Data
Governance
Network

November 2019

Working Paper 03

Data sovereignty, of whom? Limits and suitability of sovereignty frameworks for data in India

Anja Kovacs and Nayantara Ranganathan

internet
democracy
project

Data Governance Network

The Data Governance Network is developing a multi-disciplinary community of researchers tackling India's next policy frontiers: data-enabled policymaking and the digital economy. At DGN, we work to cultivate and communicate research stemming from diverse viewpoints on market regulation, information privacy and digital rights. Our hope is to generate balanced and networked perspectives on data governance — thereby helping governments make smart policy choices which advance the empowerment and protection of individuals in today's data-rich environment.

About Us

The Internet Democracy Project works towards realising feminist visions of the digital in society, by exploring and addressing power imbalances in the areas of norms, governance and infrastructure in India and beyond.

Cover page credit: Cactus Communications

Paper design: Cactus Communications

Suggested Citation

Kovacs, A., Ranganathan, N. (2019). *Data sovereignty, of whom? Limits and suitability of sovereignty frameworks for data in India*. Data Governance Network Working Paper 03.

Supported by a grant from Omidyar Network

Abstract

Sovereignty is seeing renewed relevance in the age of data in India as it has become the framework of choice in a number of data governance proposals by the Indian government. To understand the scope, import and consequences of these reassertions of sovereignty, however, it is important to unpack the nature of these claims as they have been put forward. In particular, to what extent does this type of sovereignty allow for the exercise of autonomy and choice of the Indian people? This paper will demonstrate that such assessments crucially depend on how we construct the nature of data. In most dominant discourses, data is described as a resource of some sort. However, in practice the line between our physical bodies and our virtual bodies is increasingly becoming irrelevant: data, then, emerges not so much as a resource that is simply out there, but as an extension of our bodies. In order to benefit the people of India, assertions of sovereignty in the face of data colonialism will need to take these shifting realities regarding the nature of data into account. Through an assessment of policy proposals relating to sovereignty in the realm of data and new technologies, we seek to examine to what extent policy in India does indeed recognise these new realities, and what the value of these new assertions of sovereignty for the people of India consequently is.

Table of Contents

1. Introduction

> Constitutions like our own are means by which individuals – the Preambular 'people of India' – create 'the state', a new entity to serve their interests and be accountable to them, and transfer a part of their sovereignty to it.
>
> – Justice Chelameswar in Puttaswamy vs. Union of India[i]

Sovereignty, this enduring concept, is seeing renewed relevance in the age of data as it has become the framework of choice in a number of data governance proposals by the Indian government.

Such an emphasis needn't be surprising. As Nick Couldry and Ulises Mejias (2019) have pointed out, modern forms of hegemony of big tech companies can be usefully understood through the frame of data colonialism. For a post-colonial state like India, the notion of sovereignty emerges as a potent framework to resist these new forms of colonialism.

To understand the scope, import and consequences of these reassertions of sovereignty in the face of data colonisation in practice, however, it is important to unpack the nature of these claims as they have been put forward.

Who is constructed as the body containing this sovereignty? What are the accompanying policy prospects that current assertions of data sovereignty brings? And, in particular, to what extent does this type of sovereignty allow for the exercise of autonomy and choice of the people?

This paper will demonstrate that such assessments crucially depend on how we construct the nature of data. In most dominant discourses, data is described as a resource of some sort. However, such constructions often contradict people's experiences: as van der Ploeg (2012) has argued, in practice the line between our physical bodies and our virtual bodies is increasingly becoming irrelevant. Data, then, increasingly emerges not so much a resource that is simply out there, ready to be mined, but an extension of our bodies, even a part of it.

As the words of Justice Chelameswar that we started this paper with remind us, the people of India transferred part of their sovereignty to the state they created with the intention that this state would safeguard their interests. In order to benefit the people of India, assertions of sovereignty in the face of data colonialism will thus need to take these shifting realities regarding the nature of data into account.

In what follows, we first lay out the conceptual frameworks that animate this paper. We begin by tracing the concept of sovereignty and its historical entanglements with the realm of data and then go on to briefly lay out the debate on data colonialism within which calls for greater data sovereignty in India are often situated. We then lay out the theoretical framework of a feminist politics of data that we use to guide our analysis, focusing on the growing entanglements of data and bodies.

In the second part of the paper, we investigate how these different conceptual frameworks infuse and animate sovereignty claims in data related laws, policies and debates in India, and what light they can shed on the gains to be made for different stakeholders by such claims. We start by examining regulations and proposals around data localisation, so far one of the central policies associated with the state assertion of data sovereignty in the Indian context. We then unpack sovereignty claims to, following Cohen (2018), identify and deconstruct the legal constructs that enable such claims, as well as to assess their potential for people's empowerment; in particular, we examine discourses around the economic value of data and around ownership of data. Finally, we conclude with an understanding of whether the proposed claim of sovereignty is consistent with the promotion of the rights of citizens.

There is no singular articulation of data sovereignty by the Indian government. However, what are available are fragments and threads found in legal and policy documents and in statements made by government officials on sovereignty in the realm of data and new technologies. To evaluate the gains to be made from assertions of sovereignty over data as currently put forward in India, we therefore study these policy proposals on sovereignty in the realm of data and new technologies. In particular, we analyse (1) policy documents explicitly using the sovereignty framework (2) policy documents that develop legal constructs that are linked to the emergence of data sovereignty claims, like community data, and (3) statements by government officials on sovereignty. In these articulations, sometimes technology and data are seen as a means to secure existing sovereignty, and at other times sovereignty is seen as something to be asserted over a new and strategically important kind of resource, that of data. In this paper, we consider both kinds of assertions.

1.1. Sovereignty: past, present and futures

The legal concept of sovereignty dates back many centuries. Today's dominant conceptualisation of sovereignty as accruing to the state finds its origin in the Peace of Westphalia treaties, when a new political order was recognised, organised around the co-existence of sovereign states as supposed equals. This sovereignty extended over lands, people and agents. Couture and Toupin (2019: 4) note that in the *Stanford Encyclopedia of Philosophy*, Philpott identifies four 'ingredients' for the sovereign:

> 1) it possesses authority; 2) this authority is derived 'from some mutually acknowledged source of legitimacy' — which can be God, a constitution or a hereditary law; 3) this authority is supreme; and 4) this authority is over a territory.

With changing political systems, from the rule of kings to the post-war periods, and now to a global economy, the concept of sovereignty has evolved and has taken on new dimensions. Yet, the link of sovereignty with territory continues to play a crucial role in dominant articulations of sovereignty of the state in the realm of data today, whether with regards to its external or internal aspects.

The external aspect to such sovereignty concerns, for example, changes to data flows and efforts to subject foreign firms dealing with Indians' data to compulsorily store data within the country, which our analysis of data localisation in India later in this paper will discuss in detail. The internal aspect concerns legitimacy for the state to take decisions and ownership over the data of citizens; particular moves by the Indian state to shore up its legitimacy in this regard will be discussed in the final sections of this paper. Both aspects contribute to a reconfiguration of power in the assertion of state sovereignty.

1.2. Internet as the undoing of sovereignty?

There was a time when the enduring importance of state sovereignty was in question: one of the most captivating myths in the early days of the Internet's proliferation in the nineties was that the Internet was destabilising how we understand nation states and their propensity for governance.

