

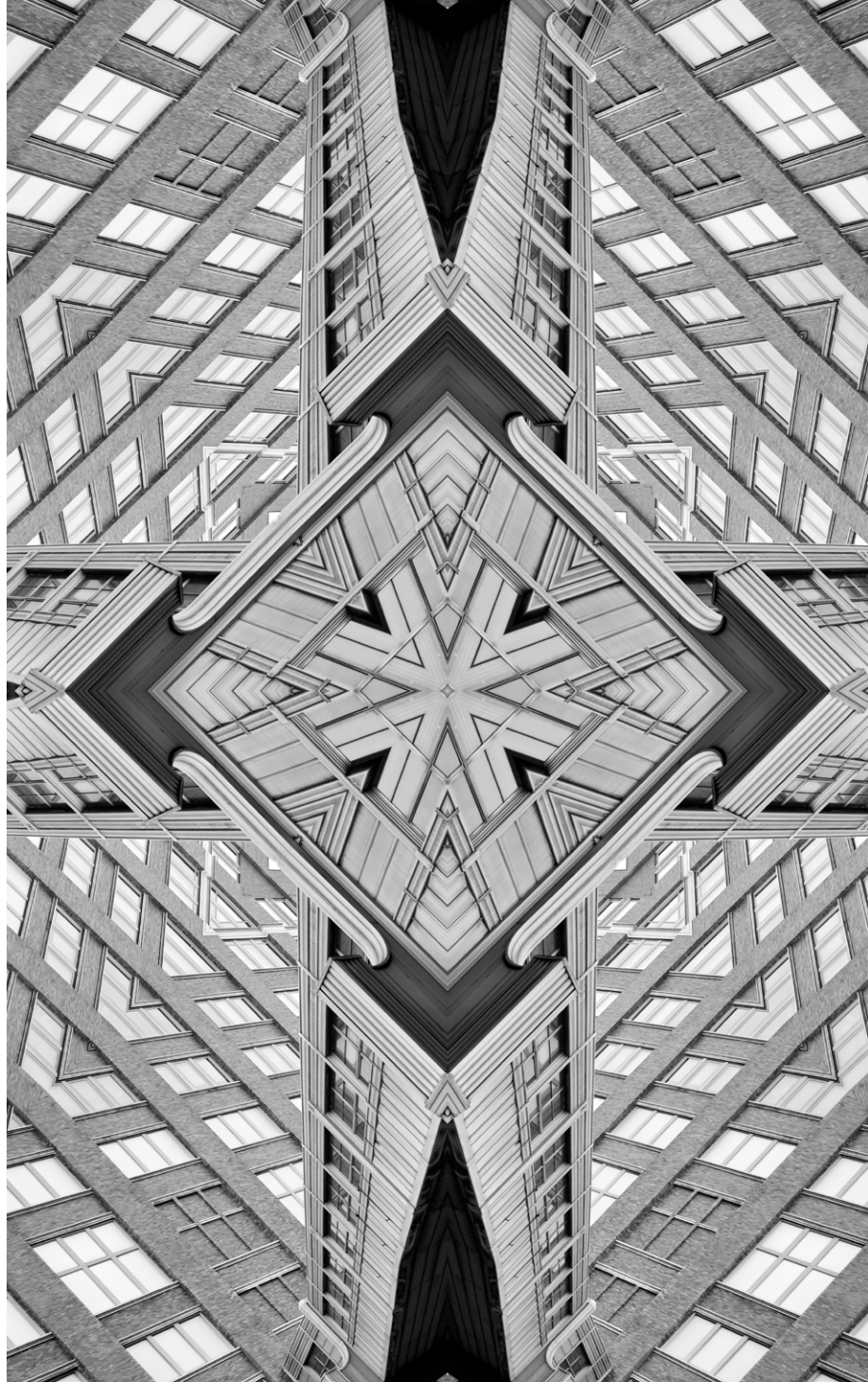
# Issue

---

# Brief

---

ISSUE NO. 631  
APRIL 2023



# Towards Evidence-Based Policymaking: India's Open-Data Initiatives

Sahil Deo and Amoha Basrur

## Abstract

India launched its Open Government Data Policy in 2012, and since then has taken further steps to increase public access to government data. However, these initiatives have met with limited success so far. Private stakeholders have had to step in to fill gaps in policy research data. This brief outlines current policies, the notable efforts made, and the existing repositories of data relevant for public policy. It examines the persistent challenges in making such data publicly accessible.

