Community Data and Decisional Autonomy:

Dissecting an Indian Legal Innovation for Emerging Economies

Amber Sinha and Arindrajit Basu

Key Takeaways

- Concerned with the power asymmetries between big tech companies and Indian citizens in terms of data sharing and processing practices, the Indian government has put in place a number of policies seeking to unlock the developmental potential of data for Indian citizens.

- While several policy instruments are still works in progress and need improvement to be in line with India's constitutional framework, international human rights law and economic welfare, they have advanced some important conceptual innovations. One such innovation is "community data," which attempts to delineate the rights and interests a community would have in its data.

- However, the existing framework does not satisfactorily define community, and does not sufficiently balance the privacy and decisional autonomy of individuals with the interests of the community and the nation in economic and social empowerment.

- The gap can be addressed by looking at Indian jurisprudence on privacy and decisional autonomy, and analysing how existing case law can be applied to the digital era. As Europe grapples with debates about "technological sovereignty," the framing of community data in line with Indian privacy jurisprudence may be valuable.

- **Policy Recommendation 1:** By studying unique Indian case law on privacy that deals with the question of individual and group rights, we find that decisional autonomy is the fulcrum of privacy jurisprudence, and thus should be the edifice for any policy framework. In a case of conflict between individual and group rights, individual rights must prevail.

- **Policy Recommendation 2:** Providing communities with adequate rights and interests while also prioritising individual rights is very much in line with human rights principles espoused by Europe, and endorsed in the General Data Protection Regulation (GDPR), and Europe should consider how an improved version of India's community data approach may be used to further its digital sovereignty vision without compromising on European human rights ethos.

1 Introduction

The last decade has witnessed a sea change in the power asymmetries that shape society and global governance structures alike. The rise of "big tech" companies that monetize individual data has triggered a global discourse on individual privacy in the digital age. Europe has been at the forefront of driving these developments by setting the benchmark on personal data protection laws with its General Data Protection Regulation (GDPR). India and other Asian economies have followed suit with their own data protection laws, enacted or proposed, modelled largely on European standards.

The policy and legal discourse in India has additionally focused on an equally important strand of this power asymmetry, relevant both for India and other emerging Asian economies. This asymmetry, appropriately called "data colonialism," describes the extractive economic practices of global technology giants that derive benefits from the data of citizens in Global South countries to consolidate their own market power, at the expense of developmental needs in these very countries.[1]

The "data for development" narrative has centred around a conception of community data, which has been referred to in multiple policy instruments, and which has been articulated most comprehensively in the recent report on Non-Personal Data, submitted to the Ministry of Electronics and Information Technology (MeitY) by the Gopalakrishnan Committee.[2] This Committee of Experts was set up by MeitY in 2019 to provide recommendations on creating a framework for governance of Non-Personal Data. At its core lies the idea that the "community" of Indian citizens, through the state, have the right to receive the welfare benefits of any data generated by other citizens; benefits that are currently being extracted solely by private technology companies.[3]

The framing of community data in the policy instruments leaves much to be desired, as we identify in the first section of this paper. However, it is also a bold legal innovation aimed at granularly addressing the rhetorical framing of data for the public good. The benefits of data processing, and the rights associated with the data one produces must be distributed equitably across defined communities, and the subgroups and individuals that make up these defined communities. These are gaps not addressed by the policy ecosystem in India which is surprising given that answers

1 Couldry, Nick and Ulijies Mejias. 2018. "Data Colonialism: rethinking big data's relation to the contemporary subject." Television and new Media. (https://eprints.lse.ac.uk/89511/1/Couldry_Data-colonialism_Accepted.pdf).

2 The committee was set up to articulate a governance framework for non-personal data in India.

3 See Basu, Arindrajit. 2019. "India's role in global cyber policy formulation." Lawfare. (https://www.lawfareblog.com/indias-role-global-cyber-policy-formulation).

can readily be extrapolated by relying on landmark Indian judgments on privacy. In this paper, we attempt to fill some of these gaps by relying on uniquely Indian legal thought. As Europe engages with technological sovereignty and tries to govern data to further the growth and equitable distribution of economic welfare, the innovation of community data has several learnings that will enable equitable distribution of rights and resources, and the fulfilment of a right to privacy.