Among the clearest articulations of this imagination was the Declaration of the Independence of Cyberspace (Barlow, 1996). Addressing governments, the Declaration claimed that outdated legal concepts of property, expression, identity, movement, and context have no place on the cyberspace:

> Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather.

But in hindsight, this vision of the Internet was naive. As the uses of the network have evolved and user bases have exponentially expanded, legal concepts of property, expression, identity, movement, and context have not only retained their force, but have often been strengthened in new ways, precisely because of the emergence of digital technology and its applications. These legal mechanisms and instruments are crucial means through which sovereignty in the digital age is operationalised, while sovereignty in turn is a key enabler of their enforcement.

For example, digitisation has altered profoundly not only how, but also when and how frequently identification takes place, as well as how identity is performed, imposed and negotiated, even if it can be negotiated at all (Sriraman, 2018). If in earlier times, identification was often an intentional act, biometric technologies such as facial recognition technologies now allow identification to happen covertly and without the (full) consent or understanding of the subjects involved. Moreover, the data collected from these identification systems is often designed to be shared with hardware suppliers, maintenance contractors and other private companies.

1.3. Sovereignty and the people

It is not just that the Internet has not meant the undoing of sovereignty, however; the concept of sovereignty has been adopted intentionally and incidentally by political technologists and activists working on reappropriation of technologies.

Tracing the history and meanings of alternative conceptions of sovereignty like food sovereignty, energy sovereignty, and body sovereignty, Couture and Toupin (2019) explain how sovereignty has found use in different movements to reclaim autonomy and self-determination over resources and one's own body. Without defining technological sovereignty, they note that in recent years, the concept of sovereignty has also been used to broadly denote forms of independence, control and autonomy over digital infrastructures, technologies and contents. Thus, in addition to the conceptions of sovereignty discussed above, they also identify technological sovereignty as defined in social movements, by indigenous peoples, and with regards to individuals.

Whether the focus is on the collective or the individual, in each of these cases sovereignty is reclaimed and asserted as a claim to authority and the legitimate exercise of power, often to further self-determination and in direct challenge to the hegemonic power of the state and/or private actors.

1.4. The colonial roots of sovereignty

What the above discussion points to, then, is that sovereignty is a norm and an ideal, as much as a concept (Bonilla, 2017). Yet, seeing the deep entanglement between empire and sovereignty, caution is advised against any uncritical adoption of sovereignty assertions as liberating. Since sovereignty has been used to defend colonialism by the colonisers, it is important to ask under what conditions it becomes possible to reclaim sovereignty despite these violent roots.

For post-colonial states, these difficulties have become evident as they found soon after independence that the promise of equality among states that the Westphalian order entailed in practice was a myth: instead, they found that differences such as those between citizens and subjects, occupation and settlement, which sovereignty had generated during colonial times, were already firmly inscribed in the landscape of international law (Bonilla, 2017).

But similar concerns apply wherever articulations of sovereignty can be found. Thus, our starting point in assessing any claims to sovereignty will need to be to ask questions about the knowledge production on sovereignty itself: 'who defines technological sovereignty and related concepts and for which purposes?' (Couture and Toupin, 2019: 5).

2. Sovereignty and data colonialism

This question gains particular importance as in the Indian context, the concept of data sovereignty often comes up in contradistinction to ideas of data colonisation, which in turn have emerged against the backdrop of global debates on the ownership of data. In particular, as private companies' free reign over the data of their users began to threaten the rights of people across the world, questions were increasingly also raised about whom the economic enrichment from data accrued to and about the potential for interference by these companies into collective internal matters such as elections. Thus, the growing dominance of foreign big tech companies has led both some key Indian tech entrepreneurs as well as members of the government to cry foul about data colonisation (see e.g. PTI, 2018; Goenka et. al, 2019).

And indeed, as Couldry and Mejias (2019) have argued, data colonialism is a useful frame to understand such modern forms of hegemony of big tech companies. At the heart of data colonialism, they argue, are data relations, a 'new type of human relations which enable the extraction of data for commodification' (337), thus 'normalising the exploitation of human beings through data' (336). At present, it is this new type of appropriation that drives capitalism, 'at every point in space where people or things are attached to today's infrastructures of connection' (337). The result is the slow emergence of a new form of capitalism, which is centred around the capitalisation of all aspects of human life, even the most intimate ones, through data.

As Couldry and Meijas (2019) explain, data colonialism presents a number of important parallels with historic colonialism, some of which have particular relevance for our investigation. First of all, as we will examine in more detail in the next section, dominant discourses today frequently construct data that has actual or potential relevance to people as a resource that is simply 'out there', up for grabs. As Cohen (2018) has noted, this naturalises the collection of data in ways that have strong parallels with the construction by colonial powers of faraway lands that were clearly inhabited as '*terra nullius*' or 'no man's land', legitimising their exploitation without legal intervention (see also Bonilla, 2017).

In addition, Couldry and Meijas (2019) argue, however, such constructions of data hide from view that for such data to exist and for its capture to become a possibility, 'the flow of everyday life must be reconfigured and represented in a form that enables its capture *as* data' (339; see also Cohen, 2018). This 'redefinition of social relations so that dispossession came to seem natural' (Couldry and Meijas, 2019: 4) forms another important parallel with historic colonialism. The big private players in the digital economy, including the social media platforms that so many of us use every day, play a fundamental role in facilitating this transformation of not only our economies, but our lives. To the extent that the state facilitates these practices, it becomes complicit in this dispossession.

Any challenge to data colonialism today can, then, only be effective to the extent that it challenges these underlying rationalities (Couldry and Meijas, 2019). And at first sight there might seem reason to hope that the assertion of Indian sovereignty in the face of data colonisation will indeed lead to increasing autonomy and freedom for the people who have constituted to sovereign state. 'To avoid data colonisation and allow for genuine empowerment, people must control the data they generate,' Nandan Nilekani, former Chairman of the Unique Identity Authority of India and co-founder of Infosys, has, for example, written (Nilekani, 2017).

But there is an added complication here: around the world, and in an important difference with historic colonialism, the elite agents of data colonialism target the populations of their home countries as much as they work externally. Moreover, such corporations are frequently framed as 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 (Couldry and Meijas, 2019).

And so the question here arises: what drives Indian assertions of data sovereignty? Are they aimed to further first and foremost its legitimate, geopolitical ambitions to finally take up its place as an equal in the community of nations? Or do they intend to facilitate genuine data decolonisation? Framed slightly differently: what reconfigurations of power and control are made possible by such claims? Who, ultimately, benefits?

3. Bodies and data

Before we dig into these questions more deeply, one more aspect of this theoretical framework needs to be further expanded upon. What makes possible the construction of data with actual or potential relevance to people as a resource that is simply 'out there?'

In the policy proposals we examine in the following sections, the dominant conceptual and metaphorical understandings construct data indeed as a resource that is simply out there, and thus up for grabs and ready to be mined. We argue that it is the erasure of the connection between data and people's bodies that is at the heart of this move. During historic colonialism, the construction of faraway lands as '*terra nullius*' required the erasure of the bodies of the people who inhabited those lands – either physically, or by ignoring their traditions of occupancy and use. Today, it is the erasure of the close connections between data and our bodies and selfhood that facilitates the construction of data as a resource.

