International Journal of Law And Information Technology, 2026, **34**, eaaf025
https://doi.org/10.1093/ijlit/eaaf025
Advance access publication 28 January 2026
Article

OXFORD

The emerging framework for non-personal data protection in India: perils, promises, and lessons for the developing world

Neha Mishra[1*], Binit Agarwal[2,]

[1]Department of International Law, Geneva Graduate Institute, Chemin Eugène-Rigot 2, 1202 Genève, Switzerland
[2]Lucio AI, Indiranagar, Bangalore, India

ABSTRACT

This paper focuses on the emerging framework for non-personal data protection in India, analysing its potential benefits and drawbacks. The Indian government has used a diverse range of policy instruments to develop a techno-political framework for non-personal data regulation, wherein non-personal data are predominantly viewed as a national economic resource that must be shared to promote domestic players and reinforce India's digital and data sovereignty. Despite several of its innovative features, including instituting nationwide platforms for non-personal data sharing, this framework lacks clarity regarding data intermediaries' roles, provides insufficient privacy protections, and creates several business uncertainties. India's experience reflects developing countries' challenges and dilemmas in regulating non-personal data, especially balancing economic goals with privacy concerns and managing diverse stakeholder interests. We conclude by highlighting the importance of a rights-based approach with clear accountability mechanisms and streamlined policy/legal frameworks to ensure effective and equitable non-personal data regulation in developing countries.

Keywords non-personal data, data governance, privacy, India, data regulation, data commons, data trustee, Data Empowerment and Protection Architecture, Open Network for Digital Commerce

INTRODUCTION

Non-personal data, often defined as data not directly linked to an individual,[1] is the new frontier of data regulation. As governments worldwide embrace data as a strategic national asset,

* Department of International Law, Geneva Graduate Institute, Chemin Eugène-Rigot 2, 1202 Genève, Switzerland. Email: neha.mishra@graduateinstitute.ch

© The Author(s) 2026. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

non-personal data have emerged as the new subject and object of regulatory interest. While the European Union (EU) has taken the lead on non-personal data regulation with the Data Governance Act already applicable,[2] and the Data Act coming into force,[3] other countries, including China, Finland, the UK, and India, are also dabbling with non-personal data regulation. Regulators are particularly attracted to regulating non-personal data as a tool to promote data-driven innovation because facilitating access to non-personal data appears less burdensome than personal data.

This paper unpacks the political economy of non-personal data regulation, using the emerging regulatory framework on non-personal data in India as a case study, to understand key opportunities, challenges, and lessons for developing countries experimenting with non-personal data regulation. Alongside the EU, India is one of the early advocates of non-personal data regulation. However, unlike the EU, the regulatory framework in India has, to date, largely remained a work-in-progress embedded in government policies rather than binding legislation. We focus on India for various reasons. First, India is a giant data economy, accounting for one-fifth of the data generated worldwide.[4] Second, the Indian government is the biggest collector of data within the country.[5] Third, being the world's largest democracy and a developing country, the Indian example can help understand regulatory challenges for other developing countries in Asia and beyond.

The global trends in data regulation, more broadly, suggest two strong pulls for a large majority of developing countries. First, several developing countries are influenced heavily by the 'digital empires' (as Bradford terms the USA, China, and the EU)[6] in developing their domestic digital regulatory models due to historical and market dependencies. In particular, with respect to data regulation, the EU has had the most prominent impact, characterized as the 'Brussels Effect' by Bradford.[7] For example, many countries across the world have adopted rules similar to the General Data Protection Regulation or 'GDPR' in their domestic laws.[8] Similar effects are possible in non-personal data regulation. We provide some examples of the potential 'influencer' effect of the EU in non-personal data regulation, but this paper does not entail a point-to-point comparative study of India and the EU.

Second, many developing economies aim to ward off data colonialism, understood as bringing together 'predatory extractive practices of historical colonialism with the abstract quantification

[1] Michèle Finck and Frank Pallas, 'They Who Must not be Identified—Distinguishing Personal from Non-personal Data under the GDPR' (2020) 10 IDPL 11.

[2] Commission, 'Proposal for a Regulation of the European Parliament and of the Council on European Data Governance' COM (2020) 767 final (Data Governance Act).

[3] Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 [2023] OJ L 2023/2854 (Data Act).

[4] Mehul Reuben Das, 'India Generates 20% of the World's Data, but Only Has 2% of Data Centres: Intel's Santhosh Viswanathan' (*Firstpost*, 16 May 2024) <https://www.firstpost.com/tech/india-generates-20-per-cent-of-the-worlds-data-but-only-has-2-per-cent-of-data-centres-intels-santhosh-viswanathan-13771439.html> accessed 29 April 2025.

[5] Nikhil Menon, 'National Sample Survey: How India Taught the World the Art of Collecting Data' (BBC, 29 June 2022) <https://www.bbc.com/news/world-asia-india-61870699> accessed 29 April 2025.

[6] Anu Bradford, *Digital Empires: The Global Battle to Regulate Technology* (Oxford University Press 2023).

[7] Anu Bradford, *The Brussels Effect: How the European Union Rules the World* (Oxford University Press 2020); See also Thomas Streinz, 'The Evolution of European Data Law' in Paul Craig and Gráinne de Búrca (eds), *The Evolution of EU Law* (3rd edn, Oxford University Press 2021); Francesco Vogelezang, 'The EU Data Governance Act: A Tale of Two Cities for the Datasphere' (Datasphere Initiative, 20 April 2022) <https://www.thedatasphere.org/news/the-eu-data-governance-act-a-tale-of-two-cities-for-the-datasphere/> accessed 29 April 2025. The Brussels Effect is not without its criticisms. See, for instance, Cara Mannion, 'Data Imperialism: The GDPR's Disastrous Impact on Africa's E-Commerce Markets' (2020) 53 Vanderbilt Journal of Transnational Law 685; Natalia Uribe, 'Is the GDPR a Form of European Imperialism?' (*EmilDAI*, 5 July 2023) <https://emildai.eu/is-the-gdpr-a-form-of-european-imperialism/> accessed 29 April 2025.

[8] Annegret Bendiek and Isabella Stuerzer, 'The Brussels Effect, European Regulatory Power and Political Capital: Evidence for Mutually Reinforcing Internal and External Dimensions of the Brussels Effect from the European Digital Policy Debate' (2023) 2 Digital Society 5.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

methods of computing'.[9] In simple words, developing countries want to ensure that digital data collected or generated within their country is used for the economic benefit of their economy and people, instead of being extracted and manipulated for profit by Big Tech companies based in countries such as the USA and China.[10] To achieve this goal, governments want to release data siloed in the data centres of big (often, foreign) companies to support their domestic digital players.

Multitudes of economic and political motivations drive non-personal data regulation. The next section, **Political Economy of Non-Personal Data and Its Regulatory Implications**, highlights the unclear dividing line between personal and non-personal data and then outlines the three main ways in which non-personal data can be conceptualized (ie as a national resource, as property, and as commons).

The next section, **Exploring India's Regulatory Framework for Non-Personal Data**, delves deeper into the existing policy framework of non-personal data in India as well as the techno-political framework for data sharing.[11] In evaluating the Indian framework on non-personal data regulation, we also examine overlaps with other areas of regulation, including personal data protection and competition law. This section argues that non-personal data is predominantly characterized as a national asset of high economic importance in India, critical for making the country a leading digital innovator, breaking the vicious cycle of data dependency on a handful of foreign companies/powers, and asserting Indian data/digital sovereignty on a global stage.[12] Further, this section argues that despite well-intentioned motives, several weaknesses exist in India's non-personal data framework, including the unclear role of data intermediaries, and the inattention towards the repercussions of the excessive vesting of power in the government to mandate data access and sharing. Finally, the potential overlaps/conflicts with domestic laws on data protection, intellectual property, and competition remain under-explored.

The section, **Learnings from India's Experience: How Can Developing Countries Regulate Non-Personal Data**, builds on the lessons learnt from the Indian experience to outline the potential challenges faced by developing countries in regulating non-personal data. It particularly highlights the need to balance economic ambitions with critical public policy concerns while managing varied stakeholder interests. This balancing exercise becomes even more complex due to the limited regulatory capacity of developing countries, inadequate experience in implementing data protection laws, and deficient digital and data infrastructure.

[9] Nick Couldry and Ulises Mejias 'Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject' (2019) 20 Television and New Media 336, 337. See generally Ulises A Mejias and Nick Couldry, *Data Grab: The New Colonialism of Big Tech and How to Fight Back* (University of Chicago Press 2024).

[10] Parminder Jeet Singh, 'Why Owning their National Data is Important for Developing Countries' (*IT for Change*, March 2019) <https://itforchange.net/index.php/why_owning_their_national_data_is_impt> accessed 29 April 2025; Neha Mishra, 'Data Governance and Digital Trade in India: Losing Sight of the Forest for the Trees?' in Anupam Chander and Haochen Sun (eds), *Data Sovereignty: From the Digital Silk Road to the Return of the State* (Oxford University Press 2023); 'AU Data Policy Framework' (2022) 47 <https://files.core.ac.uk/download/582653563.pdf> accessed 29 April 2025; Amber Sinha and Arindrajit Basu, 'The Politics of India's Data Protection Ecosystem' (*EPW Engage*, 27 December 2019) <https://www.epw.in/engage/article/politics-indias-data-protection-ecosystem> accessed 29 April 2025.

[11] Cristian Alonso et al, 'Stacking up the Benefits: Lessons from India's Digital Journey' (2023) International Monetary Fund Working Paper No. 2023/078 <https://www.imf.org/en/Publications/WP/Issues/2023/03/31/Stacking-up-the-Benefits-Lessons-from-Indias-Digital-Journey-531692> accessed 29 April 2025; 'Data' (IndiaStack) <https://indiastack.org/data.html> accessed 29 April 2025.

[12] Arindrajit Basu, 'Sovereignty in a 'Datafied' World' (2021) Observer Research Foundation Issue Brief <https://www.orfonline.org/research/sovereignty-in-a-datafied-world> accessed 29 April 2025. See also Upasana Sharma and Shreya Raman, 'AI for All, Beyond the Global North: India's Opportunity?' (*Carnegie India*, 27 November 2023) <https://carnegieindia.org/2023/11/27/ai-for-all-beyond-global-north-india-s-opportunity-pub-91110> accessed 29 April 2025; Carnegie India, 'Digital Public Intelligence: What Comes Next for DPIs & AI in India?' by Nandan Nilekani' (6 December 2023) <https://www.youtube.com/watch?v=Lvzlt-AryV4&list=PLeXQMWQXRkJXXnHrzYlBhUvO1_vPjCYKz&index=7> accessed 29 April 2025 (Nilekani is one of the architects of India's data policies); Ministry of Electronics and Information Technology ('MEITY'), 'Proposed Digital India Act, 2023' (Proposed Digital India Act) <https://www.meity.gov.in/static/uploads/2024/06/2bf1f0e9f04e6fb4f8fef35e82c42aa5.pdf> accessed 29 April 2025.

Downloaded from https://academic.oup.com/jilit/article/doi/10.1093/jilit/eaaf025/8443255 by guest on 29 January 2026

POLITICAL ECONOMY OF NON-PERSONAL DATA AND ITS REGULATORY IMPLICATIONS

To understand the complex nature of non-personal data regulation, we first outline the legal and policy uncertainty created by the thin dividing line between personal and non-personal data. We then explore possible policy rationales behind non-personal data regulation and delve deeper into three possible conceptualizations of non-personal data, using examples of three different kinds of regulatory frameworks: mandatory data-sharing frameworks, property-based frameworks, and commons-based frameworks.