The objective of this paper is not to arrive at or recommend an overarching framework for the governance of data and extraction of its economic benefits. It is limited to positing the Indian notion of "community data" as a workable legal innovation, while acknowledging and recommending solutions to the gaps in its present conception. The paper is divided into three broad sections. The first charts out the policy trajectory that defines community data and highlights lacunae in its present framing. The second charts out the historical evolution of community and group interests in Indian constitutional jurisprudence, focussing on jurisprudence around the right to privacy. Finally, the third aims to use this jurisprudence to answer some of the questions that the framing in the previous sections poses on the conceptions of community data. It also highlights the lessons Europe may draw from the Indian framing of community data.

2 "Community Data" in Indian Policy

Over the past few years, the incursion of foreign data-driven technology companies into India has resulted in clarion calls for preserving India's "data sovereignty," and championing strategies for using the data of Indian citizens for their own development. After a series of cacophonous policy moves attempting to conceptualise the notion of data for "public good," in July 2020, a committee on non-personal data set up by the Ministry of Electronics and Information Technology (MeitY) released a non-personal data framework (hereinafter "NPD Report") that attempted to comprehensively outline the contours of community non-personal data.[4] This is the first report in the world that looks to define, construct and chart out the contours of "community data," although there are several gaps in its framing. This section of our paper will critically engage with the existing legal and policy framework, and the recommendations of the NPD report, while trying to situate it within the existing policy ecosystem seeking to govern data, since the report itself fails to draw clear links.

When defining community data, the Srikrishna Committee Report (2018), which accompanied the first draft of the personal data protection bill, charts out a collective protection of privacy for an identifiable community that has contributed to community data.[5] It does not posit any specific recommendations, but suggests that a suitable law should facilitate the provision of collective protection of privacy to an identifiable community that has contributed to community data through class action remedies or group sanctions.[6] The draft E-commerce policy (2019) broadens the notion of community data as "societal

4 The report defines non-personal data as "*Firstly, data that never related to an identified or identifiable natural person, such as data on weather conditions, data from sensors installed on industrial machines, data from public infrastructures, and so on. Secondly, data which were initially personal data, but were later made anonymous. Data which are aggregated and to which certain data-transformation techniques are applied, to the extent that individual-specific events are no longer identifiable, can be qualified as anonymous data.*"

5 SriKrishna Committee. 2018. "A Free and Fair Digital Economy." (https://www.meity.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf).

6 SriKrishna Committee. 2018. "A Free and Fair Digital Economy."

commons" or a "national resource," where the undefined "community" has rights to access datasets, but the government has overriding control.[7]

A related idea that further confuses matters is the notion of "data as a public good" articulated in Chapter 4 of the 2019 Economic Survey Report, a document published by the Ministry of Finance along with the annual budget.[8] The report states that the personal data of an individual can be considered a public good when it is in government custody, and the datasets are anonymised. It does not engage clearly with non-excludability and non-rivalry – economic prerequisites for an entity to be considered a public good. Instead, it allows private corporations to bid for data being held by the government, which is fundamentally incompatible with both conditions.

Given this uncertain backdrop, the NPD report makes a fair attempt at trying to resolve some existing gaps in defining and conceptualising community data. First, the report defines a community as "*any group of people that are bound by common interests and purposes and involved in social and/or economic interactions. It could be a geographic community, a community by life, livelihood, economic interactions or other social interests and objectives and/or an entirely virtual community*." This definition casts a wide net on the kinds of groups that might get classified as a community. Further, it provides no clarity on the relationship between the individual and the community. When does an individual become a part of the community? When does membership translate to common rights over and access to resources such as data?

The report then notes that "community non-personal data" includes non-personal data, which includes both personal data that has been anonymised, and non-personal data about animate and inanimate phenomena. Interestingly, it uses the examples of data collected by municipal corporations, and private players, such as ride-hailing companies, to help clarify the point. This further troubles the definition of a community because it seems to suggest that all users of ride-hailing companies, or all individuals who provide data to municipal corporations, form a single community, even though the individuals may not have consented to community membership or a joint governance framework for ostensibly shared resources. These are important theoretical gaps that need to be filled before any governance framework for non-personal data is conceptualised. In the next section, we bring to light several theories evolved in Indian constitutional jurisprudence to do so.

At this stage, it is important to distinguish the construct of "community data" from related concepts in existing academic discourse. "Group privacy" is a limited interest that groups have in data, which is extracted using aggregated individual data via algorithmic analysis that in certain cases where the individual and the data processor are unaware of.[9] "Community data," as we describe in this paper, is a far broader set of rights and interests that is not limited to group privacy.

7 Indian Department for Promotion of Industry and Internal Trade. 2019. "Draft E-Commerce Policy." (https://dipp.gov.in/sites/default/files/DraftNational_e-commerce_Policy_23February2019.pdf). 6.