3.1. The construction of data as a resource

Such understandings are not new: their origins can be traced to the discipline of cybernetics, in which these constructions of data have been dominant since at least the late 1940s. Data, such dominant constructions maintained, is a layer of information that somehow penetrates everything, yet that can exist independently from the medium carrying it (Hayles, 1999). It is in this way that data or information has come to be thought of as both dematerialised and disembodied, and as easily and unproblematically transferable from one medium to the next.

Moreover, this seemingly independent layer of information has been accorded with enormous power: it has come to be seen as the ultimate truth teller, somehow more accurate, more objective, more representative than what has ever come before of how we, how things *really* are – raw data as 'one's very nature exposed' (Grinberg, 2017).

Much of today's dataveillance, too, is informed by such understandings of data. Rather than targeting our bodies and selves in their totality, as surveillance did in earlier eras, surveillance now takes the form of capturing purportedly disembodied information flows about our bodies and its actions (Haggerty and Ericson, 2000).

If dataveillance has become so pervasive, its portrayal as disembodied has made an important contribution to this: it has made dataveillance somehow seem more innocuous than earlier forms of surveillance. In addition, surveillance now takes a more dispersed, fragmented form than was the case before, making it more and more difficult to pinpoint harms. The purpose of surveillance today remains the same as before: to direct or govern our actions. But as we are no longer subject to the same gaze and as context disappears from view since disembodied data points become the focus, how we will be treated or how likely we will be to receive discipline and punishment or reward will now depend on which boxes we have been slotted into from the outset (Haggerty and Ericson, 2000).

3.2. Data and the social

What disembodied constructions of data lose sight of, then, is that technology and data extraction are closely tied up in power relations – and this is particularly evident where data of actual or potential relevance to people is concerned. If the bodies that generate data do not exist outside of the social world (as evidenced, for example, by the fact that we do not treat all bodies equally even if we should), neither does data itself (boyd and Crawford, 2012). Decisions about what to include and what to ignore at the design level, what to pay attention to and what to disregard during data collection and analysis, always involve processes of interpretation.

In fact, even facets of life that we often take to be 'natural' and reproduce in technologies through categorisation and naturalisation, such as gender and race, do not reflect 'base facts' of human nature, but rather continuing assumptions about the nature of society and social relations (Glabau, 2019). 'Raw data' is a myth.

In addition, even large data sets can be full of errors and gaps, only amplifying the harms when interpretation of the data happens without those doing the interpretation acknowledging their own biases or those that shape the data sample as such (O'Neil, 2016). Contrary to the cybernetic imaginary, context all too often matters: when our bodies and its actions become datafied, this doesn't expose all of us equally, because not all of us are equally vulnerable. For example, while many people might not see any harm in linking their Aadhaar numbers with their health information, there were reports of people dropping out of their HIV antiretroviral treatment after Aadhaar began to be linked to HIV patient identity cards (Rao, 2016).

These shifts are particularly challenging for those already vulnerable, as unless biases are addressed consciously, such efforts are likely to lead to the reproduction, and possibly even deepening of inequalities. At the same time, the opaqueness of most such exercises complicates challenging them.

3.3. Data and bodies, data as bodies

The point here is not merely, however, that data is social too. As van der Ploeg (2012) has argued, as even the most intimate aspects of our lives become subject to datafication, an even more fundamental shift is taking place: the distinction between our physical bodies and our data bodies is becoming increasingly irrelevant.

For example, in India, reports have highlighted instances where people have not been able to access the rations they are legally entitled to because the authentication of their finger prints, stored under India's unique ID or Aadhaar, which is mandatory to access rations, failed – sometimes with starvation deaths as a result (Johari, 2018). Under the Aadhaar (Targeted Delivery of Financial and other Subsidies, benefits and services) Act, 2016, the denial of financial benefits because of failure authentication failures is not allowed; yet, in practice, such instances continue to occur.

What these point to is a paradigmatic shift in the conceptualisation of our bodies: as decisions based on our data bodies have such far-reaching consequences for our physical bodies, even to the extent of being matters of life and death, data emerges not merely as a reflection of our bodies, but as an extension of it – not as an independent layer, but as an integral part.

For the protection of our rights in the digital age, an understanding of this paradigmatic shift has profound implications. For example, where data emerges as an extension of our body, the harms of misuse of such data might in some cases be better seen as violations of bodily integrity than as data protection violations. If the sovereign state is to continue to safeguard the interests of the people who created it, it is therefore essential that it takes these shifting realities regarding the nature of data into account.

To what extent is this the case to in the Indian context? In the following sections, we examine policy documents and debates in India that use a sovereignty framework for data or digital technologies to start answering this question.

4. Sovereignty through data localisation

As noted above, the link with territory continues to play a crucial role in the articulations of sovereignty in the realm of data today. Within the many policy documents regarding data governance that are currently under discussion in India, one crucial mechanism that has emerged as a proposed way to assert data sovereignty exemplifies this: that is data localisation. Although broader definitions are in circulation, following Bailey and Parsheera (2018), we will understand data localisation as referring to 'mandatory requirements of local storage of data', whether exclusively or in the form of mirror data copies, thus fundamentally steering, and altering, data flows.

At present, India already has data localisation requirements in a number of sectoral policies, including for payment systems data (Reserve Bank of India, 2018), subscriber data in the broadcasting sector (Department of Industrial Policy and Promotion, 2017), and insurance policy-holder data (Insurance Regulatory and Development Authority of India, 2017).

Moreover, over the past two years, proposals for comprehensive data localisation have also been put forward, in the Draft Personal Data Protection Bill (Committee of Experts, 2018) and the draft National E-commerce Policy (Department of Industrial Policy and Promotion, 2019), in addition to further sectoral requirements in the draft E-pharmacy Regulations (Department of Health and Family Welfare, 2018). In some cases, provisions have been included to allow for the conditional lifting of cross-border restrictions of data transfer.

Broad-sweep data localisation proposals, in particular, have not been without controversy; yet critical appraisals seem to have had an impact: recent reports indicate that the blanket localisation requirements in these proposals are being reconsidered (PTI, 2019).

These criticisms should not come as a surprise. After all, by reorganising data flows to gain greater control over them, broad-sweep data localisation proposals illustrate the profound reconfiguration of power that the assertion of state sovereignty in both its external and internal aspects can entail.

Particularly important for our discussion is that in doing so, data localisation proposals also already recognise, even if implicitly, that the dividing line between our physical bodies and our virtual bodies is becoming irrelevant: after all, the aim of these policies generally is not merely to gain control over data as something that is 'out there' but also as a means through which to control – or protect, depending on your perspective – the physical bodies of people, including by the state. In fact, sometimes the phrase 'data residency' is used rather than 'data localisation', emphasising the need to ensure that the residency of data is the same as the residency of the person that data connects to.

The question that then emerges is: to what extent will this reterritorialization of their data benefit citizens and restore *their* autonomy?

