the India state, as argued by Tairuru, Burch and Finlay-Smits.128

While these initiatives try to reclaim the idea of sovereignty from the exclusive realm of the state, there is also a need to problematize the idea of data sovereignty.129 Yarimar Bonilla does so by pointing out how the concept of sovereignty was used by the colonial powers in the first as ‘a legal technology with which to lay claim to putatively unowned lands’.130 Based on her argument, Couture and Toupin have argued for the need to question ‘the colonial histories or power dynamics that are being maintained or reproduced’ when social movements and indigenous struggles resort to the notion of sovereignty when thinking about data.131 Through this paper I have tried to look at the rhetoric of data sovereignty and the need to fight data colonisation as it is being promoted by the Indian state, and whether it actually opposes the ideology of data colonisation. Future research can highlight how one can uncover such an imagination of decolonisation and perhaps imagine decolonisation of data that can fundamentally oppose data colonisation to its very core. For instance, one of the ways in which we can possibly do this is by empowering the local communities through data practices at the grassroots level like the one being run by JalJeevika. It is a not-for-profit organisation based in India that works with farmers in the aquatic sector to build technological solutions from the ground up, providing farmers with location-specific advisory support, market intelligence, IVR-based support services for small farmers and linking them to agencies like banks and government programmes.132 In order to build these technologies, data are collected from farmers through community resource persons or agents who are selected from amongst the residents and trained to explain terms and conditions of the contract to them before they sign up for any government or private service.133

There is a further need to explore how state and other powerful actors within different countries in the Global South try to appropriate the concept of decolonisation. While Big Tech is rightly being opposed in many countries in the name of data colonisation, we also need to examine whether those same extractive rationalities are being deployed elsewhere by the same actors who claim to oppose it.

Copyright (c) 2024, Kushang Mishra.

Creative Commons License
This work is licensed under a Creative Commons Attribution-Non-Commercial-NoDerivatives 4.0 International License.

122 Milan and Treter (n 12).
123 Nandy (n 106).
125 Tairuru K, Burch K and Finlay-Smits S (n 124) 5.
126 Government of India (n 3) 4; Tairuru K, Burch K and Finlay-Smits S (n 124).
127 Couture and Toupin (n 1).
129 Couture and Toupin (n 1) 15.
131 Kapoor A, Mahesh S and Narayan V (n 130).