Differentiating non-personal data and personal data

Non-personal data are defined in an exclusionary manner as anything that is not covered within the scope of 'personal data'. For instance, the EU framework for the free flow of non-personal data applies to the 'processing of electronic data other than personal data'.[13] Similarly, in India, the Expert Committee on Non-Personal Data defined personal data as 'data (that) is not "personal data"' under Indian personal data protection law or 'the data is without any personally identifiable information'.[14] It provides examples of two categories: where data does not relate to an identified or identifiable person (eg weather data) and where personal data has been anonymized so that individuals are no longer re-identifiable in the dataset.[15] Both the EU and the Indian framework also specify that if a dataset contains both personal and non-personal data, the personal data protection law would apply.[16] Thus, the foundation of non-personal data regulation is the clear distinction between personal and non-personal data.[17]

In reality, the distinction between non-personal and personal data is often unclear and contextual,[18] posing legal uncertainties when devising specific regulatory frameworks for non-personal data. First, non-personal data and personal data are often inextricably linked in mixed datasets (commonly collected by most digital systems). As such, frameworks enabling the flow of non-personal data can lead to unintended disclosure of sensitive personal information, resulting in a conflict between personal data protection law and non-personal data regulation, and confusion regarding the regulatory body responsible for dealing with relevant issues.[19] Second, the distinction between personal and non-personal data is also unrealistic because anonymization or pseudonymization techniques are neither foolproof, nor do most companies collecting mixed datasets have strong incentives to separate personal and non-personal data after collection.[20]

[13] Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union [2018] OJ L 303 (EU Framework for Non-Personal Data), art 2(1).

[14] MEITY, 'Report by the Committee of Experts on Non-Personal Data Governance Framework' (December 2020) ('2020 Report') para 1 <https://ourgovdotin.wordpress.com/wp-content/uploads/2020/12/revised-report-kris-gopalakrishnan-committee-report-on-non-personal-data-governance-framework.pdf> accessed 29 April 2025.

[15] ibid para 1.

[16] EU Framework for Non-Personal Data (n 15) art 2(2); 2020 Report (n 16) para 5.1.

[17] Iana Rezlauf, 'EU Framework for Handling Big Datasets Mixed of Personal and Non-Personal Data' (2020) 21 Comput Law Rev Int 7. For example, the determination as to whether the PDP framework or the NPD framework applies in India to a specific kind of data would be determined by the identifiability of that data, see 2020 Report (n 16) para 5.2.

[18] For examples, see Finck and Pallas (n 3); Maria Lillà Montagnani and Mark Verstraete, 'What Makes Data Personal?' (2023) 56 UC Davis L Rev 93. Yves-Alexandre De Montjoye and others, 'Unique in the Shopping Mall: On the Reidentifiability of Credit Card Metadata' (2015) 347 Science 536.

[19] See De Montjoye and others, ibid.

[20] Astha Kapoor and Amrita Nanda, 'Non-Personal Data Sharing: Potential, Pathways and Problems' (2021) 9 CSI Transactions on ICT 165, 166; Nicoleta Cherciu and Teodor Chirvase, 'Non-Personal Data Processing-Why Should We Take It Personally?' (2020) Eur J Privacy L & Tech 183; Emily M Weitzenboeck and others, 'The GDPR and Unstructured Data: Is Anonymization Possible?' (2022) 12 IDPL 184; Claudia Irti, 'Personal Data, Non-Personal Data, Anonymised Data, Pseudonymised Data, De-identified Data' in Roberto Senigaglia, Claudia Irti and Alessandro Bernes (eds), *Privacy and Data Protection in Software Services* (Springer 2022) 49–57.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

Finally, aggregated data can result in group privacy concerns, which are unregulated in existing legal frameworks.[21] For example, when aggregated datasets created by anonymizing personal data are made available for broad usage, companies and governments can use a broad range of analytical techniques (such as observing behaviours or preferences of groups) to identify, profile and target groups of people, even though individual identifiers may have been removed from the dataset.[22] Thus, a possibility of the breach of privacy rights at a group level continues to exist.[23] This can potentially lead to targeted surveillance, discriminatory practices based on ethnicity, religion, or political affiliation, and exploitative business practices, ultimately impacting individual rights.[24]

The failure to regulate the overbroad and uncontrolled use of non-personal data can manifest as privacy risks in various ways. First, different non-personal datasets, particularly if they contain anonymized personal data, can be combined with other datasets (including geolocation and metadata) to re-identify individuals and draw insights about them.[25] Particularly, with easy access to advanced generative artificial intelligence (AI) tools, reidentification strategies are increasingly becoming feasible,[26] blurring the line between personal and non-personal data.[27] Second, maintaining the inherently contextual distinction between personal and non-personal data also depends on external factors such as the available technology and the sophistication of the techniques adopted by the data processor.[28] For instance, while general traffic data, public data on ethnic features, food consumption patterns, or linguistic data may not be information identifying a person for laypersons, these same datasets can be sufficient for government bodies or large companies, with access to powerful technologies and other complementary datasets, to identify individuals or at least the groups to which they belong. Third, as already mentioned above, non-personal datasets can be used for drawing insights and inferences at a group level, which can affect the privacy of a group as a whole, ultimately compromising the privacy of the individuals forming part of that group.

Approaches to non-personal data regulation: ideas, institutions and actors

The approaches to regulating non-personal data depart significantly from personal data due to fundamentally different regulatory bases. While the key goal of personal data protection law is

[21] Varunavi Bangia, 'Why "Group Privacy" Should Be Recognised, and How "Non-personal" Data Becomes a Regulatory Blindspot' (*Medianama*, 13 July 2022) <https://www.medianama.com/2022/07/223-essay-group-privacy-non-personal-data-protection-regulations-2/> accessed 29 April 2025.

[22] Brent Mittelstadt, 'From Individual to Group Privacy in Big Data Analytics' (2017) 30 Philosophy & Technology 475; Paul Ohm, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' (2010) 57 UCLA L Rev 1701; Michele Loi and Markus Christen, 'Two Concepts of Group Privacy' (2020) 33 Philos Technol *321; Irti* (n 23).

[23] ibid.

[24] See, for instance, US Government Accountability Office, 'Consumer Data: Increasing Use Poses Risks to Privacy' (13 September 2022) <https://www.gao.gov/products/gao-22-106096> accessed 29 April 2025; Francesca Bosco and others, 'National Data Protection Authorities' Views on Profiling' in Niklas Creemers and others (eds), *Profiling Technologies in Practice: Applications and Impact on Fundamental Rights and Values* (Wolf Legal Publishers 2015) 22–3.

[25] Nils Gruschka and others, 'Privacy Issues and Data Protection in Big Data: A Case Study Analysis under GDPR' (*arXiv*, 22 November 2018) <https://arxiv.org/pdf/1811.08531> accessed 29 April 2025; Boris Lubarsky, 'Re-Identification of "Anonymized" Data' (2017) 1 Georgetown Law Technology Review 202; C Christine Porter, 'De-identified Data and Third Party Data Mining: The Risk of Re-identification of Personal Information' (2008) 5 Shidler J L Com Tech 3; Kostas Drakonakis and others, 'Please Forget Where I Was Last Summer: The Privacy Risks of Public Location (Meta)Data' (Network and Distributed System Security Symposium, 24–27 February 2019).

[26] See, for instance, Alex Nyffenegger and others, 'Anonymity at Risk? Assessing Re-Identification Capabilities of Large Language Models' (*arXiv*, 19 May 2024) <https://arxiv.org/abs/2308.11103> accessed 29 April 2025.

[27] Luc Rocher and others, 'Estimating the Success of Re-Identifications in Incomplete Datasets Using Generative Models' (2019) 10 Nat Commun 3069.

[28] Emmanuel Salami, 'Balancing Competing Interests in the Reidentification of AI-Generated Data' (2022) 8 Eur Data Prot Law Rev 362; Khaled El Emam, 'Precaution, Ethics and Risk: Perspectives on Regulating Non-Identifiable Data' (*IAPP*, 24 May 2022) <https://iapp.org/news/a/precaution-ethics-and-risk-perspectives-on-regulating-non-identifiable-data/> accessed 29 April 2025.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

the protection of individual privacy, non-personal data regulation can have multiple goals including enabling data sharing for economic development and boosting innovation, regulating access over non-personal data (especially in regulating Big Tech companies that act as data monopolies), establishing national control over critical datasets, creating relevant property rights for non-personal data, protection of community rights over data, and even creating a framework for collective or group privacy.[29]

Policymakers increasingly see non-personal data as having diverse public uses such as for urban planning and infrastructure development, medical services improvement, educational and research purposes, and improved governance.[30] Local governments could leverage this data to optimize traffic flow, alleviate congestion, enhance public transportation systems, and deal with urban planning woes. Local businesses, if supplied with this data (especially when facilitated by government regulation), could utilize it to make informed decisions on setting up new establishments, tailoring services to improve profitability, and developing customized digital products addressing local community needs. These utopian objectives are often canvassed as a key driver of non-personal data regulatory frameworks.[31]

Simultaneously, non-personal data can be critical for strengthening governmental control. While several personal data protection laws make it difficult for governments to access personal data for mass surveillance or citizen policing, non-personal data remains a grey area. Governments can thus use non-personal data to augment traditional state powers such as conducting criminal investigations, precise implementation of tax laws, better targeting of laws, and improving the functioning of courts.[32] Thus, governmental use of non-personal data is a double-edged sword; on the one hand, it can improve public functions, while on the other, it can facilitate more control for authoritarian governments.[33]

These broader goals, either based on economic, political, or sovereignty-related reasons, have translated into three dominant approaches to non-personal data regulation: *non-personal data as a national resource*; *non-personal data as property*; and *non-personal data as commons*. In practice, countries may combine different elements of these approaches.

Non-personal data as a national resource

A dominant narrative on non-personal data (both private and public) is that it is a valuable national resource to be shared domestically for national growth (as is the case with India's emerging framework, discussed in detail later). This narrative leads governments to adopt mandatory measures facilitating the sharing of non-personal data, especially among market participants, for generating 'public benefit'.[34] Further, governments open up troves of public non-personal data for industry use (either on a commercial basis or free). Often, such mandatory sharing of data provides weak consideration to intellectual property rights and competition concerns. This inattention results primarily from the overstated focus of policymakers

[29] See, for instance, EU Framework for Non-Personal Data (n 15); 2020 Report (n 16).

[30] Linnet Taylor, 'The Ethics of Big Data as a Public Good: Which Public? Whose Good?' (2016) 374 Philos Trans A Math Phys Eng Sci 20160126.

[31] Lina María Díaz Vera, 'Non-Personal Data Regulation – A Latin American Perspective' (2023) 72(1) GRUR International 37–53; Vikas Kathuria, 'Comparative Assessment of Non-Personal Data Access Frameworks' in Andelka M Phillips, Edina Harbinja and Claudio Lombardi (eds), *Technologies, Law and Society* (forthcoming, Edinburgh University Press).

[32] Dennis Broeders and others, 'Big Data and Security Policies: Towards a Framework for Regulating the Phases of Analytics and Use of Big Data' (2017) 33 CLSR 309–23.

[33] See generally Mihir Kaulgud, 'India's Approach to Data Decolonization: Moving Away from the "Data as Resource" Metaphor' (*Social and Political Research Foundation*, October 2022) <https://sprf.in/wp-content/uploads/2022/10/data-decol-IB.pdf> accessed 29 April 2025.