As confirmed in the Puttaswamy judgement (2017), sovereignty lies with the people, a part of which is vested in the different apparatuses of the state. But data localisation proposals seem to see the container of sovereignty somewhat differently. Consider the following paragraph in the Personal Data Protection report drafted by the Shrikrishna Committee:

> So was the concept of the nation state bounded by territory and based on the principle of national sovereignty in the 17th century. If the unit in which sovereignty is vested and exercised is the nation state, it is inevitable that a movement towards making the nation state the central actor in Internet governance will emerge. There is no principled or practical reason to believe that the very fact of local storage or restriction to local processing itself will make the digital economy any less free or fair. On the contrary, it will ensure more effective enforcement of substantive obligations that are directed towards these objectives. It will be free and fair, but possibly different from the Internet we have today.

What are the substantive obligations that data localisation policies seek to address? In a research study on data localisation in India that is in fact subtitled 'Unpacking Policy Measures for Sovereign Control of Data', Basu, Hickok and Chawla (2019) have grouped the objectives of such measures into four categories: 1) enabling innovation; 2) improving cybersecurity and privacy; 3) enhancing national security; and 4) protecting against foreign surveillance. Bailey and Parsheera (2018) distinguish three sets of arguments: 1) those related to civil liberties; 2) those concerning government functions; and 3) those regarding economic development.

User interests, thus, do figure among the arguments presented in favour of data localisation. For example, a common argument for sectoral data localisation is that certain kinds of data – such as health data and finance data – require higher degrees of safeguards. This recognition is encouraging. But important new challenges that have emerged in the digital age remain unacknowledged. Thus, health data is no longer limited to the data that lives within files of hospitals. Interpretations from so many different varieties of data end up as health data, living far away from hospitals. Data gathered by smart watches of heartbeats, searches about conditions, states of mind derived from browsing history etc. are equally, if not more, sensitive data about health. Inevitably, any identification of health-related data therefore is bound to leave out some or the other sources of data. As elsewhere in the world, nothing in India's existing or currently proposed data governance related policies, however, acknowledges, let alone addresses, these realities.

Moreover, in addition to only providing half-hearted protection, the autonomy and choice of individuals are under question if data localisation proposals become a reality. Because not all Internet services will shift their infrastructure to relocate their data within Indian borders, a smaller selection of services will be available for use for Indians. Any discretion about the privacy and security of data that could have been exercised by individuals in making a choice about where to locate their data will no longer be available.

The consequences of this loss of choice and autonomy are far-reaching and go well beyond the question of where to locate one's data as such. With our bodies increasingly translated into data that is processed to determine at a distance who we are and how we should be treated, they 'become amenable to forms of analysis and categorisation in ways not possible before' by a multitude of actors, without us even having to even be physically present (van der Ploeg, 2012: 177). And as we outlined earlier, in such processes of deciding on our access to resources, services and power, information about our bodies is increasingly *privileged* over the presence of our physical bodies. Under these conditions, measures like blanket data localisation are not merely about losing control over where to locate one's data, but about the state, and possibly select domestic private parties, gaining an unprecedented level of access over the bodies of Indian citizens, their actions and behaviour, without any escape being possible.

In its current form, India's draft Personal Data Protection Bill, released in 2018, seems to do more to provide companies with a framework to collect, store and process personal data than it does to protect users' rights (Concerned People, 2018; Internet Democracy Project, 2018). But even a perfect Data Protection Bill would not be enough to address this problem comprehensively, since it goes much deeper. As van der Ploeg (2012: 180) has pointed out, 'we have very different regimes for protecting bodies and for protecting information from unjustified access and intrusion, however "personal" that information may be'. When the close connections between our bodies and our data are brought back into the picture, the far-reaching consequences of data localisation thus come into stark relief.

For example, doing so reminds us that data localisation makes possible intrusions in and manipulations of our bodies in hitherto unimaginable manners against we have little or no legal protection at the moment. It also foregrounds for consideration that data localisation in practice might entail a restriction on our freedom of movement that, in the current legal landscape, outweighs any possible gains in the protection of other rights. All of these affect our freedom, agency and autonomy in unprecedented ways.

Such concerns are not reflected, however, in discussions around data localisation today. Instead, studies such as those by Basu et. al, (2019) and Bailey and Parsheera (2018) as well as the relevant policies and proposals themselves, confirm that much of the rhetoric around the objectives relating to economic development and innovation in particular continues to support understanding of data as a resource.

This is exemplified by proposals in the draft National e-Commerce Policy to store data of Indian citizens within the country to make it available for Indian companies. As we (Internet Democracy Project, 2019) have pointed out in an analysis of the draft Policy elsewhere:

> In the name of enabling the country to benefit from rapid digitalization of the economy, the policy enables large scale extraction of data, while imagining frameworks like that of data protection to be models that legitimise such extraction instead of protecting against them.

A more detailed analysis of the construction of data in the draft National e-Commerce Policy will follow in the next section of this paper.

Moreover, if such proposals have emanated from the government, the drive towards data localisation by state institutions is further supported by important sections of India's tech-based industry (Mandavia, 2019). Some of these have appropriated the frame of data colonialism to promote such proposals, and through them, their own interests. For example, Reliance Jio, the biggest telecom service provider in the country, is owned by Mukhesh Ambani, one of the richest industrialists in India, who is using the banner of data colonisation against foreign companies, to support data localisation. Incidentally, the Reliance empire also has data centre infrastructure and is partnering with Microsoft to build cloud infrastructure in India (Aggarwal, 2019) as well as collecting and analysing data on its own customers (Bhatia, 2017; ET Telecom, 2019). In addition, Reliance Jio has gone on record to state that it believes Indian law enforcement should have full access to the data of Indian users, including sensitive data and decryption keys (Sathe, 2019).

There is of course nothing wrong with a company trying to make a profit as such. But it is important to remember here that businesses such as Reliance play a key role in the reconfiguration of even the most intimate aspects of our daily lives into data that can be tracked, captured, sorted, and valued by capital, as well as benefiting from this transformation themselves directly (Couldry and Meijas, 2019; Zuboff, 2019; on Reliance Jio specifically, see Bhatia, 2017).

As noted earlier, data colonialism today is not only directed towards those outside of the territorial boundaries of the state. The close entanglements, and revolving door, between government and key industry players in the development of digital India has earlier come up for scrutiny and criticism in the context of the Aadhaar project (Thaker, 2018). As close proximity to government puts these companies in an excellent position to influence both ideological discourses and material practices, such entanglements deserve to be watched for equally closely where debates around data localisation are concerned.

For now, for the citizens of Indian, data localisation seems to merely entail a transfer of power to domestic elites. While this might indeed contribute to strengthening India's profile and power in the community of nations, not an undesirable goal in itself, it does relatively little to return sovereignty to the people. Neither are the underlying rationalities of surveillance capitalism challenged in law and policy, nor are strong rights protections that centre the link between Indians' data and their bodies and selfhood emerging. As long as this is the case, the protection of Indian citizens' health-related and other data will remain half-hearted at best even with data localisation, with far-reaching consequences for the freedom and autonomy of the bodies and selves that transferred part of their sovereignty to constitute the Indian state.

5. Enabling the construction of data as a resource

As noted above, in the production of the claim of Indian data sovereignty, constructions of data as a resource continue to be dominant. For this resource extraction that is at the heart of surveillance capitalism to have become possible, enabling legal constructs had to be created. Cohen (2018) identifies and names the 'biopolitical public domain' as this construct. She defines it as

> a repository of raw materials that are there for the taking and that are framed as inputs to particular types of productive activity. The raw materials consist of information identifying or relating to people, and the public domain made up of those materials is biopolitical — rather than, say, personal or informational — because the productive activities that it frames as desirable are activities that involve the description, processing, and management of populations, with consequences that are productive, distributive, and epistemological (2).