Data is essential to the formulation of evidence-based, timely, and relevant public policies. As India kickstarted policymaking in the post-Independence era through the Planning Commission's Five-Year Plans, the dearth of available data that could facilitate the design of sound policies immediately became clear. Through its Research Planning Committee, the Planning Commission contracted out research studies and sought help in procuring data as inputs for the research requirements of its various committees. In 1969, the Indian Council for Social Science Research (ICSSR) was set up, which in turn established research institutes across the country through which social science research was conducted.<sup>1</sup> Though privately owned, these institutes—30 at last count—obtain development and maintenance grants from the ICSSR. Even so, most of the responsibility of collecting data from citizens remained with the central and state governments. Such data was mostly acquired through field surveys conducted by agencies like the Census of India and the National Sample Survey Office (NSSO).

Following the economic reforms of the early 1990s, the trend of contracting data collection to private and non-profit organisations grew.<sup>2</sup> These organisations also store local and contextual data on their focus issues and interact more frequently than before with the government to implement interventions.

The early 1990s also witnessed the Right to Information (RTI) movement in India as a response to growing frustration with the lack of transparency in government. Civil society organisations, media organisations, and activists sought to empower citizens with the means to demand accountability from the government. Their long campaign led to the passing of the Right to Information Act in 2005,<sup>3</sup> which established a framework for citizens to access information in the hands of government. The law transformed the citizen-state relationship and shifted the default position of the government to openness.<sup>4</sup> It empowered citizens to demand answers by explicitly placing the responsibility of sharing information on the state, thereby setting the background for open-data initiatives.

In 2020, the Kris Gopalakrishnan Committee Report on Non-Personal Data Governance Framework<sup>5</sup> acknowledged the public interest in collective data once it has been appropriately anonymised and secured. It suggested a data-sharing architecture to promote access to non-personal data and foster safe data-driven development.

The rise of the internet and digitalisation across India gave many private and non-profit players the infrastructure and capacity to collect, store and analyse real-time dynamic data that can be useful for governance. For example, many private insurance companies store health data of their clients, mobility companies store data on transport trends, and non-government organisations conduct surveys to gauge the levels of education of citizens. The government, therefore, is no longer the sole collector and disseminator of data. The Gopalakrishnan Committee Report<sup>6</sup> also created a new taxonomy of businesses called ‘Data Business’ that meet a certain threshold of data management, and recommended that these players provide open access to their meta-data and regulated access to the underlying data.

These changes point to the need to structure a common understanding of the stakeholders in public data along with their roles, as well as the means and ethics of data collection, analysis, and dissemination. This brief explores the various initiatives of the government and other stakeholders to collect, handle and disseminate data relevant for policymaking. It describes newer practices of data collection and the emerging ethics of dissemination, and underlines the challenges faced by all stakeholders—the government, citizens, private sector, and civil society—in using data for governance.

“Digitalisation has given private players in India the infrastructure and capacity to collect, store and analyse real-time data for governance.”

Policymaking in India, which was heavily centralised in the first few decades after independence under the oversight of the Planning Commission, has since expanded to include other players. Citizens have increasingly become interested in governance, and demand transparency in decision-making from the state. Open Government Data (OGD)<sup>7</sup> is a concept, and increasingly a set of policies, which promote transparency, accountability, and value creation, by making government data accessible to the public. The state, over time, has looked to increase the credence and trust citizens have in vast quantities of data produced and commissioned by public bodies.

By encouraging the use, reuse and free distribution of datasets, governments are able to promote business creation and innovative, citizen-centric services. OGD initiatives, specifically the development of OGD portals, began around the world in the mid-2000s.<sup>8</sup> To expand the access to data for policy decisions, India implemented an open data policy in the form of the National Data Sharing and Accessibility Policy, 2012 (NDSAP).<sup>9</sup> It laid out guidelines for the release and use of government-held data, and established a framework for data sharing between government agencies and the private sector. The goal was to increase transparency, promote innovation and economic growth, and empower citizens to make informed decisions.

‘Digital India’ is a flagship programme launched by the government in July 2015 to transform India into a digitally empowered society and knowledge economy.<sup>10</sup> It focuses on building digital infrastructure to enable all citizens to have access to government services and participate in the digital economy, regardless of their location or socio-economic status. This push led to a further proliferation of policy-relevant data across sectors. Complementing the initiative, the Open Government Data Licence<sup>11</sup> of India was approved in 2017 to ensure that data sets that are released are not misused or misinterpreted, and that all users have the same and permanent right to use the data. However, so far, meta-data and data standards<sup>12</sup> have only been published by some, and not all ministries and states.