8 IndianMinistry of Finance. 2019. "Economic Survey of India." India Budget. (https://www.indiabudget.gov.in/budget2019-20/economicsurvey/doc/echapter.pdf). 81.

9 Kammoruieh, Lanah. 2018. "Group Privacy in the age of big data." In Linnet Taylor, Luciano Floridi, and Bart Van Der Sloot (eds.) *Group Privacy: New Challenges of Data Technologies*. Heidelberg: Springer International Publishing.

3 Community Rights and Privacy in Indian Jurisprudence

A pervading theme in the Indian Constitution and its interpretation by Indian courts has been the conflicting nature of fundamental rights. While in most cases, individuals hold rights against the state, there are several instances of horizontal rights applicable against private actors, and more curiously, occasions where right holders are recognised groups, not individuals. Historically, the primary group that emerged as the bearer of group rights in India was religious communities, through the clear demarcation of public-private spheres by personal laws.[10] The primary focus in Indian jurisprudence on group rights has been on the identification of a group interest in protecting itself from external interference, rather than on laws governing groups that protect individuals from group-related harms.

Prior to the Supreme Court's judgment in *K S Puttaswamy and others v. Union of India*,[11] (*Puttaswamy*) it had not clearly established a right to decisional autonomy as a part of the right to privacy. The choice of individuals, such as women's reproductive rights,[12] dietary choices,[13] and the choice of gender,[14] had been recognised as integral to the right to privacy on various occasions, but Indian jurisprudence on this matter has been fraught with inconsistencies. It is in this regard that this judgment's clear and emphatic recognition of decisional autonomy is most significant. Three dec-

10 Flavia, Agnes. 2011. *Family Laws and Constitutional Claims.* Vol 1. New Delhi: Oxford University Press.
11 Case which established the right to privacy as a fundamental right in India.
12 *Suchita Srivastava* v *Chandigarh Administration,* AIR 2010 SC 235.
13 *Hinsa Virodhak Sangh* v *Mirzapur Moti Kuresh Jamat,* AIR 2008 SC 1892.
14 *National Legal Services Authority (NALSA)* v *Union of India,* AIR 2014 SC 1863.

ades earlier, in *T Sareetha v. Venkat Subbaiah (Sareetha),* the Andhra Pradesh High Court had held that coercing someone to live with their spouse violated their right to privacy, a judgment overturned by the Supreme Court soon after. The reasoning behind the High Court's judgment forms the basis of the Supreme Court's clear identification of decisional autonomy in *Puttaswamy*, and its centrality to the right of privacy.

This brings us to the key conflict between the individual right and group right to privacy.[15] The different dimensions of privacy often work together to protect the individual, but it bears asking which value must prevail over others when they are in conflict. First *Sareetha*, as a lone overturned High Court judgment, and decades later, *Puttaswamy*, with the full might of a nine-judge Supreme Court bench, clearly locate decisional autonomy and informed consent as the abiding principle from which other dimensions of privacy flow.

Much like decisional autonomy is a key principle for the right to privacy, group interests rely on the idea of self-determination, which is now recognised as a core tenet of public international law as well. While first formulated as a political principle during the mid-century decolonisation era, the internal aspects of self-determination have gained more importance in recent times. Shaw has described self-determination as "a people's pursuit of its political, economic, social and cultural development within the framework of an existing state."[16]

This backdrop necessitates discussion on two questions. First, how can communities be identified for the purpose of circumscribing benefits, and second, how can we identify individuals that belong to a part of that community?

15 It has been articulated precisely by Bhatia: "Does the Constitution treat groups as bearers of value in their own right, or does it view groups as instrumental to achieving individual fulfillment, and therefore guarantee group rights?" Bhatia, Gautam. 2016. "Freedom from Community: Individual Rights, Group Life, State Authority and Religious Freedom under the Indian Constitution." (https://ssrn.com/abstract=2739235).

16 In 1962, the United Nations General Assembly recognised the "right of peoples and nations to permanent sovereignty over their natural wealth and resources." It is a clear articulation not only of group interests but also a group's right to have its say over resources deemed crucial to the collective interests of the group. See Shaw, Malcolm. 2003. *International Law.* Fifth Edition. Cambridge: Cambridge University Press.