In Indian law and policy proposals, two tools are of particular importance when constructing the biopolitical public domain of 'raw data' over which data sovereignty claims are sought to be asserted: a strong emphasis on the economic value of data at the expense of other concerns, and the particular resolution of questions of ownership that the Indian government is proposing. Both contribute to the erasure of our bodies from the data governance discourse. We examine each in turn below.

5.1. The economic value of data

Data sovereignty claims in India have constructed data as a primarily economic resource to be used in the service of economic enrichment of the country. India is seen as one of the last ‘untapped’ markets of data: that is, there are large sections of the population who are unconnected or are new users of the Internet, and the accumulation of these people as users can lead to collection of their data, and is seen as one of the ways that companies can consolidate market share in a large way. The economic value of data is, thus, central to these debates, and while other concerns may be paid lip service to, in practice they are subordinated to these financial considerations. Data sovereignty is centrally one of the ways in which an effort is being made to lay claim over the data of Indians and its value.

Let us examine in more detail how this is done in practice.

5.1.1. Generation of data as natural, inevitable and desirable

Data is presupposed as being generated in ever-larger volumes. This might be true, but as we noted above, this is not a naturally occurring phenomenon that is an inevitable factor of developments in technology, but in fact a result of a market where there is both demand for more data and a promise of development from this data. However, within Indian policy documents and proposal, the availability of increasing amounts of data is framed as an unquestionable state of affairs, and business models to monetise this data are then framed as an imperative.

Consider, for example, the following from the Economic Survey 2018–2019, published by the Ministry of Finance, Government of India.ii The Survey has an entire chapter devoted to data, encouragingly titled ‘Data “Of the People, By the People, For the People”’. The chapter constructs personal data as being consensually shared:

> Put differently, people produce data about themselves and store this data on public and private servers, every day, of their own accord (79).

But this is a misrepresentation. The digital economy is plagued by issues of consent, and the lack of understanding and meaningful consent to the storage and processing of data is a well-known fact. Cohen (2018), in fact, argues that consent has been left with such little work to do that it now merely functions as a ‘status that attaches at the moment of marketplace entry’. Personal agency has little to do with this, as the architecture of our daily lives has been intensely transformed to facilitate and encourage the production of data at every step.

The Economic Survey 2018–2019 further argues that:

> the surfeit of data and a limitless capacity to store it is of no use unless one can make sense of these colossal quantities of data in a reasonable time. Fortunately, human and technical capital to process data has evolved in parallel to the data inundation (79).

But capital to process data, and the technical skills surrounding it, are not factors that come into being post the fact of generation of data; they are drivers of the generation itself. The surfeit of data is not without impetus from intentional sources, like capital or technological solutionism.

The Draft National e-Commerce Policy, released by the Department for Promotion of Industry and Internal Trade in February 2019 and subtitled 'India's Data for India's Development', confirms, 'electronic commerce and data are emerging as key enablers and critical determinants of India's growth and economic development' (5).[iii] It further adds:

> There has been an unprecedented explosion in the volume of data generated during a commercial activity or a non-commercial activity (social media, climate data, health data, etc.) over the Internet...
>
> Creating economic benefits from data, that is, monetisation of data, is an important business model adopted by many corporations to generate profits by analysing, processing and utilising data (11).

It is correct that one of the most popular ways of monetising personal data on the Internet is through targeted advertising on search engines, social media and applications. But this is a business model that has not only thrived in the digital economy, it has also reduced space for traditional business models, and has increasingly come to be seen as damaging to the use of the Internet. None of these criticisms receive attention in the draft Policy.

The construction of data as being *out there* and available in ever-increasing volumes, and its naturalisation of being a productive factor in the economy are threads that are thus strongly present in policy documents. Claims such as those discussed above overlook that the generation of data in increasing volumes is neither inevitable not natural, nor is its commercialisation necessarily desirable.

5.1.2. State facilitation of data extraction for private profits

After the construction of data as raw material available for the obvious purpose of economic enrichment, the draft National e-Commerce Policy, in fact, seems to actively encourage the problematic dominant business models that are built around extraction in order to shape behavioural modifications and that are at the heart of what Shoshana Zuboff (2019) has labelled 'surveillance capitalism', a form of capitalism in which companies are effectively making profits by taking bets on people's future behaviour. In a section titled 'The most critical factor in success of an enterprise', the draft Policy approvingly states:

> The individual's profile can be used for a variety of commercial purposes, such as precision marketing, targeted advertisements and credit worthiness assessments. The history of browsing and search by consumers also generates rich information on consumer preferences and, at times, their potential purchasing power. By tracking the search and browsing histories, online retail websites are able to target consumers with tailor-made marketing content. Companies with maximum access to data about consumers stand to make windfall profits from leveraging this through targeted advertising and product development (12).

Evidence of the problems with these practices is, however, growing (see also Zuboff, 2019). For example, complaints about the advertising industry by researchers in the United Kingdom led to the publication of a report by the Information Commissioner of the UK, which reinforced what has been known about the online advertising industry: that detailed profiles about individuals are traded between hundreds of organisations, without the knowledge of the data subjects (Information Commissioner's Office, 2019). 'Behavioural advertising is out of control, warns UK watchdog,' reads a TechCrunch article on the ICO's findings (Lomas, 2019). While still conservative, one of the next steps that the ICO sees is an 'industry review' in the near future.

And yet, this same business model is one of the ways that the draft National e-Commerce Policy sees as a way of achieving economic enrichment. Elsewhere it notes in the same optimistic vein:

> Data is the basis on which online advertisements are tailored and consumer preferences are gauged. During the last decade a further evolution has taken place. Big data and the use of Artificial Intelligence (AI) thereon have taken data crunching to the level of 'deep learning', which is automated problem solving through neural network layers. Image recognition, facial recognition software, self-driving cars, which were earlier seen as works of science fiction are now reality (11).

The precarious but unsuccessful balancing act between user rights and the promotion of data as an economic resource can be found in other government documents as well. For example, the National Digital Communications Policy of 2018, issued by the Department of Telecommunications, includes among its missions:[iv]

> To secure the interests of citizens and safeguard the digital sovereignty of India with a focus on ensuring individual autonomy and choice, data ownership, privacy and security; while recognising data as a crucial economic resource (5).

But if data sovereignty indeed intends to allow individuals to exercise autonomy and take at least some amount of control back, then a reform of the dominant business models and not their celebration is merited, even if the economic value of data continues to be explored.

At the moment, dominant discourses in India's policy landscape seem to be limited to rallying against foreign entities gathering the data of Indians, while stewarding and encouraging the extraction of the same data by Indian firms. The questioning of data colonialism's underlying rationalities that is essential to further the rights of Indian citizens in a substantive manner, and to thus protect their bodies and selves, not merely their 'data', remains absent.

5.2. Ownership of data and sovereignty

Notions of ownership of data and property also animate policy debates around data in India to a significant extent. Both government documents such as the draft National e-Commerce Policy and titans of industry such as Mukesh Ambani emphasise that Indians should be able to take ownership over their own data (PTI, 2018).