[34] In fact, as discussed below in the Section on Emerging Policy Framework for Non-Personal Data, this is the language used by the Indian government in various policy frameworks on non-personal data sharing.

Downloaded from https://academic.oup.com/jilit/article/doi/10.1093/jilit/eaaf025/8443255 by guest on 29 January 2026

on the perceived economic benefits of mandatory access to non-personal data, while ignoring how mandatory data sharing can undermine the complex nature of data and data markets. For instance, mandatory data sharing, wherein data are treated as a public utility, sidelines the enormous investments made by private companies in generating various industrial and anonymized datasets.[35] It may also further ignore proprietary protections under intellectual property law, such as trade secrets or database rights, for instance.[36] Finally, the utopian view that non-personal data are a national resource belonging to the nation (and by extension, the State), renders the interests of the individuals (who may want privacy) as secondary to the national growth objective driving mandatory data sharing.[37]

Some experts have even argued that mandatory data sharing amounts to nationalization of data.[38] Data is often compared to other natural resources and, hence, the calls for nationalizing data; however, this analogy is unsuitable in the digital context. First, without trusted intermediaries and a technologically robust platform for data sharing, data misuse is likely to increase. Second, non-personal data can affect legitimate commercial interests, especially if market participants are compelled to share free or for minimal compensation. Companies may even make a deliberate business choice to avoid compliance with a mandatory sharing requirement by not investing in technologies that help remove personal data from mixed datasets.

Non-personal data as property

Some scholars have proposed that establishing clear property rights can be helpful for non-personal data regulation.[39] While this paper does not discuss how intellectual property law applies to non-personal data,[40] this section highlights some high-level perspectives. A property-based framework for non-personal data has some obvious advantages, including better economic incentives for companies to segregate and share non-personal data. It can incentivize smaller tech companies to share data in an economically sustainable manner through licensing and facilitate clear liability rules for companies failing to adopt robust anonymization techniques.[41]

[35] See, for instance, Jeffrey Dastin, 'Amazon CEO Sets Out AI Investment Mission in Annual Shareholder Letter' (*Reuters*, 10 April 2025) <https://www.reuters.com/technology/amazon-ceo-sets-out-ai-investment-mission-annual-shareholder-letter-2025-04-10/> accessed 29 April 2025; 'Alphabet CEO Reaffirms Planned $75 Billion Capital Spending in 2025' (*Reuters*, 9 April 2025) <https://www.reuters.com/technology/alphabet-ceo-reaffirms-planned-75-billion-capital-spending-2025-2025-04-09/> accessed 29 April 2025.

[36] See generally Simon Forge, 'Optimal Scope for Free-Flow of Non-Personal Data in Europe' (*European Parliament Briefing*, April 2018) <https://www.europarl.europa.eu/RegData/etudes/BRIE/2018/618988/IPOL_BRI(2018)618988_EN.pdf> accessed 29 April 2025.

[37] Dedipyaman Shukla, 'Ownership In Non-Personal Data: A Perspective From Copyright Law' (*Indian Journal of Law and Technology*, 11 August 2022) <https://www.ijlt.in/post/ownership-in-non-personal-data-a-perspective-from-copyright-law> accessed 29 April 2025; Rekha Jain and Viswanath Pingali, 'India's Non-personal Data Framework: A Critique' (2021) 9 CSI Transactions on ICT 171.

[38] Nikhil Pahwa, 'India Must Avoid Nationalisation of Data' (*Medianama*, 15 July 2020) <https://www.medianama.com/2020/07/223-non-personal-data-nationalisation/> accessed 29 April 2025; See also Maximiliano Facundo Vila Seoane, 'Data Securitisation: The Challenges of Data Sovereignty in India' (2021) 42 Third World Quarterly 1733.

[39] See generally Wolfgang Kerber, 'A New (Intellectual) Property Right for Non-Personal Data?' (2016) MAGKS – Joint Discussion Paper Series in Economics (Band 37-2016); Srijoni Sen, Inder Gopal, and D Manjunath, 'Non Personal Data: Policy, Economics and Technology' (27 August 2021) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3911698> accessed 29 April 2025; Josef Drexl, 'Designing Competitive Markets for Industrial Data: Between Propertisation and Access' (2016) Max Planck Institute for Innovation & Competition Research Paper No. 16-13 <https://ssrn.com/abstract=2862975> accessed 29 April 2025.

[40] See Tommaso Fia, 'Resisting IP Overexpansion: The Case of Trade Secret Protection of Non-Personal Data' (2022) 53 IIC – International Review of Intellectual Property and Competition 917; Yu Xiaolan and Zhao Yun, 'Dualism in Data Protection: Balancing the Right to Personal Data and the Data Property Right' (2019) 35 CLSR 105318; Guobin Cui, 'The Eligibility Requirements for Legal Protection of Publicly Accessible Datasets' (2022) 14 LIT 157.

[41] See Kyle Wiggers, 'Reddit Says it's Made $203M so far Licensing its Data' (*Techcrunch*, 22 February 2022) <https://techcrunch.com/2024/02/22/reddit-says-its-made-203m-so-far-licensing-its-data/> accessed 29 April 2025.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

Some evidence suggests that adequate economic incentives can motivate companies to share non-personal data and the infrastructure to use it, as demonstrated by open-source large language models such as Meta's Llama-2.[42] For instance, the pre-trained Llama-2 AI model is available free as an open-source infrastructure, and the fees only apply to apps or licensors with at least 700 million monthly users. Thus, most small and medium-sized apps and businesses can use Llama-2 for free, thereby getting access to both insights from non-personal data and to the code and infrastructure that processes that data.[43]

A key consideration in adopting a property-oriented approach is determining the right market value of data.[44] Relying solely on regulators to calculate the market value can lead to errors and regulatory overdrive. For example, the EU's Data Act, aimed at enabling fair and equitable data sharing by manufacturers of Internet of Things (IoT) products, has been criticized for taking away IoT providers' discretion and contractual freedom to determine compensation for their non-personal data.[45] Eventually, this can disincentivise data sharing and create business uncertainty.[46] However, leaving data pricing entirely to market mechanisms also necessitates sufficient regulatory guidelines and checks to prevent market failures like information asymmetry, rampant in digital markets.

Non-personal data as commons

Finally, scholars have argued for a commons-based approach to regulating non-personal data.[47] Non-personal data are seen as a cooperative infrastructural resource to be liberated from private ownership to yield the maximum benefits for society. Simply put, this means that no single entity generates non-personal data, but that it results from cooperative interaction between many entities and individuals.[48] For example, Google Maps is able to generate non-personal data not only because of its technologies, but also because of the constant interaction of external players such as smartphone manufacturers, internet service providers, individuals who use Google Maps during travel, and the government's traffic system data. Furthermore, proponents of data commons highlight that almost all privately held data has its roots in publicly available data.[49] Thus, non-personal data arguably result from society-wide cooperation, warranting equal and common access to all.[50]

To avoid possible market failures due to disproportionate market power of a handful of companies, commons proponents argue that non-personal data regulation requires data

[42] Competition & Markets Authority, 'AI Foundation Models Review: Short Version' (2023) <https://assets.publishing.service.gov.uk/media/65045590dec5be000dc35f77/Short_Report_PDFA.pdf> accessed 29 April 2025.

[43] ibid.

[44] See generally Wolfgang Kerber, 'Governance of IoT Data: Why the EU Data Act Will Not Fulfill Its Objectives' (2023) 72 GRUR International 120.

[45] ibid; Foo Yun Chee, 'Businesses Criticise New EU Data Rules, Consumer Group Sees Missed Opportunity' (*Reuters*, 29 June 2023) <https://www.reuters.com/technology/businesses-criticise-new-eu-data-rules-consumer-group-sees-missed-opportunity-2023-06-28/> accessed 29 April 2025.

[46] Kerber (n 32).

[47] Parminder Jeet Singh, 'Data and Digital Intelligence Commons' (Making a Case for their Community Ownership) (1 November 2019) Data Governance Network Working Paper 02 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3873169> accessed 29 April 2025; Tomasso Fia, 'An Alternative to Data Ownership: Managing Access to Non-Personal Data Through the Commons' (2021) 21(1) Global Jurist 181.

[48] Fia, ibid; Patrick Hummel, Matthias Braun and Peter Dabrock, 'Own Data? Ethical Reflections on Data Ownership' (2021) 34 Philos Technol 545; Singh, ibid; Trishi Jindal and Aniruddh Nigam, 'Data Stewardship for Non-Personal Data in India: A Position Paper on Data Trusts' (2020) Vidhi Centre for Legal Policy Submissions & Working Paper <https://vidhilegalpolicy.in/wp-content/uploads/2020/11/201120_Data-Trusts_Position-Paper_Final-1.pdf> accessed 29 April 2025.

[49] Parminder Jeet Singh and Anita Gurumurthy, 'Data Sharing Requires a Data Commons Framework Law' (Data Governance Network Policy Brief 2.0, 2020) <https://www.datagovernance.org/files/research/Policy_Brief_02.pdf> accessed 29 April 2025; Jindal and Nigam, ibid.

[50] Jacqueline Hicks, 'The Future of Data Ownership: An Uncommon Research Agenda' (2023) 71 Sociol Rev 544.

Downloaded from https://academic.oup.com/jilit/article/doi/10.1093/jilit/eaaf025/8443255 by guest on 29 January 2026

management techniques, not data ownership/property rules.[51] These data management techniques are borne out of the theory of the commons, which often incorporates elements of stewardship and trusteeship.[52] In the context of data commons, several scholars have proposed that data management should be entrusted to neutral intermediaries such as data trusts and data cooperatives, who would ensure fair, ethical, and transparent data use.[53] In this framework, independent data trusts or data intermediaries are primarily responsible for managing non-personal data sharing, instead of a State or private sector mechanism. These neutral entities deploy data management mechanisms such as proposing non-intrusive mechanisms for collecting data, implementing data quality norms, regulating fair use and sharing of such data, ensuring that the data held by them is not put to exploitative use, and fostering public trust through accountability mechanisms.[54] The Indian policy framework on non-personal data partially endorses a commons approach (with imperfections and a strong State-centric bias, as discussed below).

A government-supported non-personal data commons is a well-intentioned policy in theory, but it can face public choice failures in practice. For instance, a non-personal data common can be misused, especially if government organizations or bodies closely linked to the government act as data intermediaries. Governments could also use non-personal data commons to control their citizens' behaviour.[55] Particularly, governments that have ready access to a large amount of personal data of their citizens can integrate the same with non-personal datasets for greater impact. For instance, China's social credit system, which uses both non-personal and personal data, demonstrates the potential for using data to shape citizen behaviour, such as rewarding compliant behaviour and penalizing non-compliance and dissent.[56] Another example is the Gangs Matrix used by the London Metropolitan Police to prevent gang violence; although public information remains deficient, experts believe it combines personal data with aggregated/group level indicators such as music tastes, social media activities, and presence in certain neighbourhoods, and primarily targets Black people.[57]

Further, a commons framework can lead to capture by special interest groups. Powerful industries or organizations may influence the design and implementation of non-personal data commons at the expense of the broader public interests.[58] In addition, when governmental agencies manage the non-personal data commons, it is harder to attribute responsibility for

[51] Parminder Jeet Singh and Anita Gurumurthy, 'Economic Governance of Data: Balancing Individualist-Property Approaches With Community Rights Framework' (2021) Data Governance Network Working Paper <https://itforchange.net/sites/default/files/1880/Economic-governance-of-data.pdf> accessed 29 April 2025; Jennifer Shkabatur, 'The Global Commons of Data' (2019) 22 Stan Tech L Rev 354.