To strengthen policymaking, the government has also launched initiatives to promote the use of open data, such as the Smart Cities Mission,<sup>13</sup> which uses data and technology to improve urban services and infrastructure. Data is also at the core of flagship programmes such as Swachh Bharat Mission,<sup>14</sup> Housing for All,<sup>15</sup> One Nation One Ration Card,<sup>16</sup> Pradhan Mantri Ujjwala Yojana,<sup>17</sup> and fertiliser distribution programmes.<sup>18</sup>

# Recent Data Initiatives

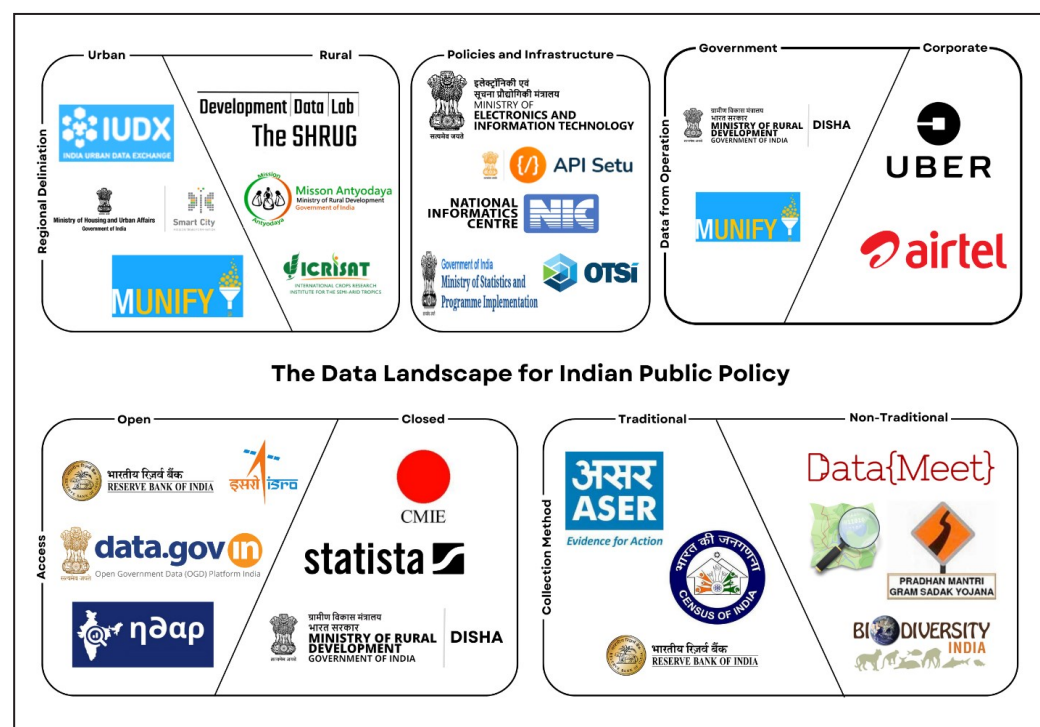
The National Informatics Centre (NIC) under the Ministry of Electronics and Information Technology (MeitY) is the agency tasked with providing technology-driven solutions to the central and state governments. It has developed a number of tools to facilitate the data economy. Its in-house tool called Darpan<sup>19</sup> is a web-based platform for administrators to monitor projects in real time without the need for coding or programming. It improves analytical capabilities by gathering data from multiple sources and consolidating it into a centralised, user-friendly platform. Similarly, Prayas<sup>20</sup> is a data-centric framework that allows the government to track and visualise the progress of its key programmes on a single platform. MeitY also created the API Setu,<sup>21</sup> an open API platform, in 2020, to build an open and interoperable digital platform to enable seamless service delivery across government. It also aims to promote innovation by making data from e-governance applications and systems available to industry and the public.

“Citizens have increasingly become interested in governance, and demand transparency in decision-making from the state.”

The data landscape in India has a number of stakeholders, each playing a unique role in the collection, dissemination and analysis of data. The primary player is the government and its various agencies involved in policymaking as well as implementing the OGD principles. Despite past and current initiatives, there are a number of gaps in the nature and quality of the data available for policy research.