However, the paradigm of ownership in the context of data, too, does little to challenge the rationalities that underlie data colonialism, while continuing the myth that data is always at a remove from our bodies and personhood.

5.2.1. Ownership and the individual

As noted earlier, consent has limited value when the terms of the architecture of data generation, collection and storage are not open for debate or criticism or overhaul (Cohen, 2018). The paradigm of ownership, too, only provides limited control if the parameters of the market in which it has to operate are already established. In particular, within this market, much of our data has value not on its own, but when combined with data points from a large number of other people. Moreover, based on the analysis of such large, aggregated data sets, inferences may be made about us even if our own data is not included in the original data set (Haggerty and Ericson, 2000; Tisné, 2018; van der Ploeg, 2012; Zuboff, 2019). Having ownership over our data will not stop this from happening.

Bringing bodies back into the debate can once again further elucidate what is at stake here. In democratic societies, questions of human dignity and bodily integrity have never been reduced to questions of 'ownership'. Thus, Indian laws make it impossible, for example, to sell yourself into slavery even if you want to. While private property might be protected, ultimately the values of freedom, agency and dignity gain primacy.

In a similar vein, protecting our freedom, agency and dignity in the digital age requires that our data is not merely reduced to a resource that we can trade in a deeply asymmetrical market in which we hardly have any power. As Tisné (2018) has argued, in this context, a far more substantive protection of data-related rights is needed. Although developing this in more detail is beyond the scope of this paper, bringing our bodies back into the picture highlights the importance of values beyond ownership to be highlighted in such a bill of rights, and can, thus, be a useful starting point to help elucidate what these substantive rights should consist of.

5.2.2. Ownership and community data

Questions of ownership of data do not only figure in debates about individual users, however. In 2018, the government-constituted Committee of Experts under the Chairmanship of Justice B.N. Srikrishna released a draft Personal Data Protection Bill, as well as an accompanying report titled 'A Free and Fair Digital Economy: Protecting Privacy, Empowering Indians'.[v] This report articulated for the first time a new category of data, called 'community data.'

> In relation to data fiduciaries, there is an emerging need to recognise a new category of information as community data. This is information that is valuable owing to inputs from the community, which might require protection in addition to individuals' personal data (24).

This new category of data is said to relate to a 'group dimension of privacy;' at the same time, it also creates a space of dispersed ownership by saying that ownership is difficult to ascertain.

> It is a body of data that has been sourced from multiple individuals, over which a juristic entity may exercise rights. Such data is akin to a common natural resource, where ownership is difficult to ascertain due to its diffused nature across several individual entities. It is relevant for understanding public behaviour, preferences and making decisions for the benefit of the community (45).

Deploying 'community' in a very loose manner, the report distinguishes community data from big data sets depending on the degree of involvement of 'the larger community' in building the dataset. It gives the example of the data gathered by products of private companies like Google Maps as an example of community data. Two issues are recognised with such data collection:

> Though these services are incredibly useful, two concerns arise. First, an individual's sharing of her personal data (such as current location) may lead to the sharing of similar personal data of her spouse, friends or family, without

> their consent. Second, juristic entities make use of Big Data and can identify patterns of behaviour. This can have spill-over effects on the entire community as decisions may be taken on the basis of such patterns. Thus, community data may deserve protection (46).

The vision for higher protection is laid out in more detail to include class action remedies for data breaches involving community data, where the harm is social and systemic. This is a potentially important step in moving from a western liberal individual framework of privacy that does not have much room for both the value as well as the harms that can accrue from the aggregation of data.

But the evolution of the concept of 'community data' in other policy documents, such as the draft National e-Commerce Policy, is, unfortunately, concerning. The draft Policy follows a trend in recent technology policymaking where the concept of community or the commons is used strategically in the realm of the digital to create a vacuum of ownership, which is then followed by such ownership being asserted by the government.

Thus, the draft Policy starts by invoking Indian citizens only to help contrast and distinguish the claims of foreign entities. Once that is achieved, the ownership over this data is subsumed by the state theoretically, and by the government in practical terms.

> India 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). This understanding flows from the acknowledgement that data about an Indian, is his/her own. Even after anonymization, the interests of the individual cannot be completely separated from the derivatives that may be obtained by analysing and drawing inferences from a certain set of data. Data can, therefore, best be likened to a societal 'commons.' (14)

The draft Policy then continues to argue that rights are 'permitted' over this resource that the government holds in trust, thus making rights secondary to the government setting priorities for the data:

> The data of a country, therefore, is best thought of a collective resource, a national asset, that the government holds in trust, but rights to which can be permitted. The analogy of a mine of natural resource or spectrum works here… (14)

> National data of various forms is a national resource that should be equitably accessed by all Indians. The same way that non-Indians do not have access to the national resources on the same footing as Indians, non-Indians do not have equal rights to access Indian data. However, access to it can be negotiated, in national interest. Thus, the e-commerce policy is about how best to exploit this national resource, for maximizing growth and for delivering greatest benefits to all sections of society (15).

Even as community data forms a large part of the policy, it is not defined. But what is specified is how it may be used by the government and private parties in the name of public good. Without definitions, an understanding of how to grapple with competing interests of different communities, and the fact that data can simultaneously be personal data as well as community data, the category only serves to create a class of data over which individual claims for data protection can be weakened.

The e-commerce policy also sidesteps defining public interest, while it is a concept running through the entire document as justification. Public interest is used as a justification to allow for commercial exploitation of this data. This in effect gives no tools for arriving at workable definitions of emergent or existing communities, nor does it define public interest in a way that can guide arriving at these definitions.

> Suitable framework will be developed for sharing of community data that serves larger public interest (subject to addressing privacy-related issues) with start-ups and firms. The larger public interest or public good is an evolving concept. The implementation of this shall be undertaken by a 'data authority' to be established for this purpose (17).

Recently, the Ministry of Electronics and Information Technology (Meity) has formed a new Committee of Experts to come up with a governance framework for 'non-personal' and 'community' data (Agarwal, 2019). The committee is headed by Infosys co-founder Kris Gopalakrishnan, creating alarming conflicts of interest.

- *Community control over data or indigenous data sovereignty*

The draft e-commerce policy lists examples of where there is a reclaiming of control over data. Among these are the Maori Data Sovereignty Network, Project Decode in

Barcelona and Amsterdam and First Nations Information Governance Centre. But these examples are both ideologically and practically different from what the draft policy suggests. The starting point for indigenous data sovereignty is an understanding of who comprises the community. This baseline understanding is missing in the government's proposal and weakens the case for community data.

Indigenous communities in many parts of the world have used data as a strategic resource for self-determination and self-governance. This framework has emerged from demands for control over their own data. Moreover, this has not emerged as simply a grab of data about indigenous peoples but reflects a more intentional engagement with data, that includes the understanding that the usefulness of data depends on other factors:

> Reliance on data that do not reflect tribal needs, priorities, and self-conceptions threatens tribal self-determination. Tribal data sovereignty through governance of data on indigenous populations is long overdue. (Rainie, Schultz, Briggs, Riggs & Palmanteer-Holder, 2017)

One question this engenders is how the benefits from data aggregation and processing should be distributed. When we speak about the larger public good, what exactly is the constitution of the public? There are honest attempts at solving these. However, the co-option of indigenous data sovereignty, community data and such concepts does not look promising in the Indian context.