[52] Peter G Brown, 'Toward an Economics of Stewardship: The Case of Climate' (1998) 26 Ecol Econ 11.

[53] Bart van der Sloot and Esther Keymolen, 'Can We Trust Trust-Based Data Governance Models?' (2022) 4 Data & Policy e45; Sylvie Delacroix and Neil D Lawrence, 'Bottom-up Data Trusts: Disturbing the "One Size Fits All" Approach to Data Governance' (2019) 9 IDPL 236.

[54] Francesca Perucci and Eric Swanson, 'Building Trust and Facilitating Use of Data' (2024) 40 Statistical Journal of the IAOS 71; Puneeth Nagaraj, Varsha Rao and Dedipyaman Shukla, 'Community Rights over Non-Personal Data: Perspectives from Jurisprudence on Natural Resources' (Data Governance Network, 2020) <https://www.datagovernance.org/files/research/1611826214.pdf> accessed 29 April 2025.

[55] See Lina Al-Hathloul, 'Dictators in Egypt and Saudi Arabia love smart cities projects — here's why' (*Access Now*, 1 March 2022) <https://www.accessnow.org/smart-cities-projects/> accessed 29 April 2025; Evie Lucas and Seamus Simpson, 'Perspectives on Citizen Data Privacy in a Smart City – An Empirical Case Study' (2024) Convergence 13548565241247413; Zhijian Wang and Yoshiko Naiki, 'Smart Cities and Privacy: Comparative Analysis of Japan and China on the Use of Facial Recognition Technology in Public Spaces' (2025) IJLIT (forthcoming).

[56] Xu Xu, Genia Kostka and Xun Cao, 'Information Control and Public Support for Social Credit Systems in China' (2022) 84 J Politics 2230.

[57] Amnesty International, 'Trapped in the Matrix: Secrecy, Stigma, and Bias in the Metropolitan Police Gangs Database' (Amnesty International UK, May 2018) <https://www.amnesty.org.uk/london-trident-gangs-matrix-metropolitan-police> accessed 29 April 2025.

[58] Aria Thaker, 'The New Oil: Aadhar's mixing of public risk and private profit' (The Caravan, 1 May 2018) <https://caravanmagazine.in/reportage/aadhaar-mixing-public-risk-private-profit> accessed 29 April 2025.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

privacy breaches and other failures, as governmental bodies may enjoy immunity under domestic laws.[59]

Finally, a commons-based framework does not always ensure equitable outcomes. For instance, Big Tech companies with access to sophisticated digital systems and technologies are likely to be able to use non-personal data more meaningfully to generate digital intelligence as compared to smaller companies (assuming all non-personal data is equally accessible to all).[60] In adopting a mantra of freeing up as much non-personal data as possible, governments lose sight of other critical factors, such as accuracy, completeness, and quality of data.[61] Another possible pitfall is that the excessive focus on data sharing may take away attention from individual privacy rights.[62]

Conclusively, the choice of the appropriate regulatory approach on non-personal data entails difficult policy choices. The generation and use of non-personal data requires a deeper understanding of how market players generate and use non-personal data. Further, the basis of non-personal data regulation is challengeable as personal and non-personal data are not neatly segregable. Privacy risks underlie the sharing of both government and private sector non-personal data. As we discuss in the context of India below, a myopic perspective of non-personal data can result in skewed policy and legal outcomes, ultimately harming rather than benefiting the public.

EXPLORING INDIA'S REGULATORY FRAMEWORK FOR NON-PERSONAL DATA

This section evaluates the non-personal data framework in India, including its institutional mechanisms for data sharing. In a somewhat ingenious but potentially risky move, the Indian government has chosen a policy-led approach to governing non-personal data, stepping away from traditional lawmaking. Instead of drafting a comprehensive legislation that would require legislative debates and approval,[63] the Indian government has approached non-personal data regulation primarily through executive actions and by putting in place infrastructures for data sharing, as discussed in detail below.

We see this strategy as a 'techno-political' governance framework that circumvents the traditional route of developing a dedicated legislation on non-personal data.[64] While this approach can be viewed as pragmatic and flexible, especially given India's exploding digital economy and delayed implementation of its nascent data protection law, it carries significant risks. In particular, the establishment of complex infrastructure for data sharing without appropriate legal

[59] Digital Personal Data Protection Act, 2023 <https://www.meity.gov.in/static/uploads/2024/06/2bf1f0e9f04e6fb4f8fef35e82c42aa5.pdf> accessed 29 April 2025.

[60] See Katerina Yiannibas, 'Inequality of Arms, Power & Remedy: The Case for Diverse and Input-Based Dispute Management System' (*Cambridge Core Blog*, 30 March 2020) <https://www.cambridge.org/core/blog/2020/03/30/inequality-of-arms-power-remedy-the-case-for-diverse-and-input-based-dispute-management-systems/> accessed 29 April 2025.

[61] Eszter Hargittai, 'Is Bigger Always Better? Potential Biases of Big Data Derived from Social Network Sites' (2015) 659 Ann Am Acad Political Soc Sci 63; Daron Acemoglu and others, 'Too Much Data: Prices and Inefficiencies in Data Markets' (2022) 14 AEJ: Microeconomics 218.

[62] Nadya Purtova and Gijs van Maanen, 'Data as an Economic Good, Data as a Commons, and Data Governance' (2024) 16 LIT 1.

[63] Previously, a Bill had been proposed in India that would provide a common legislative framework for non-personal and personal data, but it was later withdrawn. See Indranath Gupta and Paarth Naithini, 'Separating Personal Data Protection from Non-personal Data Governance' (2023) 58 Economic and Political Weekly <https://www.epw.in/journal/2023/36/commentary/separating-personal-data-protection-non-personal.html> accessed 29 April 2025.

[64] Sunil Abraham, 'Systematic Government Access to Private-Sector Data in India' in Fred H. Cate and James X. Dempsey (eds), *Bulk Collection: Systematic Government Access to Private-Sector Data* (Oxford University Press 2017) 267.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

safeguards can have significant pitfalls for the protection of individual rights and ensuring governmental accountability.

Emerging policy framework for non-personal data

The draft framework for non-personal data in India

In 2019, the Indian government set up a committee of experts for non-personal data regulation (Kris Gopalakrishnan committee) that was tasked with recommending a regulatory framework for non-personal data.[65] The expert committee had recommendatory capacity but lacked any direct legal authority. The Kris Gopalakrishnan committee published two recommendatory reports: the first report came out in July 2020,[66] and a revised report in December 2020 ('2020 Report').[67] Both reports endorse an umbrella national legislation for enabling access to and sharing of non-personal data of the private sector to boost opportunities for Indian businesses, thereby creating economic and social benefits for the Indian community. The Committee justified the legal basis of non-personal data regulation by referring to one of the Directive Principles of State Policy in the Indian Constitution, which requires the government to ensure that 'ownership and control of the material resources of the community are so distributed as best to subserve the common good'.[68]

The 2020 Report lays out the core guiding principles for non-personal data protection in India.[69] Although it is a draft report that does not by itself represent the government's official position and has not been incorporated into a specific law, the high-level principles and the general approach recommended by the Report is reflected in various governmental actions such as the implementation of the Data Empowerment and Protection Architecture (DEPA) and the formulation of the Draft Electronic Commerce Policy,[70] as discussed below. The core guiding principles of the 2020 Report are:[71]

(i) sovereignty: India has rights over the data of India, its people and organizations;
(ii) benefit India: benefits of data must accrue to India and its people;
(iii) benefits the world: innovation, new models and algorithms for the world;
(iv) privacy: Misuse, reidentification and harms must be prevented;
(v) simplicity: the regulations should be simple, digital, and unambiguous;
(vi) innovation and entrepreneurship: the data should be freely available for innovation and entrepreneurship in India.

Two aspects of these principles are particularly intriguing. First, although they seem to steer towards data sovereignty (which we explain further below) and fostering domestic innovation, the report also states that digital innovations must be of global benefit. This could be a possible indicator of India's ambition to be a global player in data-driven technologies. Second, although there are strong notions of sovereign control over data, the report appears to pay at

[65] Incidentally, Gopalakrishnan is the co-founder of a highly successful Indian IT company, Infosys.
[66] MEITY, *Report by the Committee of Experts on Non-Personal Data Governance Framework* (July 2020) <https://ourgovdotin.wordpress.com/wp-content/uploads/2020/07/kris-gopalakrishnan-committee-report-on-non-personal-data-governance-framework.pdf> accessed 29 April 2025.
[67] 2020 Report (n 16).
[68] The Constitution of India 1950, art 39(b).
[69] For a detailed discussion, see Rishab Bailey and Renuka Sane, 'A Missed Opportunity' (*The Hindu*, 3 September 2020) <https://www.thehindu.com/opinion/op-ed/a-missed-opportunity/article32507522.ece> accessed 29 April 2025.
[70] Draft National e-Commerce Policy: India's Data for India's Development, sec I <https://static.investindia.gov.in/s3fs-public/2019-03/DraftNational_e-commerce_Policy_23February2019.pdf> accessed 29 April 2025.
[71] 2020 Report (n 16) para 3.4.

Downloaded from https://academic.oup.com/jilit/article/doi/10.1093/jilit/eaaf025/8443255 by guest on 29 January 2026

least some attention to privacy and related harms and occasionally refers to the 'rights of India and its communities over non-personal data'.[72]

The articulation of data sovereignty in the 2020 Report merits further consideration. Although the concept of data sovereignty remains contested,[73] the 2020 Report frames the principle of data sovereignty in the context of non-personal data as India's right to control the non-personal data generated within its borders, including the geographical flow of the data, who accesses it and how, and how this data is processed and leveraged to primarily benefit India economically. The Report's approach is closely aligned with the idea that a government must be able to apply its laws and regulations and control the data that is generated within its borders, thus retaining self-determination over its data.[74] Data sovereignty is also used to imply that the State, and not individuals or organizations, have the ultimate oversight over the data being generated within its jurisdiction.[75]

In the context of the 2020 Report, the expression of data sovereignty for non-personal data regulation is predominantly viewed from an economic and development perspective, such as India's sovereignty in determining how to unlock the value of data, generate new digital innovations and start-ups, and promote Indian businesses.[76] While there are occasional references to addressing privacy concerns, including risks of re-identification,[77] this aspect is underaddressed compared to the predominant economic objective. As an example, the report mandates companies to provide complete open access to their metadata to ensure more economic innovation,[78] without taking into account that such metadata can harm their competitive/commercial interests and pose privacy risks. Another example is the Indian AI mission, wherein the government aims to set up a platform for non-personal data sharing to create 'sovereign AI models', [79] without accounting for potential privacy risks.[80]

In fact, the two sets of rights over non-personal data are also curiously framed:[81] (i) right to derive economic and other value and maximizing data's benefits for the community and (ii) right to eliminating or minimizing harms from the data to the community.

The above framework, recognizing the rights of a vaguely defined 'community' (as discussed further below), includes both the right of the community to derive economic value from non-personal data and to be protected against data-related harms, without explaining how the balance between the two will be achieved in practice. There is no clear indication that the objective of preventing digital harms would prevail over the economic rights of the community.

The report is further vague regarding who actually can exercise such rights. While it suggests that individuals (or data principals, as termed in India's personal data protection law)

[72] ibid para 2.3.

[73] Jens Meijen, 'Future-Proofing the People? A Comparative Analysis of Data Sovereignty as a Discursive Practice in Western European Right-Wing Populism's Digital Policies' (2024) 29 Information Polity 73–91.