The state has begun to partner with external organisations to fill these gaps. The organisations could either be analysing or adding value to government data, producing data themselves through traditional methods such as surveys, or using non-traditional methods such as crowdsourcing. Civil society also engages in independent projects. The ubiquity of digital or digitally enabled services today enables private businesses to have access to highly relevant data for policymaking across areas of health, mobility, and everyday life of users of their platforms. To highlight both the efforts that have been made as well as the gaps in the current data ecosystem, the next section analyses various data-related initiatives across sectors. It is not a comprehensive list of data sources and projects, but an illustrative one. Many sectors have both data and challenges in common, even if this is not explicitly mentioned.

## Figure 1: Stakeholders in Data for Public Policy and Governance



## a. Open Access vs. Closed Access

### *Open Data*

The OGD platform makes a wide range of government-generated data available to the public, including data on demographics, economic indicators, and social statistics. The data is available in various formats such as CSV, JSON, and XML, and can be accessed by anyone for free.

The National Data and Analytics Platform (NDAP)<sup>22</sup> is a data lake<sup>a</sup> developed by a private company, Object Technology Solutions India (OTSI), and launched by the government think tank, NITI Aayog. It is a single platform collection of all government-based data sets. It gives access to government-owned shareable data, along with information about its usage, in an open and machine-readable format. It publishes datasets, documents, tools and applications collected by the government, and encourages community participation with visualisation tools, APIs, and alerts.

The Reserve Bank of India (RBI) is the primary source of economic data in the country. It makes financial data available on its website, including data on monetary policy, inflation, and foreign exchange rates. The Database on Indian Economy<sup>23</sup> is its data warehouse—publishing state finance data, real-time data, and time series data on aggregates in a flexible and reusable format for analysis.

There are a number of specialised repositories for data relating to the natural sciences. The Indian Space Research Organisation (ISRO), for instance, runs the ISRO Science Data Archive (ISDA)<sup>24</sup> which stores data from Indian space missions. Bhuvan<sup>25</sup> is another ISRO initiative—an open data archive that provides free satellite data, products download facility, and thematic datasets. Meanwhile, the National Centre for Polar and Ocean Research (NCPOR) has its National Polar Data Centre<sup>26</sup> (NPDC)—an authoritative platform for managing and sharing polar research data, alongside tools to create visualisations online.

India has made significant efforts towards realising its OGD objectives. However, citizen engagement and data stewardship on most OGD platforms has been limited. There is no mechanism either to ensure that community-requested datasets are added, or for citizens to flag datasets following concerns about privacy, security, or potential misuse of data. Some other countries have provided this facility—Australia's *data.gov.au*<sup>27</sup> platform, for example,

---

a A data lake is a centralised, scalable repository that allows storage of structured and unstructured data in its original format for analysis and processing.



allows users to report issues with data sets and suggest improvements using either a feedback form or contacting the data custodian directly. The platform also provides a forum for users to discuss data-related topics and share their experiences and best practices.

On the OGD, there is also an excess of data aggregation that makes it challenging to find geographically specific information. On some platforms like ISRO's Bhuvan, development and application data tools to evaluate the ecosystem components are not yet complete. Bhuvan does not have the capacity to distinguish the finer details of disaggregated data. These capabilities need to be enhanced. For social science data, most datasets available are only at the state level, or occasionally the district level, and not at the individual village level. This hides important variations that exist at granular levels and limits the potential for innovation. Current data-sharing policies only apply to non-sensitive data and prioritise highly aggregated and anonymous data, preventing the sharing of detailed and valuable datasets.

### *Closed Data*

The Centre for Monitoring Indian Economy<sup>28</sup> (CMIE) is an economics think tank and business information company. It conducts a continuous Consumer Pyramids Household Survey<sup>29</sup> – a fast frequency indicator of living standards of Indian households. It also provides other valuable products including the Prowess Application for Credit Evaluation, databases on investment intentions (CapEx), and India's foreign trade (Tradedx). However, all these data sets are priced since CMIE is a private organisation. Websites like Statista<sup>30</sup> also offer a wide range of data but most of it is behind a paywall.

The data ecosystem cannot be heavily reliant on private organisations because these do not have the same obligations to make data freely accessible. This is all the more so when there are no valuation frameworks for data, no criteria for valuation or reference valuation models, and no guidance on pricing datasets.

## **b. Regionally Delineated Data**

### *Urban Data*

Data is at the forefront of urban planning decisions in initiatives such as the Smart Cities Mission. The Indian Urban Data Exchange<sup>31</sup> (IUDX) is an initiative of the Ministry of Housing and Urban Affairs, which offers a platform for the exchange of data among Indian cities. The IUDX platform is designed as an

interface for data providers and users, including urban local bodies, to share, seek and access data sets related to cities, urban governance and services. The India Urban Observatory<sup>32</sup> (IUO) website is another repository of data, visual resources and use cases for the urban ecosystem.