4 Implications of Sareetha and Puttaswamy

4.1 Prioritising Individual Rights over Other Interests

If we look at the full import of *Sareetha*, and *Puttaswamy*, as its jurisprudential successor, it must be accepted that while group rights and individual rights further each other, where they are in conflict, it is the individual rights which must prevail. *What implications must this have for community data rights?*

Much of the debate around community and non-personal data has to do with the privacy implications for anonymised data. So far, anonymised and pseudonymised data has existed in a regulatory vacuum between personal data protection laws and open data mandates. In a 2008 paper,[17] Narayanan and Shmatikov demonstrated issues that have emerged with anonymisation of data with the advancement in math and algorithm techniques. They argue that increasingly, the datasets we deal with are high-dimensional in nature, which allows greater scope for algorithms to correlate them with other databases, making anonymisation ineffective. Even so, "seemingly" anonymised datasets fall squarely outside the scope of personal data protection laws, putting individual rights at risk. Paul Ohm echoes these fears in his 2010 paper, dramatically titled, "Broken promises of privacy".[18] In Europe, the General Data Protection Regulation (GDPR) has wrestled with the legal question about anonymised data.[19] The GDPR, under Recital 26, adopts a risk-based approach to determine whether data is personal or not – an approach that has been endorsed by the British Information Commissioner's Office (ICO.) When risk assessment suggests that identifica-

17 See Narayanan, Arvind and Vitaly Shmatikov. 2008. "Robust De-anonymisation of Large Sparse Datasets." (https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf).

18 Ohm criticises the robust anonymisation assumption – the idea that anonymisation techniques could adequately change data so as to convert personal information into anonymised or aggregated information. The thrust of the robust anonymisation assumption was that these techniques could protect the privacy of the data subjects. The balance between personal data and open data policies has been upset by techniques which threaten to neutralise the effects of anonymisation. See Ohm, Paul. 2010. "Broken Promies of Privacy: Responding to the surprising failure of anonymization." UCLA Law Review 57, 1701.

19 Finck, Michele and Frank Pallas. 2020. "They who must not be identified-distinguishing personal from non-personal data under the GDPR." International Data Privacy Law: 10.

tion is "reasonably likely" to occur, anonymised data must receive GDPR protection in its entirety. The definitions of personal data adopted by the Article 29 Working Party of the European Union (now the European Data Protection Board) differs from that adopted by the national authorities of various EU countries,[20] and it adopts a higher threshold, arguing that anonymised personal data can only qualify as non-personal data when "irreversible identification" is present.[21]

The approach taken by the NPD Report in India opts for a midway between the contrasting European definitions. The report recognises the difficulties in irreversibly anonymising datasets, and instead of setting an impossible threshold for anonymisation, seeks to get around this problem by extending personal data and privacy rights even to anonymised data of an individual. While this may have been a regulatory strategy to circumvent the issue of the impossibility of irreversible de-identification, it, perhaps unwittingly, echoes *Sareetha* and *Puttaswamy* in clearly prioritising individual right of privacy in personal data over community rights or public interest in leveraging the economic or social value of datasets.

4.2 Nature of Individual and Collective Interests in Community Data

Unlike the prior conflicts between privacy and group interests, the group interests in community data revolve around the following factors:

a) Defining a community and its collective right to privacy;
b) A community interest in itself using community data for economic benefits, including through processing by other actors such as the state; and
c) An individual's right to privacy vis-a-vis the group.

Defining a Community and its Collective Right to Privacy

A group right to privacy is often described as arising from the failure of traditional personal data protection frameworks to protect the interests of the group.[22] This is so because big data and algorithmic analyses focus on the attributes of personal data, which involves bringing attention to the membership of individuals to specific groups.[23] Even where individuals may have provided informed consent, their data may be used to derive inference and make decisions about a group as a whole. Second, the granular amount of data available about individuals makes groups vulnerable by making more information discoverable about them. Finally, in many cases, even the data controller may be able to discern the correlations within and between groups identified by algorithms. As a result of these factors, despite the group's individual members having a working right to privacy, any protection to the group as a consequence of that right is rendered ineffective.

20 Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP216) 0829/14/EN, 11–12, 23–25.
21 Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP216) 0829/14/EN, 11–12, 23–25.
22 Floridi, Luciano. 2017. "Group Privacy: A Defence and an Interpretation." In Linnet Taylor, Luciano Floridi and Bart Van Der Sloot (eds). *Group Privacy: New Challenges of Data Technologies.* Heidelberg: Springer International Publishing.
23 Floridi, Luciano. 2017. "Group Privacy: A Defence and an Interpretation."