[74] Melissa Lukings and Aarash Habibi Lashkari, *Understanding Cybersecurity Law in Data Sovereignty and Digital Governance* (Springer 2022) 117; Luca Belli, Walter B Gaspar and Shilpa Singh Jaswant, 'Data Sovereignty and Data Transfers as Fundamental Elements of Digital Transformation: Lessons from the BRICS Countries' (2024) 54 CLSR 106017.

[75] Mark Ryan, Paula Gürtler and Artur Bogucki, 'Will the Real Data Sovereign Please Stand Up? An EU Policy Response to Sovereignty in Data Spaces' (2024) 32 IJLIT 1.

[76] 2020 Report (n 16) para 3.6.

[77] ibid para 3.55.

[78] ibid para 6.1.

[79] The Economics Times, 'Will be Developing Our Own Indian Models: Rajeev Chandrasekhar Unveils AI Mission Roadmap' (11 March 2024) <https://www.youtube.com/watch?v=HOAUM8wlO08> accessed 29 April 2025.

[80] For example, even personal data collected through scraping is non-personal data in Indian law. See Digital Personal Data Protection Act 2023, s. 3(c) (Illustration).

[81] 2020 Report (n 16) para 7.1.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

continue to exercise their rights over personal data, non-personal data rights belong to the community.[82] The community is defined as:[83]

> …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.

The community can exercise the above rights through a 'data trustee', defined as a government body or a non-profit organization.[84] First, since the regulation relates to non-personal data, the role of the individual privacy is mostly irrelevant,[85] despite the 2020 Report stating, 'data collectors at the time of collecting personal data should provide a notice and offer the data principal the option to opt out of data anonymization'.[86] Secondly, most Indian users are unlikely to understand the implications of not opting out of anonymization due to low levels of digital literacy.[87]

The definition of the community is also vague; for instance, individuals could belong to multiple communities based on the context in which the use of non-personal data affects them, which would make it harder for them to identify the relevant data trustee and exercise their community rights.[88] Further, where a digital harm relates to non-personal data shared by governmental bodies, it is unclear whether community rights would be effectively exercisable if the trustee were a government body. In certain scenarios, these two bodies could also be the same. For instance, the Health and Welfare Department could share the non-personal data while also being the trustee for the community. The regulator may also find it difficult to detect malpractices of non-profit data trustees set up by powerful companies.

The 2020 Report introduces several new concepts to refer to stakeholders in the digital economy, adding further confusion. For instance, it defines a data custodian as any entity (public or private) that engages in 'collection, storage, processing, use, etc. of data' from consumers.[89] It is unclear why the term 'data fiduciary' was not used instead, similar to the Indian personal data protection law.[90] It imposes various obligations on the data custodian to share non-personal data in response to data requests from the government or other prescribed third parties.[91] The category of 'data custodian' appears to overlap with 'data business', which is defined as any governmental or private body that 'collects, processes, stores or otherwise manages data'.[92] In fact, the 2020 report states that 'a data business can be a data custodian'.

The 2020 Report further states that data custodians must engage in 'responsible data stewardship' and exercise a 'duty of care' to the community in relation to handling their non-personal data.[93] It provides no further explanation as to what these terms entail in practice for data custodians; for example, the kind of technical, organizational or legal measures

[82] ibid para 7.1.
[83] ibid para 7.2.
[84] ibid para 7.2.
[85] However, 2020 Report (n 16) para 8.6 states that non-personal data sharing must not violate privacy of individuals, groups, or communities. This is however not supplemented with guidelines or other legal mechanisms.
[86] ibid para 5.4.
[87] Aafreen Mitchelle Collaco, 'Contours of Data Protection in India: The Consent Dilemma' (2024) IRLCT 1.
[88] Rishab Bailey and Renuka Sane, 'Regulating 'Non-Personal Data': Developments in India' in Bart van der Sloot and Sascha vann Schendel (eds), *The Boundaries of Data* (Amsterdam University Press 2024) 241.
[89] 2020 Report (n 16) para 7.4.
[90] Digital Personal Data Protection Act 2023, s 2(i).
[91] 2020 Report (n 16) para 7.4.
[92] ibid para 6.1.
[93] ibid para 7.4.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

that data custodians must put in place to ensure secure anonymization and sharing of non-personal data, and how they can be held accountable by the community.

The 2020 Report also introduces the concept of a high-value dataset, vaguely defined as 'a dataset that is beneficial to the community at large and shared as a public good'.[94] A wide range of use cases of high-value datasets are identified in the 2020 Report, including providing public services and citizen engagement; financial inclusion; poverty alleviation; healthcare; urban planning; energy use; creating new business opportunities, innovations and jobs; and research. Given that the report identifies such a broad range of use cases, a large number of non-personal datasets could qualify as high-value datasets.

Data trustees are intermediaries responsible for the creation, maintenance, and sharing of these high-value datasets (eg through APIs). Several scholars have argued that data trustees must act as neutral stewards of data and operate with the ultimate objective of ensuring that non-personal data is used purposefully.[95] In practice, they operate as an intermediary entity responsible for collecting data from data custodians (private companies collecting non-personal data) and enabling access to that data to other potential users. To enable this intermediary framework under the 2020 Report, data trustees are empowered to request any non-personal data from data custodians on the grounds of 'public interest',[96] which can be requested from them by third parties termed as 'data requesters'.[97] In providing such services to data requesters, data trustees can charge a nominal fee.[98] Data custodians may be paid a small fee to facilitate the sharing of high-value datasets but cannot refuse to share such data (eg on the grounds of losing competitiveness) or demand a fair market value of their data.[99] Any such refusals can be subject to adjudication, as discussed below. Notably, the 2020 Report makes only symbolic references[100] to relevant property rights of businesses such as copyright and trade secrets (incidentally, India also does not have a law on trade secrets but remains bound by the TRIPS agreement of the WTO, which contains an explicit provision on trade secrets).

As previously discussed, mandatory data sharing does not account for important economic incentives of companies, and can lead to inefficient outcomes, including deliberate measures to avoid sharing. Further, data trustees hold significant power to mandate non-personal data sharing for high-value datasets, as 'public interest' can be interpreted very broadly, but the 2020 Report does not provide a clear accountability framework for these bodies. While it mentions 'responsible data stewardship' and 'duty of care' applying to data trustees,[101] no clear mechanism is set out to action these duties, besides a possible grievance redressal mechanism set up by the data trustee.[102] In any case, the framework proposed in the 2020 Report does not provide any clear choice to the end users/data subjects to choose a particular data trustee to manage their data.[103]

The non-personal data regulator (details of which are outlined further below) can adjudicate in scenarios where the data custodian refuses to share certain non-personal data with the

[94] ibid para 7.6.
[95] Iris Schneider, 'Data Stewardship by Data Trusts: A Promising Model for the Governance of the Data Economy?' in Claudia Padovani, Véronique Wavre, Arne Hintz, Gerard Goggin and Petros Iosifidis (eds), *Global Communication Governance at the Crossroads* (Palgrave Macmillan 2024).
[96] 2020 Report (n 16) para 7.8.
[97] ibid para 7.7.
[98] ibid para 7.7.
[99] ibid para 8.5.
[100] See ibid para 8.6.
[101] ibid para 7.7.
[102] ibid para 7.7.
[103] See the proposal by Delacroix and Laurence (n 57) 241.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

data trustee. If the adjudication lies against a government body acting as a data trustee, it is doubtful how effective the remedies would be to protect any concerns of the data custodian. These examples are not hypothetical; for instance, a press report in 2019 reported that the Ministry of Road Transport and Highways sold two databases to private third parties for profit without getting appropriate consent from the users.[104] The 2020 Report could have taken an alternative route of setting up robust requirements for any entity to qualify as a data trustee, including assessing their technical and organizational capabilities and unequivocally proving that there was no conflict of interest possible in their functioning as a data trustee.[105]

The 2020 Report recommends the institution of a Non-Personal Data Protection Authority (NDPA) to supervise the implementation of non-personal data regulation.[106] The NDPA is designated an enabling role (by providing frameworks and formats for data sharing) and an enforcing function (to establish and protect vaguely defined rights arising in relation to non-personal data).[107]Although the 2020 Report recommends alignment between the NDPA and other bodies such as the Competition Commission of India and the Data Protection Board of India, there are no further guidelines on the same, including managing jurisdictional conflicts between regulators. It is also unclear whether the NDPA will function independently.[108]

The 2020 Report recommends a broad exception for governments requesting data for 'sovereign' purposes, which is also undefined. Non-personal data can be a double-edged sword. While governments can use this exception to request data to improve several public services, they can also use it to control citizens or monitor their behaviour. Since the 2020 Report proposes that NDPA cannot adjudicate any non-personal data requests relating to sovereign powers, this also provides unbridled powers to the government.[109] The 2020 Report has not yet been implemented, although data-sharing mechanisms are already in practice through mechanisms such as the DEPA and Open Network Digital Commerce ('ONDC'), discussed in more detail below.

Draft national governance framework policy

The Government of India also published the Draft National Data Governance Framework Policy, 2022,[110] to establish a policy framework for streamlining and augmenting the availability of government-collected non-personal data for broader research and economic use, and creating an India Datasets programme to bring together non-personal data from both government and private sources.[111] The underlying rationale is the same as the 2020 Report, ie non-personal data are a valuable economic/national resource to drive drastic digital innovation in emerging areas such as AI. This report was revised and published in May 2022, replacing the Draft India Data Accessibility and Use Policy.[112]

[104] Shivam Srivastav, 'Indian Govt Is Selling Vehicle Owner Data To Companies And Citizens Don't Have A Clue' (*Inc42*, 13 July 2019) <https://inc42.com/buzz/indian-govt-is-selling-vehicle-owner-data-to-companies-and-citizens-dont-have-a-clue/> accessed 29 April 2025.

[105] Commission, 'Staff Working Document - Impact Assessment Report Accompanying Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act)' EC(2020) 405 final, 26.

[106] 2020 Report (n 16) para 7.10.

[107] ibid para 7.10.

[108] The Data Protection Board of India (enforcing India's personal data protection law) can be a parallel example; its constitution, budget and functioning is subject to a high degree of control by the Central government.

[109] 2020 Report (n 16) para 8.1.

[110] MEITY, *National Data Governance Framework Policy* (2022) <https://www.thehinducentre.com/resources/67557000-National-Data-Governance-Framework-Policy_compressed.pdf> accessed 29 April 2025.

[111] Private entities can also share information in the dataset. See ibid para 6.3.

[112] MEITY, *Draft Data Accessibility and Use Policy* (2022) <https://www.thehinducentre.com/resources/article38451598.ece/binary/Draft%20India%20Data%20Accessibility%20and%20Use%20Policy.pdf> accessed 29 April 2025; See Rudra Chaudhuri and Arjun Kang Joseph, 'Living in a Fragmented World: India's Data Way' (2024) 23(2) India Review 154, 164.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

The policy seeks to ensure that government data collection is modernized and systematized, enabling better sharing of non-personal data countrywide through rapid anonymization[113] and fostering AI research, in particular. This policy also envisages an Indian Data Management Office ('IDMO'), which would be responsible for setting the standards for the storage and collection of public data and enabling the creation of an Indian datasets programme.[114] The previous draft of the policy faced severe public criticism due to privacy concerns pertaining to the monetization of public data, particularly in light of the extensive amount of data collected by the Indian government.[115] Although this policy sets out a broad mandate for governance of government data sharing, including privacy, security and trust concerns, the mechanisms to achieve this in practice are unclear. Unless future guidelines provide more detailed mechanisms for protecting data security and privacy, this non-personal data-sharing initiative would also likely raise privacy concerns.