However, IUDX as a repository is fractured and incomplete, making it difficult to access the most recent and accurate information. IUO is also far from comprehensive. It hosts 134 data sets; New York City's open-data portal, for example, has 3,588.<sup>33</sup> There is no well-structured system in place to monitor data-sharing efforts by government bodies, which is essential to ensure that data sets are of high quality, and are useful for research and other purposes. The enthusiasm and vision with which open data initiatives began need to be sustained to ensure that these efforts are meaningful in the long run.

### *Rural Data*

Digitalisation missions are moving rapidly to generate data from India's rural districts. The Ministry of Rural Development's Mission Antyodaya<sup>34</sup> is a convergence and accountability framework to manage resources allocated by 27 ministries and departments under various programmes for rural development. In the non-government sector, the Development Data Lab's flagship project, the Socioeconomic High-resolution Rural-Urban Geographic Platform for India<sup>35</sup> (SHRUG), is a geographic platform based on administrative data that enables data-sharing among researchers studying India. It is an open-access platform comprising multiple data sets that provide details of 500,000 of India's villages and 8,000 of the country's towns over the past 25 years, using common geographic identifiers.

A crucial concern is that India currently lacks clear data privacy safeguards. The implications are two-fold. For one, micro-data that may be non-sensitive is often kept out of the public domain. Data sets such as unit-level cost of cultivation data, for example, are not made available to the public at all because of their relevance to crucial policy decisions like setting the minimum support price (MSP) of crops.<sup>36</sup> Meanwhile, individual-level sensitive data, such as electoral rolls or beneficiaries' lists of the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA), which contain substantial personally identifiable information, are available on a discretionary basis without any checks or safeguards.<sup>37</sup>

Ideally, all non-identifying data should be made publicly available at no cost. Sensitive data should be released in an aggregated form at the lowest level that will not cause harm, and personally identifiable data should be available to researchers and policymakers only through secure, standardised procedures, such as anonymisation, and should be accessible remotely on a controlled server. Official standards and tools should be put in place to assist data cells and chief data officers in protecting personal information.<sup>38</sup>

### c. **Traditional and Non-traditional Data Collection Methods**

#### *Traditional Data Collection*

Traditional data collection is carried out using surveys, interviews, scrutiny of public records, and administrative data. Education-related data is a mine field in Indian policy that is usually collected using traditional methods. Since 2005, for instance, the Annual Status of Education Report<sup>39</sup> (ASER)<sup>b</sup> has been providing reliable annual estimates of children's schooling status and basic learning levels for each state and rural district in India. In 2018-19, the Ministry of Education also initiated a large-scale management information system called Unified District Information System for Education Plus<sup>40</sup> (UDISE+). It has been collecting real-time data from all recognised schools that provide formal education from pre-primary to Class XII.

Various state ministries also collect similar data and the two sets are not always consistent. The Maharashtra government's School Education and Sports Department data sets, for example, are not consistent with those on UDISE+.<sup>41</sup> Data platforms and portals working in silos, with multiple government bodies maintaining their own data portals, result in manual, inconsistent, and delayed integration.<sup>42</sup> Data collection needs to be streamlined.

#### *Non-traditional Data Collection*

Non-traditional data collection has gained popularity because it is swift and cost-effective. The sources include crowdsourced science data, mobile phone data, social media data, Internet of Things (IoT) data, and remote-sensed data. Remote-sensed data, in particular, has a wide range of applications from terrain mapping and land use tracking to disaster management. The Pradhan Mantri

---

b ASER is a private effort carried out by the NGO Pratham through volunteers.

Geo Sadak Yojana<sup>43</sup> (PMGSY), for example, a nation-wide road connectivity plan, uses Online Management, Monitoring and Accounting System (OMMAS) software as a web-based system with a centralised database. However, the data is only available as individual pages on specific roads, making it hard to assemble into an analysable data set.

The need for usable data resources has also led to organisations, on their own, producing crowdsourced open data like DataMeet.<sup>44</sup> Members of this community have undertaken several projects to collate disparate government data, convert it into usable formats, and provide previously unavailable data such as geo-referencing PDFs and image maps of parliamentary constituencies and village boundaries. Another organisation, OpenStreetMap,<sup>45</sup> has built the biggest open geographical database in India. It is a free, publicly accessible geographic database created and maintained by volunteers who gather information through surveys, tracing aerial imagery, and incorporating data from other sources available under free licensing.