- *Problems with the commons not recognised*

This is also true because these proposals disregard that the commons has never been without its problems in the Indian context, as elsewhere. From public space to wells of drinking water, the commons are spaces where there is exclusion on the basis of caste, and to an extent gender and other barriers (Nath, 2019). Even where there are no clear lines of discrimination, there are competing interests over the commons.

Without surfacing these complexities, what community data serves to do is become a category where the ownership is somewhat dispersed, the outline of who forms a community and how decisions over it can be taken is absent, and room is essentially created for other manners of claims to be asserted: those from the state and from the private sector, as well as possibly from other powerful actors within this ‘community’.

We see time and again that profit motivations of private companies, motivations for control by the government, or even individual preferences for values like convenience can be harmful to the rights of citizens. For example, changes to the law on Aadhaar have consistently been driven by increases in business use cases of the data, while instances of exclusion are ignored as ‘edgecases’ to be ironed out. The use of the community data paradigm in the Indian context so far unfortunately seems to hold little promise that within this context, control over data relevant to them will be devolved to historically marginalised communities within the country to improve the protection of their freedom, autonomy and dignity in the age of datafication. Instead, the concept of community data seems to be foregrounded to make available ever more data in the service of the ‘national interest’, and to assert sovereignty to further geopolitical aspirations of global dominance rather than the freedom and autonomy of all the communities that make up the Indian people.

Conclusion

In this paper, we sought to understand to what extent assertions of data sovereignty in the Indian context might be able to contribute to substantively promote the rights of Indian citizens in the digital age. We asked who defines data sovereignty in the Indian context and what reconfigurations of power and control are made possible by such claims.

As this paper has illustrated, data sovereignty in India is a vision created and asserted by arms of the government, with support from select sections of India’s tech industry, and imagines the state as the vessel of such sovereignty. Legal constructs and normative understandings that are entangled with this proposed vision for sovereignty treat data as an abundant resource, ideal for exploitation by the market. While individual privacy and autonomy of citizens do find mention in policy documents envisioning data sovereignty, they are not fleshed out, or are seen as secondary to larger collective agendas like economic enrichment, the terms of which are defined by the state, often in close conjunction with powerful private actors.

By promoting such visions, the government and industry do little to undermine the underlying rationalities of the data colonialism that they at times claim to be responding to. Rather, the portrayal of data as a resource, the emphasis on its economic value at the expense of other considerations, and the centrality of the notion of ownership (rather than say dignity, freedom, and/or integrity) are all legal constructs that further enable the structural perpetuation of such data colonialism of the Indian people, however now by Indian entities instead of foreign ones.

Ultimately, by increasing India's economic prowess, such policies might indeed contribute to the country's weight in the community of nations, the structure of which itself continues to be profoundly influenced by the legacy of colonialism. But by continuing to perpetuate the erasure of our bodies and selfhoods from debates about data governance, they do little to substantially further the autonomy, freedom and dignity of the Indian people in the digital age.

Instead, executive arms of governments in power, and private companies that are cooperating with the government in the collection and processing of data, are emboldened in taking far-reaching decisions about data of Indian citizens, and in the process, about their physical bodies and lives. Moreover, because of the erasure of bodies from data governance discourses, this is allowed to happen without the substantial protections and accountability measures that typically attached to such decisions and activities in democratic countries in the pre-digital age.

In their current form, rather than providing greater autonomy and freedom to the people, data sovereignty claims in the Indian context, thus, seem to facilitate first and foremost greater control of the people by state and domestic private actors alike. For the people, if not the state, data sovereignty, for now, continues to remain a dream.

References

Agarwal, A. (2019, 16 September). Indian govt forms committee to recommend governance norms for non-personal data, Infosys' Gopalakrishnan to head it. *Medianama*. https://www.medianama.com/2019/09/223-meity-non-personal-data-committee/

Aggarwal, V. (2019, August 12). Reliance Jio partners Microsoft for cloud infrastructure. *The Hindu*. https://www.thehindubusinessline.com/info-tech/reliance-jio-partners-microsoft-for-cloud-infrastructure/article29025461.ece

Barlow, J. P. (1996, February 8). A Declaration of the Independence of Cyberspace. Retrieved from https://www.eff.org/cyberspace-independence

Bailey, R. & Parsheera, S. (2018, 31 October). Data Localisation in India: Questioning the Means and Ends. New Delhi, National Institute of Public Finance and Policy, Working Paper No. 242. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3356617

Basu, A., Hickok, E. & Chawla, A. S. (2019). *The Localisation Gambit: Unpacking Policy Measures for Sovereign Control of Data in India*. Bangalore: Centre for Internet and Society. https://cis-india.org/internet-governance/resources/the-localisation-gambit.pdf

Bhatia, R. (2017, 7 August). Is Big Data Turning the Wheels at Reliance Jio – Inside the Youngest Mobile Operator's Big Data Strategy. *Analytics India Magazine*. https://analyticsindiamag.com/big-data-turning-wheels-reliance-jio-inside-youngest-mobile-operators-big-data-strategy/

Bonilla, Y. (2017). Unsettling Sovereignty. *Cultural Anthropology*, 32: 330–339. https://doi.org/10.14506/ca32.3.02

boyd, D. & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication and Society*, 15(5): 662–679. https://doi.org/10.1080/1369118X.2012.678878

Cohen, J. E. (2018). The Biopolitical Public Domain: The Legal Construction of the Surveillance Economy. *Philosophy and Technology*, 31(2): 213–233. http://dx.doi.org/10.1007/s13347-017-0258-2

Concerned Citizens (2018). Solving for Data Justice: A Response to the Draft Personal Data Protection Bill. New Delhi, Internet Democracy Project. https://internetdemocracy.in/reports/datajustice/

Couldry, N. & Mejias, U. A. (2019). Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject. *Television & New Media,* 20(4): 336–349. https://doi.org/10.1177/1527476418796632

Couture, S. & Toupin, S. (2019). What Does the Notion of 'Sovereignty' Mean When Referring to the Digital? *New Media and Society*, 21(10): 2305–2322. https://doi.org/10.1177/1461444819865984

ET Telecom (2019, 22 August). Reliance Jio Taps Guavus' AI-based Data Analytics to Improve Customer Experience. *Economic Times*. https://telecom.economictimes.indiatimes.com/news/reliance-jio-taps-guavus-ai-based-data-analytics-to-improve-customer-experience/70788340

Glabau, D. (2019, March 18). Natural's Not in It: Countering Biological Essentialism with a Biological Futurism. *Real Life Magazine*. https://reallifemag.com/naturals-not-in-it/

Goenka, V., Patil, V. M., Shekatkar, D. B., Khandare, V., Bhatia, V., Ranade, J., & Panchal, B. (2019). *Data Sovereignty: The Pursuit of Supremacy*. Delhi: Penman Books.

Grinberg, Y. (2017). The Emperor's New Data Clothes: Implications of 'Nudity' as a Racialised and Gendered Metaphor in Discourse on Personal Digital Data. In Daniels, J., Gregory, K. & McMillan Cottom, T. (Eds.). *Digital Bodies.* Bristol: Policy Press.