Infrastructure and institutional framework for data sharing

The Indian government has established India Stack, a data governance framework to facilitate access and sharing of data between various entities.[116] It consists of the digital ID/Aadhar, designated to all Indian citizens, coupled with other mechanisms/layers such as the government-administered Digilocker to share certificates and documents, a payments interface called the Unified Payments Interface ('UPI') and the DEPA to enable widespread sharing of data.[117]

The India Stack relies on a layered approach to data governance, with the Aadhaar as its foundation. Aadhar is not only a unique ID for every Indian citizen but also a near-universal requirement for day-to-day transactions, ranging from buying basic services, including banking, to registering for government schemes. Consequently, and unsurprisingly, Aadhaar has become an ever-expanding hotspot of personal data. In fact, by some estimates, the Indian government is the biggest collector of personal data in India.[118] As per reports, the Aadhaar system has been hacked multiple times.[119]

The most important and widespread services currently linked to the Aadhaar database are digital payments and finance. Popularly known as UPI, this layer of the India Stack relies on interlinkages between an individual's mobile number, their bank account, and their Aadhaar. The UPI is managed by the National Payments Corporation of India ('NPCI'), a non-profit organization set up by the Reserve Bank of India (India's central bank—'RBI') in collaboration with an association of 30 major Indian banks. The NPCI has created an interoperable digital ecosystem enabling instantaneous transfer between banks using a QR code or mobile number, with the back-end infrastructure ensuring that the mobile numbers are connected with Aadhaar and a bank account. While it has

[113] No detailed standards or protocols for anonymisation and data sharing are set out in this policy.
[114] MEITY (n 112) para 5.
[115] Chaudhuri and Joseph (n 121) 164; Soujanya Sridharan and Shefali Girish, '(Em)Powering India through data: Thoughts on the National Data Governance Policy Framework' (*Medianama*, 9 August 2022) <https://www.medianama.com/2022/08/223-india-national-data-governance-policy-framework-idmo/> accessed 12 December 2024.
[116] Smriti Parsheera, 'Stack is the New Black?: Evolution and Outcomes of the 'India-Stackification' Process' (2024) 52 CLSR 105947.
[117] ibid.
[118] Vindu Goel, 'On Data Privacy, India Charts Its Own Path' (*The New York Times*, 10 December 2019) <https://www.nytimes.com/2019/12/10/technology/on-data-privacy-india-charts-its-own-path.html> accessed 29 April 2025.
[119] Anushka Sengupta, 'How Did Massive Aadhaar Data Leak Happen, its Impact' (*The Economic Times*, 2 November 2023) <https://bfsi.economictimes.indiatimes.com/news/industry/how-did-massive-aadhaar-data-leak-happen-its-impact/104901967> accessed 29 April 2025.

Downloaded from https://academic.oup.com/jilit/article/doi/10.1093/jilit/eaaf025/8443255 by guest on 29 January 2026

brought down transaction costs and increased efficiency,[120] it has also increased risks of online fraud and data leaks.[121]

The DEPA, which is the electronic consent management system for a data-sharing layer of the India Stack, is particularly important in the context of this paper. It enables real-time movement of non-personal data (also, personal data) between various platforms, entities, businesses, and institutions. These transfers are implemented through a consent management scheme overseen by entities called 'Consent Managers' (or Account Aggregators in the fintech sector).[122] These entities, licensed by India's central bank, the Reserve Bank of India, act as intermediaries between the users, the entity required to share data, and the entity attempting to receive the data.[123] These intermediaries enable the sharing of data by obtaining, managing, and revoking consent from data subjects.[124] In practice, Account Aggregators collect and assimilate information from different financial services used by an individual and provide a systematic and easy overview for users who are registered on their system.

Consent managers have been ascribed specific obligations under India's personal data protection law, such as registering with the Data Protection Board of India, and adopting prescribed requirements for the adoption of technical, operational, and financial standards.[125] They are also ultimately accountable to end users for the data that they share regarding them (but such complaints can specifically only relate to personal data).[126] The DEPA highlights the importance placed on the exchange and interoperability of data as an economic resource in the Indian economy. Further, as data can easily change hands under the DEPA framework through a highly simplified yes/no query, the consent mechanism under DEPA is diluted. In fact, most apps do not even provide an explicit yes/no query since the data protection law is not yet enforced. For instance, most applications in the financial sector, instead of asking whether a user would like to share or give access to their prior financial data, proceed to secure consent by default through pre-selected checkboxes when a user signs up with the application, with vague statements such as 'I, (name of individual), give my consent to download my KYC Records from the Central KYC Registry (CKYCR), only for the purposes of (stated purpose)'. The simplified illustration below shows how these transactions are structured in practice (Fig. 1).

The DEPA is critical because it provides a basis for the exchange of *inter alia* non-personal data through various open networks, such as the ONDC, which forms the fourth layer of India Stack. These open networks act as a 'platform of platforms' to pool data and insights from countless member platforms.[127] For example, the ONDC will act as an open network with

[120] Shailesh Rastogi and others, 'Unified Payment Interface (UPI): A digital innovation and its Impact on Financial Inclusion and Economic Development' (2021) 9(3) UJAF 518.

[121] Deepika Chelani, 'Rising UPI Scam Trend: Follow These Tips to Keep Yourself Safe from Scammers' (*Mint*, 28 November 2023) <https://www.livemint.com/money/personal-finance/upi-scams-on-the-rise-follow-these-tips-to-keep-yourself-safe-from-scammers-11701168127203.html> accessed 29 April 2025; 'UPI and Managing the Surge in Digital Fraud' (*The Indian Express*, 29 November 2023) <https://indianexpress.com/article/opinion/editorials/upi-and-managing-the-surge-in-digital-fraud-9046374/> accessed 29 April 2025.

[122] 'What are Account Aggregators' (*Sahamati*) <https://sahamati.org.in/what-is-account-aggregator/> accessed 29 April 2025.

[123] Vikas Kathuria, 'Data Empowerment and Protection Architecture: Concept and Assessment' (2021) Observer Research Foundation Issue Brief <https://www.orfonline.org/research/data-empowerment-and-protection-architecture-concept-and-assessment> accessed 29 April 2025.

[124] Tanay Mahindru and Anushka Mittal, 'Data Empowerment and Protection Architecture: Side-Stepping Empowerment for Convenience?' (December, 2021) IT for Change Research Paper <https://itforchange.net/sites/default/files/2015/Data-Empowerment-and-Protection-Architecture-Policy-Brief-Dec-2021.pdf> accessed 29 April 2025.

[125] Digital Personal Data Protection Act 2023, s 6.

[126] ibid.

[127] Tejasi Panjiar and Prateek Waghre, 'Open Network for Digital Commerce (ONDC): An Explainer' (*Internet Freedom Foundation,* 2023) <https://internetfreedom.in/ondc-an-explainer/> accessed 29 April 2025.

Downloaded from https://academic.oup.com/jiit/article/doi/10.1093/jiit/eaaf025/8443255 by guest on 29 January 2026

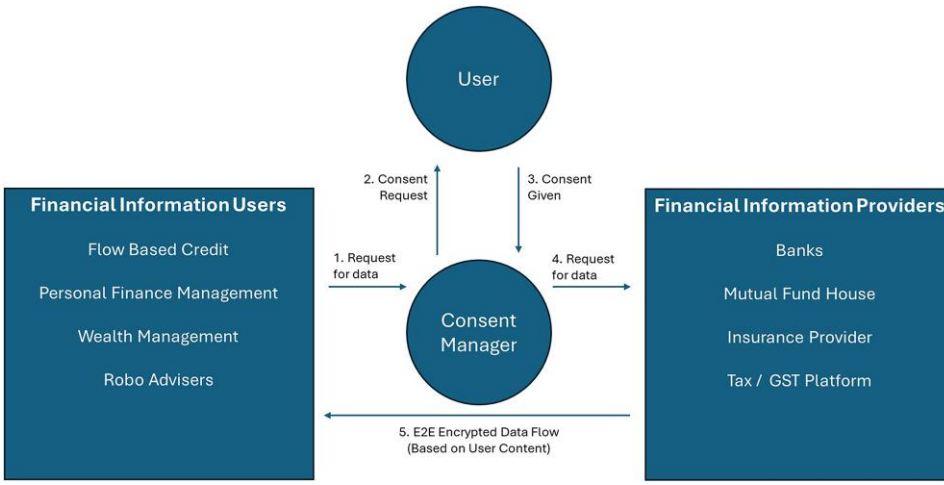

Figure 1 Developed by Mishra and Agrawal using visualization software.

e-commerce platforms, end-customers, sellers and delivery service providers.[128] Under this framework, a merchant is not required to register with a specific food delivery platform or an e-commerce company. Rather, once they join the ONDC platform, their data are automatically interoperable with all platforms and service providers that are part of ONDC.[129] Further, all the transactions taking place on ONDC would leave behind insights that would be made publicly available.[130] The overall aim of the open networks policy is to replace the prevalent walled-garden model of digital platforms with open platforms.

An example of an existing open platform is the Namma Yatri ride-hailing app, supported by the ONDC. All participants on this platform make their insights and metrics available publicly on their website and the ONDC network.[131] Aggregated and anonymized data containing information regarding live traffic data, total user statistics, net rider density and availability, cumulative earnings and total spend, number of ongoing trips, and number of requested bookings are made publicly available by Namma Yatri on its open data web portal. While the government does provide some subsidies/incentives to industry players to join these platforms, there is a perceptible element of coercion and restriction on contractual freedom.[132] This has also been occasionally reflected in comments of office bearers; for instance, the Minister of Commerce was quoted as saying, that India will use the 'full force of the government' to promote these open networks.[133] Another press report highlights how platforms are facing pressure to join the ONDC and provide their customer data or to repent later.[134]

[128] Open Network for Digital Commerce ('ONDC'), 'Democratizing Electronic Commerce in India' (2022) <https://www.medianama.com/wp-content/uploads/2022/03/ONDCStrategyPaper.pdf?ref=static.internetfreedom.in> accessed 29 April 2025.

[129] ibid 8-9.

[130] ibid 12.

[131] See 'Namma Yatri Open' <https://nammayatri.in/open/> accessed 29 April 2025.

[132] Deepsekhar Choudhury, 'Govt Wants Main Platforms to Come to ONDC, Not New E-comm Apps for Joining' (*Moneycontrol,* 26 April 2023) <https://www.moneycontrol.com/news/business/announcements/we-want-main-platforms-to-come-to-ondc-dont-come-with-new-e-commerce-app-piyush-goyal-10472551.html> accessed 29 April 2025.

[133] Andy Mukherjee, 'An Indian Biryani Won't Democratize E-Commerce' (*Bloomberg Opinion,* 18 May 2023) <https://www.deccanherald.com/opinion/an-indian-biryani-won-t-democratise-ecommerce-1219705.html> accessed 29 April 2025.

[134] Manish Singh, 'India Warns E-Commerce Giants, Urges Adoption of Open Network' (*TechCrunch*, 25 April 2023) <https://techcrunch.com/2023/04/25/india-warns-e-commerce-giants-urges-adoption-of-open-network/> accessed 29 April 2025.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

The focus of such initiatives seems to be increasing market opportunities and creating a level playing field, [135] without proper consideration of how non-personal data shared on these networks may prejudice individual rights or commercial interests of market players.