A community interest in using data about itself arises from the skewed nature of the data ownership paradigm. Through broadly drafted terms and conditions, it is usually the data collectors who exercise all economic rights over data generated. The shared nature of data created by the data subject's interaction with an interface created by a data holder makes the answer to the question "who is rightfully entitled to control over personal data" complex. Singh questions that if "individuals are supposed to [control] their data, why should data about groups/communities not, similarly, be [controlled] by the corresponding group/community?"[24] However, the key challenge here, as discussed above, is devising a process for identifying communities and groups, and deciding who does this identification.

Taking a cue from the decisional autonomy lens put forward in *Sareetha* and *Puttaswamy*, the decisions must be taken by both the community as a whole, and the individuals that make up the group. Neither the state nor any entity external to the community should make any decisions on the membership or formation of a community. Therefore, assuming, as the NPD report does, that all users of ride-sharing apps are a community, and accordingly casting data collected about them as "community non-personal data" is not a move that respects decisional autonomy. While individuals may have consented to some of their data being shared with a ride-sharing company, that cannot be taken as consent to being treated as part of a community of ride-sharing app users. This is a very different scenario from tightly knit communities such as farming communities or indigenous groups who might explicitly consent to being treated as a community. This consent could be gleaned from claims by the community as a whole over specific resources, declarations made by communities to be treated as one when it comes to exercising rights and obligations over unspecified issues, or pivots towards self-sustaining modes of governance and a call for non-interference from the state. In the absence of this consent, any decision attempting to box a group of people into a community violates the autonomy, and consequently the right to privacy of all the individuals that form a part of the group, and by extension that of the collective as a whole as well.

The framing we provide here does not apply to groups that do not self-identify as a community, but are treated as one due to algorithmic decision-making. For example, algorithmic decision-making may create groups of individuals residing in similar areas, and having similar income even though the individuals making up the group and the group itself does not identify as one. In this case, they would have a group right to privacy but as they do not have a collective interest in the data itself, would not qualify as a "community" for the purpose of circumscribing "community data."

A Community Interest in Using Data for Economic Benefits

This brings us to our next point of guidance from *Sareetha* and *Puttaswamy* on community data. Within this existing constitutional scheme, how must one think of the idea of community and its corresponding interests in data. While the groups in question are very different from the religious institution of matrimony discussed in

24 Singh, Parminder. 2019. "Community data in the draft e-commerce policy." Medianama (https://www.medianama.com/2019/03/223-community-data-in-the-draft-e-commerce-policy/).

Sareetha, the constitutional principles of decisional autonomy as well as equality were clearly established in the context of any group privacy by the judgment.

Let us first consider the nature of interest contemplated in the NPD Report. The report fashions data as a resource in which the community (and other stakeholders, such as the state) have a legitimate interest. This ownership model of data requires some examination. Unlike other kinds of property, data is non-rivalrous, and the idea of "privacy based on 'ownership' of an 'informational space' are metaphorical twice over."[25] This idea of data ownership lacks conceptual congruity, both legally and economically.[26]

Floridi advocates an "identity"-driven idea of group privacy, and "each individual person or group as constituted by his, her or its information, and hence by understanding a breach of an individual's informational privacy as a form of aggression towards that individual's identity."[27] This view finds a symmetrical echo in the ratio in *Sareetha* which states "any plausible definition of right to privacy is bound to take [...] human body as its first and most basic reference for control over personal identity."

A natural extension of this argument would entail that if privacy (individual and group) is to be seen as protection from aggressions towards the identity of the right holder, then we must answer our question about what constitutes the relevant group and what its protected "resources" are drawing from this understanding. Depending upon context, the relevant unit, and its informational space would both depend upon the identity sought to be protected. If individuals making up the community feel that the best way to protect their individual identity, and enforce associated rights would be through the community, then that would be the most appropriate mechanism.

While several communities may choose to process, interpret and manage all data they create, this may be a challenging task given that unlocking the real value of data requires sophisticated processing power, which communities might not possess.[28] Accordingly, the community may delegate its interest in certain datasets to the state to process it and extract value for the community's benefit, with explicit consent. This approach is fraught with danger, and magnifies the difficulties mentioned above multifold by taking away agency from individuals and groups, and instead handing it to the state. As a result, strict safeguards including a clear definition of the community, an agreement delineating the relationship between the state (or other bodies) and the community and an option to opt-out of this relationship should be provided to each community.