Citizen-led aggregation of science data includes the India Biodiversity Portal<sup>46</sup> (IBP), initiated by the Ashoka Trust for Research in Ecology and the Environment (ATREE), to provide information on all aspects of biodiversity in India. It harnesses collective knowledge, seeks voluntary participation of users, and has established a participatory platform for content generation, verification and usage, making it accessible through mapping technologies and geographic visualisation software. However, a concern with real-time and crowdsourced data is that it can be easily misused. Biodiversity data, for example, can become an aid for poachers.<sup>47</sup> There must be a balance between the speed at which all sensitive data, even when not related to personal identities, is made available. Care has to be taken on the extent of granularity to be provided. Other countries have found solutions. The Singapore government, for example, has issued an official guide to basic anonymisation<sup>48</sup> and also provides a free anonymisation<sup>49</sup> tool.

“The data ecosystem cannot be reliant on private organisations, which do not have the same obligations to make data freely accessible.”

#### d. **Data from Government vs. Data from Business Operations**

##### *Government Data*

Access to data related to outcomes of government schemes is vital to public accountability. Macromoney Research Initiatives, a Hyderabad-based private enterprise, created Munify<sup>50</sup> to tackle the inaccessibility of municipal reports and budgets. The Munify database is designed to be readable, and allows for query, comparison, and analytics of the financial performance of the different municipal corporations of India.

The Ministry of Rural Development has constituted District Development Coordination and Monitoring Committees<sup>51</sup> (DISHAs), which have a dashboard enabling elected representatives to track the performance of schemes in their constituencies. It brings together granular data from 42 national government schemes in real time in an accessible and structured format. The platform has room for improvement, however. More importantly, despite the fact that the data it hosts is not sensitive in nature, access is restricted to government officials. This limits the potential value of the platform for firms, developers, think-tanks, researchers and private citizens.

##### *Data from Corporations*

Uber and Airtel are two of many examples of private corporations sharing their data for use in public policy. The ride-hailing app Uber collects data that can provide insights into transportation patterns and used to improve transportation infrastructure and services. Recognising the public value of its data, Uber created Uber Movement,<sup>52</sup> an open and interoperable platform that provides data and tools for cities to understand and address urban transportation challenges more deeply. The telecom provider Airtel<sup>53</sup> has contributed its data to a project that uses mobile network data to help identify geographical locations at risk of increasing tuberculosis incidence. Airtel also has anonymised, aggregated mobile network data showing regular population movements. The scale, granularity and immediacy of mobile data enables identification of areas that have low TB incidence rates, but are at risk because they have high levels of communication with areas of high TB incidence.


# The Data Landscape

Despite the huge potential of such privately held data, the absence of a unified framework for sharing it with the government has impeded more prolific use. India still does not have a specific data protection law, with the bill seeking to introduce one repeatedly running into difficulties. Without such a law, it is vital to have a clear framework to protect citizen privacy and prevent misuse of data. The European Union's (EU) General Data Protection Regulation (GDPR)<sup>54</sup> sets out several articles that allow for sharing personal data with government authorities for lawful purposes, subject to appropriate safeguards. These include Articles 6(1)(c) and 6(1)(e), which provide the basis for processing personal data for compliance with legal obligations or tasks carried out in the public interest, and Article 89, which allows exemptions for archiving, research, or statistical purposes. The GDPR emphasises that all data sharing should be subject to appropriate safeguards through technical and organisational measures that ensure data minimisation, anonymisation, and accountability.

“Absent a data protection law, it is vital to have a clear framework to protect privacy and prevent misuse of data.”

India's OGD policy framework has strong fundamentals and a robust technological backbone. Recent initiatives like NDAP are powerful platforms with potential to facilitate analysis. However, gaps in the data landscape exist. They are further back in the pipeline and need to be addressed for access points like NDAP to be used to their fullest potential. Data collection needs to be standardised across organisations to ensure high quality and interoperability.

Data also needs to be regularly updated and shared, including inter-government sharing, to allow for discovery of data and prevent duplication of data assets. This calls for a model data-sharing toolkit to assist data officers in assessing and managing the risk associated with sharing and releasing data sets. In the absence of a data protection law, it is particularly important that privacy standards are prioritised from the source itself.