The manner in which platforms such as DEPA and ONDC envisage the role of data indicates a supposedly commons-based approach (albeit predominantly State-centric with little autonomy to citizens or market players). While open networks can be revolutionary and drive economic efficiency, the element of State coercion and disregard for both individual and commercial rights are counterproductive. In fact, given the huge amount of State support behind such mechanisms, it is hard to assess whether these networks are economically viable or more efficient than other alternative mechanisms.

Further, no clear guidelines exist regarding how bodies such as the ONDC will manage and secure non-personal data. The ONDC Strategy Paper states that:

> ONDC will foster and promote established and viable principles for platforms to emit anonymized performance metrics enabling informed policymaking and network robustness, empowering network participants fairly and equitably.[136]

Thus, open networks like ONDC will be critical in promoting the sharing of non-personal data within their ecosystem. However, without any relevant legislation, organizations like ONDC and IDMO will decide the protocols and standards for the sharing of non-personal data.[137] There are no substantive guidelines on data sharing protocols, data formats, and security/technical standards in any policy or legal instrument for these organizations.[138] Experts have raised concerns relating to re-identification due to the extensive sharing of anonymized data through such open networks.[139] Finally, since organizations like ONDC are set up as public-private partnership companies, they are outside the purview of key accountability mechanisms such as the Right to Information Act, 2005 (a law allowing citizens to access information held by public authorities), and thus not subject to public scrutiny and accountability.[140]

Interplay between non-personal data regulation and domestic laws

In addition to the above policy and institutional mechanisms, some domestic laws in India enable the government to access (or enable access to) non-personal data from private companies for law enforcement and regulatory supervision.[141] The RBI has also proposed to

[135] Sharad Sharma and others, 'Digital Public Infrastructures: Lessons from India for a Thriving Digital Economy' (*IE Center for the Governance of Change*, 2022) <https://static.ie.edu/CGC/12.%20ISPIRT%20-%20Digital%20Public%20Infrastructures%20(November%202023).pdf> accessed 29 April 2025.

[136] ONDC (n 130).

[137] Parsheera (n 118).

[138] The IT Act is currently being revamped for a new Digital India Act, but this has not yet been implemented. In fact, it is still at a legislative proposal stage.

[139] Panjiar and Waghre (n 129).

[140] For example, the Government refused public disclosures in the case of Digi Yatri (a not for profit dealing with airline passenger data). See Nisha Anand, 'Does Govt-Backed DigiYatra Store Passengers' Personal Data? Centre Explains' (*Business Standard*, 6 February 2024) <https://www.business-standard.com/india-news/does-govt-backed-digiyatra-store-passengers-personal-data-centre-explains-124020600320_1.html> accessed 29 April 2025.

[141] Information Technology Act 2000, s 29; Information Technology (Procedures and Safeguards for Interception, Monitoring or Decryption of Information) Rules, 2009; Reserve Bank of India Act 1934, ss 45K and 45L; Banking Regulations Act 1949, s 27. See also 'FAQs: Storage of Payment System Data' (Reserve Bank of India, 26 June 2019) <https://www.rbi.org.in/commonperson/English/Scripts/FAQs.aspx?Id=2995> accessed 29 April 2025.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

localize data in India by creating its own cloud infrastructure in partnership with 'incorporated Indian companies with prior cloud solutions experience'.[142] It is unclear if access to non-personal data enabled by the above domestic laws/initiatives will also be used for data sharing and reuse for economic reasons. However, several of the above authorities (such as the MEITY and the RBI) also oversee the implementation of economic policies concerning non-personal data, leading to overlapping functions.

The non-personal data protection framework emphasizes that boosting domestic digital innovation requires providing better access to smaller players to non-personal data held by bigger technology companies, including Western Big Tech companies.[143] For example, the India AI Mission aims to create a platform for non-personal data, available only for Indian companies,[144] although the operational details remain unclear. The Digital Competition Bill of 2024 also highlights the same approach by fostering more data accessibility, including data portability mechanisms to create a level playing field for smaller digital companies.[145] The competition regulator in India has also held that anti-competition harms can arise from closed data ecosystems of non-personal data.[146] However, currently, the non-personal data policy framework does not address the interface with competition law, despite overlapping objectives.

Non-personal data regulation in India: a weak work-in-progress?

The Indian government has emphasized the importance of data in building a competitive, robust digital economy and fostering development.[147] However, a robust regulatory framework that protects the digital rights of citizens and provides an ecosystem that facilitates secure and trustworthy data sharing is equally important.[148] As Streinz argues, frameworks that increase access to data are ripe with regulatory conflicts.[149] The lack of a clear rights-based approach, particularly the negligible focus on individual privacy rights and the absence of clear accountability mechanisms in the use and management of non-personal data by data intermediaries, is a glaring deficiency of the Indian framework. The proposal for the Digital India Act highlights the importance of managing the complexities of the Internet and the rapid expansion of intermediaries.[150] This would necessarily entail the implementation of a robust rights-based framework in India to protect digital users, including the businesses exchanging data through these interfaces. The EU is seen as a regulatory leader in that aspect.[151] While it is beyond this paper's scope to compare India and the EU's non-personal data regulatory framework, some key differences indicate specific policy considerations.

[142] 'RBI to Take on Google, Microsoft with Its First-Ever Cloud Platform: Report' (*Livemint*, 18 November 2024) <https://www.livemint.com/news/india/rbi-fund-local-affordable-cloud-data-storage-pilot-program-2025-financial-firms-compete-global-providers-it-india-news-11731915287499.html> accessed 29 April 2025.

[143] Sen, Gopal and Manjunath (n 43); Jindal and Nigam (n 53).

[144] Ashutosh Mishra, 'Centre to Develop Non-Personal Data Collection Platform for India Inc' (*Business Standard*, 12 March 2024) <https://www.business-standard.com/industry/news/centre-to-develop-non-personal-data-collection-platform-for-india-inc-124031101007_1.html> accessed 29 April 2025; 'India AI: A MEITY Initiative' <https://indiaai.gov.in/hub/indiaai-datasets-platform> accessed 29 April 2025.

[145] Ministry of Corporate Affairs, 'Report of the Committee on Digital Competition Law' (2024) <https://prsindia.org/files/parliamentry-announcement/2024-04-15/CDCL-Report-20240312.pdf> accessed 29 April 2025.

[146] In Re: Updated Terms of Service and Privacy Policy for WhatsApp users, Suo Moto Case No. 1 of 2021, paras 28.7 and 245 <https://www.cci.gov.in/antitrust/orders/details/1158/0> accessed 29 April 2025.

[147] See various examples discussed in Bailey and Sane (n 90) 229-231; See also Chaudhuri and Joseph (n 114) 154, 169, discussing India's message at the G20 of using disaggregated data for development.

[148] Amrita Nanda and Astha Kapoor, 'Understanding Non-Personal Data Sharing a Principle First Approach' (The Data Economy Lab, 22 July 2021) 43.

[149] Streinz (n 9) 916.

[150] Proposed Digital India Act (n 14).

[151] Anu Bradford (n 9) chapter 9.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

First, unlike India, the Data Governance Act of the EU provides a comprehensive compliance framework for bodies acting as data trustees (or 'data intermediaries' in EU law),[152] including several requirements to ensure that data intermediaries function without any conflict of interest and by following clear requirements for data privacy and security.[153] Second, non-personal data sharing is largely voluntary under the Data Governance Act (termed 'data altruism').[154] Although the EU harbours a long-term ambition of developing sectoral data spaces,[155] the route chosen is through data altruism organizations that manage large data pools, subject to strict transparency and accountability requirements.[156] Unlike the Indian policy framework, the EU framework places heavy emphasis on technical requirements necessary to facilitate a data-sharing ecosystem, such as data formats, interoperability, and cybersecurity standards.[157] This is evident in both the Data Governance Act and the Data Act.

The Data Act, dealing with the sharing of IoT data in the EU, provides several safeguards to ensure that data are not shared to harm market competition. For example, it restricts explicitly 'gatekeepers' designated under the Digital Markets Act, 2022, (eg Google) from soliciting or commercially incentivizing users to share the data the user has received from another data holder (eg a banking app), or encouraging users to request a data holder to share their data with a designated gatekeeper, or receiving such data directly from the users.[158] In the case of business-to-business data sharing, the Data Act allows non-personal data to be shared only on fair, reasonable, non-discriminatory, and transparent terms.[159] Sharing of non-personal data between governments and businesses is restricted to emergency needs, such as a public emergency.[160] In the case of the Data Act, governments can request data in the prescribed manner from IoT companies only in the case of an 'exceptional need'.[161] Further, this request mechanism cannot be applied to SMEs/micro-enterprises.[162]

In contrast, the Indian policy framework neither has a clear focus on assessing the market value data nor does it provide a well-reasoned framework for the sharing of non-personal data for B2B or B2G data sharing. In fact, these omissions will strengthen existing market players as they can access the non-personal data of smaller companies. The Indian framework also provides unfettered discretion to government bodies and data trustees to mandate non-personal data sharing for public interest and sovereignty-related objectives.

While developing countries can learn from the experience of the EU, especially to foster a rights-based framework for data regulation, they must not mimic the EU approach. For instance, the Data Act in the EU places a new set of restrictions on cross-border flows of non-personal data,[163] an approach unlikely to be suitable for India, given its bureaucratic

[152] See generally Gabriele Caravano and Michele Finck, 'Regulating Data Intermediaries: The impact of the Data Governance Act on the EU's Data Economy' (2023) 50 CLSR 105830. Carovano and Fink argue that the EU Data Governance Act introduces the concept of data intermediaries with three primary aspirations: to form a European model for principles-based data sharing, reduce concentration in the data economy, and reduce the market power of non-European tech firms.

[153] See various conditions set out in Data Governance Act (n 4) arts 12 and Art 26: they cannot use the data for any other purposes except as specified by users, to prevent data intermediaries to impose pricing conditions based on using related services, etc. and any regulatory body overseeing data intermediaries must be legally distinct and functionally independent of data intermediaries.

[154] Data Governance Act (n 4) Chapter IV.

[155] See Commission, 'Commission Staff Working Document on European Data Spaces' SWD (2024) 1 final <https://digital-strategy.ec.europa.eu/en/library/second-staff-working-document-data-spaces> accessed 29 April 2025.

[156] See particularly Data Governance Act (n 4) arts 20, 21, 24.

[157] See Data Governance Act (n 4) arts 12(e), 12(g), 12(i), 12(l), 5(3), 12(c), 22.

[158] Data Act, art 5(3).

[159] ibid art 9.

[160] Data Governance Act (n 4) art 15(1)(a).

[161] Data Act, arts 14, 15 and 17.

[162] ibid.

[163] ibid art 32.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

complexity and unclear impact on digital innovation. Similarly, some aspects of the Data Governance Act and Data Act are bureaucratically complex;[164] the Indian government does not have the same degree of regulatory capacity to implement such complex measures. Most importantly, the Indian government must adopt a regulatory approach suited to its public and community interests, in the context of its socio-cultural values, digital development capabilities, digital education, and industry interests. For example, asking Indian digital users to opt out of anonymization is a largely futile exercise. Similarly, given innovative tech entrepreneurship in India, mandatory data sharing is likely to adversely majority of tech SMEs.

LEARNINGS FROM INDIA'S EXPERIENCE: HOW CAN DEVELOPING COUNTRIES REGULATE NON-PERSONAL DATA

The Indian experience demonstrates the uphill task that developing countries may face in developing non-personal data regulation. While non-personal data is undoubtedly a valuable resource, especially in developing countries facing a data poverty trap, its regulation is rife with several legal, policy, and market risks. Below, we highlight some key lessons that other developing countries must learn from India's brief tryst with non-personal data regulation.