An Individual's Rights and Interest in Data vis-a-vis the Group

Even if the community and its associated rights and obligations are defined clearly, there are clear learnings for protecting individuals or sub-groups from *Sareetha*, which were discussed at some length by Justice Chandrachud in *Puttaswamy*, where he consid-

25 Floridi, Luciano. 2017. "Group Privacy: A Defence and an Interpretation."
26 Radin, M. J. 2002. "Incomplete Commodification in the Computerised World.". In Elkin-Koren, Niva and Neil Weinstock-Netanel (eds). *The Commodification of Information.* The Hague: Kluwer Law International.
27 Radin, M. J. 2002. "Incomplete Commodification in the Computerised World." 94.
28 Smith, Diane. 2016. "Governing data and data for governance: the everyday practice of indigenous sovereignty." In Talu Kakutai. *Indigenous Data Sovereignty: Toward an agenda.* Canberra: ANU Press.

ers the feminist critique of privacy.[29] The presumption challenged by Sareetha was that the individual privacy interests (decisional autonomy) are not necessarily synonymous with the group privacy interests (non-interference of state in religious matters), and in fact the group interests protected those in the most advantaged position within the group at the expense of others. Thus, protection provided to the group, in the case of personal laws, sought to provide protection to certain members of the group at the expense of others. This formulation of a group interest was justifiably deprioritised before an individual interest.

By extending the protection of personal data rights to anonymised data within any legal scheme that seeks to monetise data in the hands of a group or an entity other than the data principal, we see the first steps to avoid similar outcomes where community rights over data only advantages those who are most powerful within the community. A consistent application of this principle – that where the two are in conflict, the individual right to privacy will prevail over the group right to privacy or the group interest in data – can go a long way in thwarting the dangerous implications of community data, along with clear positive obligations to protect individual privacy. Therefore, the mere fact of them being a consensual (or otherwise) member of a group, does not result in them giving up the inviolable right to privacy.

29 "Many writers on feminism express concern over the use of privacy as a veneer for patriarchal domination and abuse of women. Patriarchal notions still prevail in several societies including our own and are used as a shield to violate core constitutional rights of women based on gender and autonomy. As a result, gender violence is often treated as a matter of 'family honour' resulting in the victim of violence suffering twice over – the physical and mental trauma of her dignity being violated and the perception that it has caused an affront to 'honour'. Privacy must not be utilised as a cover to conceal and assert patriarchal mindsets. Catherine MacKinnon in a 1989 publication titled 'Towards a Feminist Theory of the State' adverts to the dangers of privacy when it is used to cover up physical harm done to women by perpetrating their subjection. Yet, it must also be noticed that women have an inviolable interest in privacy. Privacy is the ultimate guarantee against violations caused by programmes not unknown to history, such as state imposed sterilization programmes or mandatory state imposed drug testing for women. The challenge in this area is to enable the state to take the violation of the dignity of women in the domestic sphere seriously while at the same time protecting the privacy entitlements of women grounded in the identity of gender and liberty." (para 140) in Chandrachud J.'s plurality opinion in Puttaswamy.

5 Takeaways for Europe

The concept of "community data," and its legal evolution has several connections to, and possible recommendations for the data governance ecosystem in Europe. First, the European Strategy for Data, a draft of which was published in March 2020, seeks to turn Europe into "a society empowered by data to make better decisions – in business and the public sector"[30] and recognises "data as the lifeblood of economic development."[31] It also talks up the significance of technological sovereignty in "key enabling technologies and infrastructures for the data economy." Further, as per reports dated 30 September 2020, a future version of the Europe Digital Services Act will mandate large technology companies to share data with their rivals.[32] Just as India is grappling now with rights over data, the fundamental question the data strategy needs to ask is: whose sovereignty and for whom? The European Strategy for Data goes on to suggest that data pools may be centralised or distributed, but it does not clarify how the benefits of this data can be distributed across communities and individuals. This is where a recognition of community data, which addresses the gap we identified in the Indian framework and accordingly prioritises decisional autonomy, will result in the most equitable distribution of rights and resources across communities in Europe.

Like Chapter 4 of India's 2019 Economic Survey, the European Strategy for Data also invokes the concept of "data as a public good." It argues that there is not enough data available for reuse that can foster innovation, particularly those involving the use of artificial intelligence. It underscores this point by stating that private sector organisations do not share enough data with each other or make available these datasets for use by the public sector in order to improve evidence-driven policy-making and public services. Therefore, the strategy recommends the creation of "common European data spaces" in strategic sectors, and domains of public interest. However, like India's Economic Survey (2019), the strategy ignores community, and by extension individual rights and interests in public datasets. While it eloquently

30 European Commission. 2020. "A European Strategy for Data."(https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf). 1.