Official meta-data and data quality standards can be set across sectors, ensuring that these standards are met while uploading data on any OGD portal. Creating platforms and developing the technology to easily work with data is ultimately only fruitful if the platforms are regularly and accurately updated. In the past, the main drawback across sectors and repositories was the lack of comprehensive and timely data. Ensuring that departments follow through on their OGD commitments is the single most important step to improve India's data landscape. 

**Dr. Sahil Deo** is a co-founder of CPC Analytics, a boutique data-driven policy consulting firm with offices in Berlin and Pune.

**Amoha Basrur** is a research analyst at CPC Analytics.

*The authors thank Arindam Das for providing valuable insights and feedback during the interview process; and Abhay Pethe and Ovee Karwa for their constructive comments on the draft of this brief.*

- 1 Kuldeep Mathur, *Public Policy and Politics in India: How Institutions Matter*.(India: Oxford University Press, 2016).
- 2 Mathur, *Public Policy and Politics in India*
- 3 Ministry of Law and Justice, “Right to Information Act, 2005”, *The Gazette of India*, (2005), <https://rti.gov.in/rti-act.pdf>.
- 4 Glover Wright et al., *Open Government Data Study: India*, (2010), <https://cis-india.org/openness/publications/open-government.pdf>.
- 5 Gopalkrishnan Committee
  - *Committee Report: Report by the Committee of Experts on Non-Personal Data Governance Framework*, 2020, <https://ourgovdotin.files.wordpress.com/2020/07/kris-gopalakrishnan-committee-report-on-non-personal-data-governance-framework.pdf>.
- 6 Gopalkrishnan Committee
  - *Report*
- 7 Barbra Ubaldi, “Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives”, *OECD Working Papers on Public Governance*, (2013), <https://doi.org/10.1787/5k46bj4f03s7-en>.
- 8 OECD, “Open Government Data”, <https://www.oecd.org/gov/digital-government/open-government-data.htm>.
- 9 Government of India, “National Data Sharing and Accessibility Policy (NDSAP) – 2012” *The Gazette of India*, (2012), <https://dst.gov.in/sites/default/files/gazetteNotificationNDSAP.pdf>.
- 10 Government of India, “Digital India“, <https://digitalindia.gov.in/>.
- 11 Ministry of Electronics and Information Technology, “Government Open Data License – India“. *The Gazette of India*, (2017), [https://data.gov.in/sites/default/files/Gazette\\_Notification\\_OGDL.pdf](https://data.gov.in/sites/default/files/Gazette_Notification_OGDL.pdf).
- 12 Ministry of Electronics and Information Technology, “e-Governance Standards & Guidelines“, <https://egovstandards.gov.in>.
- 13 Ministry of Housing and Urban Affairs, “Smart Cities Mission“, <https://smartcities.gov.in/>.
- 14 Department of Drinking Water and Sanitation, “Swachh Bharat Mission“, <https://swachhbharatmission.gov.in/sbmcms/index.htm>.



- 15 Ministry of Housing and Urban Affairs, “Pradhan Mantri Awas Yojana“, <https://pmaymis.gov.in/>.
- 16 National Informatics Centre, “One Nation One Ration Card“, [https://www.nic.in/infographs\\_post/one-nation-one-ration-card/](https://www.nic.in/infographs_post/one-nation-one-ration-card/).
- 17 Ministry of Petroleum and Natural Gas, “Pradhan Mantri Ujjwala Yojana 2.0“, <https://www.pmu.gov.in/index.aspx>
- 18 Government of India, “Direct Benefit Transfer Mission“, <https://dbtbharat.gov.in/successstory/view?id=TkE9PQ>
- 19 National Informatics Centre, “Darpan: NIC- Dashboard Services“, <https://darpan.nic.in/>
- 20 National Informatics Centre, “Prayas“, <https://prayas.nic.in/>.
- 21 Ministry of Electronics and Information Technology, “API Setu“, <https://apisetu.gov.in/home/about>.
- 22 NITI Aayog, “National Data and Analytics Platform“, <https://ndap.niti.gov.in/>.
- 23 Reserve Bank of India, “Database on Indian Economy: RBI’s Data Warehouse“, <https://dbie.rbi.org.in/DBIE/dbie.rbi?site=home>.
- 24 ISRO, “ISRO Science Data Archive (ISDA). Indian Space Science Data Center“, <https://www.issdc.gov.in/isda.html>.
- 25 ISRO, “Bhuvan: Indian Geo Platform of ISRO“, <https://bhuvan.nrsc.gov.in/home/index.php>.
- 26 National Centre for Polar and Ocean Research, “National Polar Data Center“, <http://data.ncaor.gov.in/>.
- 27 Australian Government, “data.gov.au“, <https://data.gov.au/page/privacy-policy>.
- 28 Centre for Monitoring Indian Economy, “CMIE“, <https://www.cmie.com/>.
- 29 Centre for Monitoring Indian Economy, “Consumer Pyramids dx“, <https://consumerpyramidsdx.cmie.com/>.
- 30 Statista, “The Statistics Portal“, <https://www.statista.com/>.
- 31 Indian Urban Data Exchange, “Data Exchange Platform“, <https://iudx.org.in/>.
- 32 Ministry of Housing and Urban Affairs, “India Urban Observatory“, <https://iuo.mohua.gov.in/portal/apps/sites/#/data>.