Developing countries face critical challenges in building a successful digital economy, especially due to the dominance of successful multinationals and winner-takes-all market tipping in digital/data markets. This has resulted in repeated calls for dealing with data and digital colonialism. A report by IT for Change, a think tank based in India, argued that the free flow of data across borders hinders the ability of developing countries to facilitate equitable data sharing and thereby benefit from the economic value of data.[165] It further recommends that economic benefits can be derived from data sharing with data localization requirements, which require big foreign technology companies to store domestic data locally.[166]

In other words, data localization establishes a degree of territorial control over data, thus enabling it to better enforce its data sharing policies. Since unrestricted data flows are often equated to economic expropriation, policies on data localization and mandatory data sharing are increasingly being deployed together. Such thinking amongst policymakers in India and other developing countries has spurred data localization laws.[167] However, the Indian government's position has occasionally softened on this issue.[168] For instance, while the 2019 and 2022 versions of the draft data protection law contained explicit provisions on the localization of personal data, the law enacted in 2023 lacked a data localization requirement.[169] This position has, however, remained in a state of flux, with the draft Digital Personal Data Protection

[164] For example, data altruism organisations under the Data Governance Act are subject to complex requirements for registration (Article 19). Some experts have also expressed concern that the constitution of data intermediaries creates such high costs that it may not facilitate more data sharing and may not be economically viable. See Caravano and Finck (n 139).

[165] 'Cross-border 'Data Flow With Data Rights' (2022) IT for Change Policy Brief <https://itforchange.net/sites/default/files/2208/Cross-Border%20%E2%80%98Data%20Flow%20With%20Data%20Rights%E2%80%99.pdf> accessed 29 April 2025.

[166] ibid.

[167] Nigel Cory and Luke Dascoli, 'How Barriers to Cross-Border Data Flows Are Spreading Globally, What They Cost, and How to Address Them' (2021) Information Technology & Innovation Foundation Report <https://www2.itif.org/2021-data-localization.pdf> accessed 29 April 2025; Martina Ferracane, 'Data Localisation: Global Trends' (July 08, 2024) Robert Schuman Centre for Advanced Studies Research Paper, RSC 2024/21 <https://ssrn.com/abstract=4888145> accessed 29 April 2025.

[168] Suraksha P, 'Data Localisation Was Never a Mandate in India: IT MoS Chandrasekhar' (*The Economic Times*, 20 January 2023) <https://economictimes.indiatimes.com/tech/technology/data-localisation-was-never-a-mandate-in-india-it-mos-chandrasekhar/articleshow/97144650.cms?from=mdr> accessed 29 April 2025.

[169] Karthika Rajmohan, 'Data Localization: India's Tryst with Data Sovereignty' (*Tech Policy Press*, 23 January 2025) <https://www.techpolicy.press/data-localization-indias-tryst-with-data-sovereignty/> accessed 29 April 2025.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

Rules, 2025, specifying that 'significant data fiduciary' shall not transfer certain types of personal data 'specified by the Central government' outside the territory of India.[170]

Given the market realities and the government's constant struggle with controlling the flows of personal data and leveraging it for broad-based use, non-personal data appears to be an easy target for policymakers looking to address digital inequalities by treating it as a national resource. This approach can be both simplistic and counterproductive, as global data flows are essential for the rapid scaling of digital innovations across borders. However, the digital/data asymmetry and the global digital divide are challenges.[171] Therefore, in developing countries, non-personal data regulation has a relevant role in fostering growth and innovation, but must simultaneously account for critical public policy concerns and varied stakeholder interests.

First, protecting individual rights should lie at the core of non-personal data regulation. Many developing countries have nascent frameworks on personal data protection and insufficient regulatory experience. Further, they could have constrained budgets in setting up administrative frameworks necessary for the robust implementation of domestic laws on privacy and data protection. Nonetheless, without a solid legal framework for protecting personal data, it would be impossible to develop a framework for facilitating non-personal data sharing that addresses risks such as re-identification of individuals and data leaks. It may perhaps be prudent to adopt a 'wait and watch' strategy to assess how nascent frameworks on data protection are implemented and optimized, before setting up a comprehensive framework for non-personal data. This would also provide clarity regarding the scope of application of a country's personal data protection law and the role and success rate of its data protection authority in addressing privacy harms.

Second, as setting up data intermediaries or trustees is a common tool to enable non-personal data sharing, clear guidelines are necessary to ensure that such trustees can act transparently, accountably, and without any clear conflict of interest. Data trustees, at least theoretically, are expected to safeguard the community interests of the community whose data they hold and act towards their best interest. They are also expected to be key executors of data governance policies by ensuring non-personal data are governed in a way that privacy is not breached, and that data use is economically fair to various stakeholders.[172] Thus, data trustees must adopt sound data management techniques to protect an enormous amount of data from breaches, while also ensuring fair and equitable use of data.

In the Indian context, it is unclear if and how data trustees would fulfil the above requirements pertaining to stewardship and good governance. It is also likely that the characterization of data trustees in the Indian context could give rise to conflicts of interest, given that several existing and potential data trustees are linked (directly or indirectly) to the government. Further, no clear benchmarks exist to ensure their accountability, especially in the context of how they collect, facilitate or enable the sharing of non-personal data. Data trustees need to balance different public and commercial policy considerations to enable meaningful data sharing.

[170] Digital Personal Data Protection Rules, 2025, Rule 12(4).

[171] See generally World Bank, 'Digital Trends and Progress Report' (2023) <https://openknowledge.worldbank.org/server/api/core/bitstreams/95fe55e9-f110-4ba8-933f-e65572e05395/content> accessed 29 April 2025; Marilia Maciel, 'The Renaissance of Industrial Policy and Its Articulation With Data Governance' (*International Institute for Sustainable Development*, 15 January 2023) <https://www.iisd.org/articles/policy-analysis/industrial-policy-data-governance> accessed 29 April 2025.

[172] Ingrid Schneider, 'Data Stewardship by Data Trusts: A Promising Model for the Governance of the Data Economy?' in Claudia Padovani, Véronique Wavre, Arne Hintz, Gerard Goggin and Petros Iosifidis (eds), *Global Communication Governance at the Crossroads* (Palgrave Macmillan 2024).

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

Any dataset or data trust put in place by data trustees must use robust technological and organizational mechanisms for data security and privacy, including benchmarks for anonymization. The Indian government has set up infrastructures for extensive data sharing without implementing sufficient legal safeguards, risking individual rights, and reducing digital trust in the economy. Further, the Indian framework fails to evaluate economically efficient incentives for companies to participate in data-sharing platforms. For instance, the potential benefits of data licensing mechanisms, especially for SMEs, remain completely unexplored in the current framework.

Third, developing countries should not create overly complex or bureaucratic frameworks for non-personal data regulation and sharing. For instance, in India, the current framework does not address possible tension points between non-personal data regulation and other regulatory frameworks, such as competition law. The costs of managing these regulatory/jurisdictional conflicts can be particularly burdensome in developing countries. Similarly, managing asymmetric data-sharing frameworks (eg more regulatory burden on bigger companies to share data) is complex and challenging to implement for developing economies with minimal regulatory capacity. Further, in several developing economies, including those in Asia, regulators are tied to a particular government ministry rather than functioning as a completely independent body. Therefore, before adopting a comprehensive framework on non-personal data, mapping how existing regulatory frameworks (including sectoral frameworks) apply to non-personal data is essential. This includes assessing if a new regulator is necessary to achieve functions such as enabling more data sharing to foster competition in the domestic digital economy.

Fourth, the Indian experience suggests that alternatives to State-controlled and administered frameworks on non-personal data sharing exist. For instance, developing countries might benefit from pilot projects or data-sharing sandboxes (ideally in collaboration with trusted companies) to assess the regulatory and technological frameworks necessary to facilitate meaningful and high-quality data sharing. This may also provide practical knowledge regarding the categories of non-personal data that are safer to share, the best practices on anonymization, the capacity/resources of local companies to adopt such practices, and the ability of the regulators to understand technological and policy risks. In that regard, a multistakeholder participatory approach is more useful, especially involving relevant voices from the civil society and industry bodies (including SMEs) to build up bottom-up initiatives for data sharing.[173] Regulators must also focus on important factors such as the quality and relevance of such data. This may also be relevant in many developing countries where English may not be the most commonly spoken language.

CONCLUSION

Non-personal data regulation presents a new challenge for countries across the world, as policymakers grapple with balancing the potential economic and public benefits from enabling non-personal data sharing, and the potential risks arising from oversharing and overuse of such data. Further, as we have argued above, there are deep-seated concerns regarding weakening of privacy protection, enabling frameworks for state surveillance, possibilities of data expropriation, and adopting ill-informed frameworks on data governance. Further, the extent to which governmental intervention in non-personal data markets is effective remains unclear, especially accounting for the chilling effect of mandatory data sharing on pricing and market competition.

[173] Nanda and Kapoor (n 150) 20, 43, 54.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026

In this paper, we examined India's emerging framework for non-personal data protection and the key learnings it offers for developing countries. Although some countries, such as India, treat non-personal data primarily as a national resource of economic importance that can be reused and repurposed for varied uses, the paper discusses alternative ways of conceptualizing the role of non-personal data. For example, non-personal data can be shared using market mechanisms or a bottom-up commons-based approach. Market mechanisms can include the creation of a bundle of property rights over non-personal data, allowing private players to make efficient decisions regarding the generation, use, and sharing of such data, with the government overseeing the fairness of such transactions. This could create a functional and effective data market. Alternatively, a bottom-up commons-based approach can allow for decentralization of data governance by allowing a multitude of data trusts and intermediaries, provided they are held to high standards of accountability and that there is some degree of merit-based competition between these intermediaries. All these different approaches come with their challenges and benefits.

The Indian framework's overwhelming focus on economic benefit, its overreliance on anonymization, and its broadly state-led techno-political framework for non-personal data regulation present several challenges. As discussed above, some of the weaknesses of the Indian framework include unclear scope of non-personal data; sidelining of privacy concerns; improper alignment of economic incentives for commercial players in the market, especially small-sized tech companies; and inadequate transparency and accountability mechanisms for data intermediaries. Additionally, the framework vests a disproportionate amount of power in the government to control and demand access to non-personal data. Further, the framework does not address potential overlaps with competition law and intellectual property law. In the absence of a binding framework, India has already started adopting a techno-political infrastructure for widespread data sharing through mechanisms such as the DEPA and ONDC, despite no legal safeguards being in place.

Using the example of India, we propose four key lessons for developing countries. First, we emphasize the importance of individual rights while developing a framework for non-personal data in developing countries. Second, we highlight the importance of having clear and robust guidelines for any mechanism related to creating data trusts and data intermediaries. Third, we outline why developing countries need sufficient regulatory experience before implementing overly complex or bureaucratic frameworks for non-personal data regulation. Finally, we highlight the importance of bottom-up, experimental, and multistakeholder approaches to assess the extent to which developing countries have the regulatory capacity and domestic resources to utilize non-personal data sharing in a meaningful manner.

ACKNOWLEDGEMENTS

The authors would like to thank Niveditha Prasad for her research assistance and participants at the APSN 2024 conference, the special issue editors, Ching-Fu Lin and Han-Wei Liu, and Julia Hörnle for providing comments and suggestions on previous versions of the article. All errors remain ours.

FUNDING

None declared.

Downloaded from https://academic.oup.com/ijlit/article/doi/10.1093/ijlit/eaaf025/8443255 by guest on 29 January 2026