31 European Commission. 2020. "A European Strategy for Data."

32 "(The likes of Amazon and Google 'shall not use data collected on the platform ... for[their] own commercial activities ... unless they [make it] accessible to business users active in the same commercial activities,' said the draft.)" Espinoza, Javier. 2020. "Brussels drafts rules to force Big Tech to share data." Financial Times (https://www.ft.com/content/1773edd6-7f1d-4290-93b6-05965a4ff0db).

bats for sharing of privately held data among companies and with the government, this framing does not account for the fact that companies are not the primary creators of this data.[33] With this in mind, the European Strategy for Data should work with communities, which might include farming collectives, religious, ethnic and sexual minorities, indigenous populations and migrants to identify individual and collective interests in data they create. This must be done while considering the rights individuals have vis-a-vis the group stemming from the construct of decisional autonomy, including rights of redress, opting out, and enforcement of individual rights. Bearing this in mind, we recommend a three-pronged principled approach to protecting both individual and group interests in data as Europe looks to unlock the economic potential of data:

1. Rights and interest in datasets must be accorded to communities who self-identify as one, and establish such rights and interests.
2. Individuals who are treated as being part of the community must consent to being part of the community, and to their data being treated as "community data."
3. The community as a whole consent to third parties – either state or non-state actors – processing community data on their behalf.

As discussed above, despite the seemingly straightforward distinction between personal and non-personal data in Recital 26 of the GDPR, several question marks remain over the practical ramifications of this distinction and the challenges of anonymisation and pseudonymisation in terms of identifiability and consequently, violation of privacy. The approach of the NPD report, which inadvertently uses the framing of decisional autonomy, could help preserve individual rights over data even if the practical implications of the Recital 26 distinction are not resolved.

Like with India, Europe's decision-making and approach to data governance is a product of negotiations between companies, the regulator, and the consumer, and much like with the GDPR itself, the strategic interests of each stakeholder group will drive future negotiations and approaches. However, it is clear that Europe wants to chart a citizen-centric approach in its approach to digital governance. Not all the policy measures coming out of Europe are perfect, as we have discussed in this paper as well, but Europe has demonstrated that it is willing to listen to stakeholders both within and outside Europe before finalising any approach. The uniqueness of a European way to digital governance was captured most poignantly by European Council President Charles Michel in a speech delivered on 29 September 2020, where he stated[34]:

> **"Between the American model of 'business above all', and the Chinese state-controlling authoritarian model, there is plenty of room for an attractive and human-centred model."**

33 "Data generated by the public sector as well as the value created should be available for the common good by ensuring, including through preferential access, that these data are used by researchers, other public institutions, SMEs or start-ups. Data from the private sector can also make a significant contribution as public goods. The use of aggregated and anonymised social media data can for example be an effective way of complimenting the reports of general practitioners in case of an epidemic." See Economic Survey of India. 2019. 6–7.

34 European Council. 2020. "The digital in a fractious world: Europe's way-speech by President Charles Michel at the FT-ETNO Forum." Press Release, 29 September. (https://www.consilium.europa.eu/en/press/press-releases/2020/09/29/the-digital-in-a-fractious-world-europe-s-way-speech-by-president-charles-michel-at-the-ft-etno-forum/).

6 Conclusion

Establishing a balance between economic value and collective or individual rights and interests is a challenge that countries both in Europe and Asia continue to wrestle with. As stakeholders and countries join the data sovereignty bandwagon, and aspire to utilise data for citizens' interest, a sound theoretical conception of collective interests in data that adequately respects both community and individual interests is the need of the hour. Community data could be this theoretical framework, although at present it is plagued by several lacunae, most notably a lack of guidance on identifying communities that have rights or interests in data, and the individuals that form it.

By studying unique Indian case law on privacy that deals with the question of individual and group rights, we find that decisional autonomy is the fulcrum of privacy jurisprudence, and thus should be the edifice for any policy framework. We find that in a case of conflict between individual and group rights, Indian jurisprudence finds that individual rights must prevail. The NPD report that provides the most concrete framing of community data to date unwittingly adopts this approach, and extends privacy and personal data protection rights to anonymised datasets that might be treated as community data. This adopts a middle ground between two regulatory approaches currently being discussed in Europe on anonymisation – between Article 29's threshold of irreversible identification, and Recital 26 GDPR's risk-based approach, which we discussed in Section III of this paper.

The concept of community data has rich value for Europe, which is beginning to shape its own strategy for leveraging economic benefits from data. Providing communities with adequate rights and interests while also prioritising individual rights is very much in line with human rights principles espoused by Europe, and endorsed in the GDPR. The principles we identified for governing community data are, at this stage, still abstract. Future research must focus on case studies through which this theoretical innovation can be piloted. These case studies would likely reveal further cases of conflict with these principles, which an overarching governance framework must address. If conceptualised effectively, community data could be the policy innovation that charts out the path for the next digital decade.