Haggerty, K. D. & Ericson R. V. (2000). The surveillant assemblage. *British Journal of Sociology*, 51(4): 605–622. https://doi.org/10.1080/0007131002001528

Hayles, K. N. (1999). *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: University of Chicago Press.

Information Commissioner's Office (2019, 20 June). B*log: ICO Adtech update report published following industry engagement.* Retrieved from https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/06/blog-ico-adtech-update-report-published-following-industry-engagement/

Internet Democracy Project (2018). Is the Fourth Way Going Far Enough? Our Submission to MeitY on Draft Personal Data Protection Bill 2018. New Delhi, Internet Democracy Project. https://internetdemocracy.in/reports/pdpb/

Internet Democracy Project (2019). Submission in Response to the Draft e-Commerce Policy. New Delhi, Internet Democracy Project. https://internetdemocracy.in/reports/submission-in-response-to-the-draft-e-commerce-policy/

Johari, A. (2018, 3 February). Yet Another Aadhaar-linked Death? Denied Rations for 4 Months, Jharkhand Woman Dies of Hunger. *Scroll*. https://scroll.in/article/867352/yet-another-aadhaar-linked-death-jharkhand-woman-dies-of-hunger-after-denial-of-rations

Lomas, N. (2019, 20 June). Behavioural advertising is out of control, warns UK watchdog. *Tech Crunch*. https://techcrunch.com/2019/06/20/behavioural-advertising-is-out-of-control-warns-uk-watchdog/

Mandavia, M. (2019, 22 February). How desi tech lobby is giving Silicon Valley giants a run for their money. *Economic Times*. https://economictimes.indiatimes.com/tech/internet/how-desi-tech-lobby-is-giving-silicon-valley-giants-a-run-for-its-money/articleshow/68102813.cms

Nath, A. (2019, 22 August). Tamil Nadu: Funeral procession blocked, Dalits airdrop body for cremation. *India Today.* https://www.indiatoday.in/india/story/dalit-man-funeral-procession-denied-in-vellore-community-says-not-the-first-time-1590160-2019-08-22

Nilekani, N. (2017, July 27). Why India needs to be a data democracy. *Livemint*. https://www.livemint.com/Opinion/gm1MNTytiT3zRqxt1dXbhK/Why-India-needs-to-be-a-data-democracy.html

O'Neil, C. (2016). *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy.* New York: Crown Publishers.

PTI (2019, 23 July). Personal Data Protection Bill: IT Ministry may back storage curbs for critical, sensitive data. *The Hindu*. https://www.thehindubusinessline.com/info-tech/personal-data-protection-bill-it-ministry-may-back-storage-curbs-for-critical-sensitive-data/article28682941.ece?fbclid=IwAR1WgGPacmj55DkeW8m-VBT-XfT6D78EdNqKO0aBm83sFRa0-q8ETSjZ7Po

PTI (2018, 19 December). Mukesh Ambani says 'data colonisation' as bad as physical colonisation. *Economic Times*. https://economictimes.indiatimes.com/news/company/corporate-trends/mukesh-ambani-says-data-colonisation-as-bad-as-physical-colonisation/articleshow/67164810.cms?from=mdr

Rainie, S. C., Schultz, J. L., Briggs, E., Riggs, P. & Palmanteer-Holder, N. L. (2017). Data as a Strategic Resource: Self-determination, Governance, and the Data Challenge for Indigenous Nations in the United States. *The International Indigenous Policy Journal, 8(2).* https://ir.lib.uwo.ca/iipj/vol8/iss2/1

Rao, M. (2017, 17 November). Why Aadhaar is prompting HIV positive people to drop out of treatment programmes across India. *Scroll*. https://scroll.in/pulse/857656/across-india-hiv-positive-people-drop-out-of-treatment-programmes-as-centres-insist-on-aadhaar

Sathe, G. (2019, 30 January). Reliance Jio, Paytm Tell TRAI They Will Spy on Us for the Government. *Huffington Post*. https://www.huffingtonpost.in/entry/reliance-jio-paytm-tell-trai-the-government-should-be-able-to-access-your-data_in_5c4f30e9e4b0f43e4109647c

Sriraman, T. (2018). *In Pursuit of Proof: A History of Identification Documents in India.* New Delhi: Oxford University Press.

Thaker, A. (2018, 1 May). Aadhaar's Mixing of Public Risk and Private Profit. *The Caravan*. https://caravanmagazine.in/author/796

Tisné, M. (2018, 14 December). It's Time for a Bill of Data Rights. *MIT Technology Review*. https://www.technologyreview.com/s/612588/its-time-for-a-bill-of-data-rights/

van der Ploeg, I. (2012). The body as data in the age of information. *Kirstie Ball, Kevin Haggerty and David Lyon (eds.), Routledge Handbook of Surveillance Studies.* New York: Routledge.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Public Affairs.

Endnotes

[i] Justice K. S. Puttaswamy (Retd.) and Anr. v. Union of India. WP (Civil) No 494 of 2012.

[ii] Available at https://www.indiabudget.gov.in/economicsurvey/.

[iii] Available at https://dipp.gov.in/sites/default/files/DraftNational_e-commerce_Policy_23February2019.pdf.

[iv] Available at http://dot.gov.in/whatsnew/national-digital-communications-policy-2018.

[v] Available at https://meity.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf.

Acknowledgements

The authors would like to thank Elizabeth Coombs, an anonymous peer reviewer and all participants at the Data Governance Network roundtable held in New Delhi on 4 September, 2019 for their valuable inputs and comments.

About the Authors

Dr. Anja Kovacs directs the Internet Democracy Project in Delhi, India. Her research and advocacy currently focuses on questions regarding data governance, surveillance and cybersecurity, and regarding freedom of expression — including work on gender, bodies, surveillance, and dataveillance, and gender and online abuse. She has also conducted extensive research on the architecture of Internet governance.

Dr. Kovacs has worked as an international consultant on Internet issues, including for the Independent Commission on Multilateralism, the United Nations Development Programme Asia Pacific and the UN Special Rapporteur on Freedom of Expression, Mr. Frank La Rue, as well as having been a Fellow at the Centre for Internet and Society in Bangalore, India, and a CyberBRICS Fellow at the Fundação Getulio Vargas (FGV) in Rio de Janeiro, Brazil. She currently a member of the Board of Governors of Veres One.

Prior to focusing her work on the information society, Dr. Kovacs researched and consulted on a wide range of development-related issues. She has lectured at the University of East Anglia, Norwich, UK, and Ambedkar University, Delhi, India, as well as guest lectured at universities in India and Brazil, and has conducted extensive fieldwork throughout South Asia. She obtained her PhD in Development Studies from the University of East Anglia in the UK.

Nayantara Ranganathan is a researcher and lawyer interested in the politics and culture of technologies. She has explored questions of what it might mean to have a feminist politics of data at the Internet Democracy Project, where she developed work related to freedom of expression, surveillance, net neutrality, and the gendered use of the internet. For example, she co-conceptualised Gendering Surveillance, a research project that exposes the gendered nature of surveillance in the Indian context. She has also been working with groups in India to explore how digital technologies are affecting fights for social justice. Beyond her work at the Internet Democracy Project, Nayantara has co-founded ad.watch, a project compiling and presenting political ads from Facebook and Instagram.

datagovernance.org dgn@idfcinstitute.org

@datagovnetwork /datagovnetwork /datagovnetwork