- 33 NYC Office of Technology and Innovation, “NYC Open Data“, <https://opendata.cityofnewyork.us/overview/>.
- 34 Ministry of Rural Development, “Mission Antyodaya“, <https://missionantyodaya.nic.in>.
- 35 Development Data Lab, “The SHRUG“, <https://www.devdatalab.org/shrug>.
- 36 Arindam Das (Joint Director, Foundation of Agrarian Studies), in discussion with the author, Bangalore, India, March 2023.
- 37 Sam Asher et al., “Big, Open Data for Development: A Vision for India.” *India Policy Forum*, (2021), <https://www.ncaer.org/wp-content/uploads/2022/09/IPF2021Paper1.pdf>.
- 38 Asher, et al., *Big, Open Data for Development: A Vision for India*
- 39 ASER Centre, “ASER 2022“, <http://www.asercentre.org/>.
- 40 Department of School Education & Literacy, “UDISE+“. <https://udiseplus.gov.in/#/home>.
- 41 School Education and Sports Department, Government of Maharashtra, “Maharashtra Education Portal“, <https://education.maharashtra.gov.in/>.
- 42 Abhay Pethe (Professor, Mumbai University), in discussion with the author, January 2023.
- 43 Pradhan Mantri Gram Sadak Yojna, “Online Management, Monitoring and Accounting System“, <http://www.omms.nic.in/home/citizenpage>.
- 44 Data Meet, “Data Meet“, <https://datameet.org/>.
- 45 Open Street Map, “OpenStreetMap“, <https://www.openstreetmap.org/#map=4/21.84/82.79>.
- 46 Ashoka Trust for Research in Ecology and the Environment , “India Biodiversity Portal“, <https://indiabiodiversity.org/page/show/4246005>.
- 47 Adam Welz, “Unnatural Surveillance: How Online Data is Putting Species at Risk.” *Yale Environment 360*, September 6, 2017, <https://e360.yale.edu/features/unnatural-surveillance-how-online-data-is-putting-species-at-risk>.
- 48 Personal Data Protection Commission Singapore *Guide to Basic Anonymisation*, 2022, <https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Advisory-Guidelines/Guide-to-Basic-Anonymisation-31-March-2022.pdf>.

# Endnotes

- 49 Personal Data Protection Commission Singapore, “Data Anonymisation Tool“, <https://www.pdpc.gov.sg/help-and-resources/2018/01/basic-anonymisation>.
- 50 Munify, “Macromoney Municipal Database“, <https://www.munify.in/index.php>.
- 51 Ministry of Rural Development, “District Development Coordination and Monitoring Committees (DISHA)“, <https://dishadashboard.nic.in/>.
- 52 Uber, “Uber Movement“, <https://movement.uber.com/?lang=hi-IN>.
- 53 GSMA, “Helping end tuberculosis in India by 2025“, [https://www.gsma.com/betterfuture/wp-content/uploads/2018/09/Helping\\_end\\_tuberculosis\\_in\\_India\\_by\\_2025.pdf](https://www.gsma.com/betterfuture/wp-content/uploads/2018/09/Helping_end_tuberculosis_in_India_by_2025.pdf).
- 54 European Parliament, “Regulation (EU) 2016/679 (General Data Protection Regulation)“, (2016), <https://gdpr-info.eu/art-89-gdpr/>.

*Images used in this paper are from Getty Images/Busà Photography.*



Ideas . Forums . Leadership . Impact

20, Rouse Avenue Institutional Area,  
New Delhi - 110 002, INDIA  
Ph. : +91-11-35332000. Fax : +91-11-35332005  
E-mail: [contactus@orfonline.org](mailto:contactus@orfonline.org)  
Website: [www.orfonline.org](http://www.orfonline.org)