Authors

Amber Sinha

Amber Sinha is the Executive Director of the Centre for Internet and Society, India (CIS). At CIS, he has led projects on privacy, digital identity, artificial intelligence and misinformation. Amber's research has been cited with appreciation by the Supreme Court of India. His first book, The Networked Public, was released in 2019. Amber studied law and humanities at National Law School of India University, Bangalore.

Arindrajit Basu

Arindrajit Basu is a Research Manager at the Centre for Internet and Society, India, where he focuses on the geopolitics and constitutionality of emerging technologies. He is a lawyer by training and holds a BA LLB (Hons) degree from the National University of Juridical Sciences, Kolkata, and an LLM in public international law from the University of Cambridge, UK.

References

A Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP216) 0829/14/EN.

B Basu, Arindrajit. 2019. "India's role in global cyber policy formulation." Lawfare. (https://www.lawfareblog.com/indias-role-global-cyber-policy-formulation).

Bhatia, Gautam. 2016. "Freedom from Community: Individual Rights, Group Life, State Authority and Religious Freedom under the Indian Constitution." (https://ssrn.com/abstract=2739235).

C Couldry, Nick and Ulijies Mejias. 2018. "Data Colonialism: rethinking big data's relation to the contemporary subject." Television and new Media. (https://eprints.lse.ac.uk/89511/1/Couldry_Data-colonialism_Accepted.pdf).

E European Commission. 2020. "A European Strategy for Data." (https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf).

European Council. 2020. "The digital in a fractious world: Europe's way-speech by President Charles Michel at the FT-ETNO Forum." Press Release, 29 September. (https://www.consilium.europa.eu/en/press/press-releases/2020/09/29/the-digital-in-a-fractious-world-europe-s-way-speech-by-president-charles-michel-at-the-ft-etno-forum/).

Espinoza, Javier. 2020. "Brussels drafts rules to force Big Tech to share data." Financial Times. (https://www.ft.com/content/1773edd6-7f1d-4290-93b6-05965a4ff0db).

F Finck, Michele and Frank Pallas. 2020. "They who must not be identified-distinguishing personal from non-personal data under the GDPR." International Data Privacy Law.

Flavia, Agnes. 2011. *Family Laws and Constitutional Claims*. Vol 1. New Delhi: Oxford University Press.

Floridi, Luciano. 2017. "Group Privacy: A Defence and an Interpretation." In Linnet Taylor, Luciano Floridi and Bart Van Der Sloot (eds). *Group Privacy: New Challenges of Data Technologies.* Heidelberg: Springer International Publishing.

I Indian Department for Promotion of Industry and Internal Trade. 2019. "Draft E-Commerce Policy." (https://dipp.gov.in/sites/default/files/DraftNational_e-commerce_Policy_23February2019.pdf).

Indian Ministry of Finance. 2019. “Economic Survey of India.” India Budget. (https://www.indiabudget.gov.in/budget2019-20/economicsurvey/doc/echapter.pdf).

K Kammoruieh, Lanah. 2017. “Group Privacy in the age of big data.” In Linnet Taylor, Luciano Floridi, and Bart Van Der Sloot (eds.) *Group Privacy: New Challenges of Data Technologies.* Heidelberg: Springer International Publishing.

N Narayanan, Arvind and Vitaly Shmatikov. 2008. “Robust De-anonymisation of Large Sparse Datasets.” (https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf).

O Ohm, Paul. 2010. “ Broken Promies of Privacy: Responding to the surprising failure of anonymization.” UCLA Law Review 57, 1701.

R Radin, M. J. 2002. “Incomplete Commodification in the Computerised World.” In Niva Elkin-Koren, Niva and Neil Weinstock- Netanel (eds). *The Commodification of Information*. The Hague: Kluwer Law International.

S Shaw, Malcolm. 2003. *International Law*. Fifth Edition. Cambridge: Cambridge University Press.

Singh, Parminder. 2019. “Community data in the draft e-commerce policy.” Medianama (https://www.medianama.com/2019/03/223-community-data-in-the-draft-e-commerce-policy/).

Smith, Diane. 2016. “Governing data and data for governance: the everyday practice of indigenous sovereignty.” In Talu Kakutai. *Indigenous Data Sovereignty: Toward an agenda*. Canberra: ANU Press.

SriKrishna Committee. 2018. “A Free and Fair Digital Economy.” (https://www.meity.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf).