

# **Inputs to MeitY on the Draft National Data Governance Framework Policy**

Submission from IT for Change

June 2022

# Inputs from IT for Change<sup>1</sup> to MeitY on the Draft National Data Governance Framework Policy

11 June 2022

IT for Change has already provided its detailed comments on the earlier draft of this policy, which is now named the draft National Data Governance Framework Policy (NDGFP), released for public consultation by the Ministry of Electronics and Information Technology (MeitY). Since there are only a few substantive changes from that earlier document, our main response remains unchanged. This present response may be read in continuation to our earlier response that can be found [here](#).

At the outset, we appreciate the NDGFP for its recognition of the crucial role of data, specifically non-personal data (NPD), in transforming governance, achieving social development, and as the basis of economic innovation, growth, and equity. Being the central resource for the functioning of various sectors in a digital society and economy, it is critical that data is made available and accessible to a larger range of social and economic actors for obtaining the best public, social, and economic value from it. This needs to be done while ensuring the right to privacy, and protection against other data-related harms and exploitations. To that extent, as and where applicable, principles like data minimization, checking for data biases, being cautious about over-reliance on data (recognizing its limits), etc., are important to be applied. The policy document can do much more and better to present and discuss such principles, which pertain not just to personal data but also to various kinds of non-personal data, whether originating from personal data (through anonymization) or otherwise.

We see two important changes from the earlier draft; (i) the part on monetizing NPD in government hands has been dropped, and (ii) there are new mentions of the need for privately held data to be a part of 'India Datasets', with the hope that such data will be voluntarily shared by the data-holding private parties.

## Monetizing Access to a Public Good was Always Problematic

**We greatly welcome dropping the part on monetizing NPD that is in government hands.** Public goods, like important NPD, are best delivered for free. This is even more true for non-rivalrous resources like information or data. Seeking fees for access to such resources tend to lead to under-utilization of these

---

<sup>1</sup> IT for Change is an Indian NGO having a global footprint and is also in Special Consultative Status with the United Nations ECOSOC. For further information, visit our website (<https://itforchange.net/>) or write to us at [ITFC@itforchange.net](mailto:ITFC@itforchange.net).

resources, as well as their inequitable distribution. The loss in this way, with regard to the overall creation of public/social/economic value, and negative consequences for distributional equity, tends to be many times higher than the revenues that governments can generate by monetizing access to such resources.

The rest of our response is devoted to change (ii) as mentioned above, about the need for sharing privately held data for developing India Datasets, with the hope that it will be shared voluntarily.

## Today, the Most Important Data is Privately Held

First, **we welcome the important recognition by the draft NDGFP that privately held data is required for developing the much-needed India Datasets.** However, we are of the view that it is still not adequately understood and appreciated in government's data-related documents that **increasingly the most important, as well as extensive data in every sector is held by a few platform companies.** There is an erroneous assumption that governments still hold the most extensive and important data. This was true earlier, but definitely not in the emerging digital society. Just off the cuff, to convey the enormity of the shift, a single digital platform like Google collects many times more data in a city in a single day, perhaps even in an hour, than all government agencies put together would do in many years. Such data with platform companies in different sectors, and across sectors, is not only a part of the most extensive datasets, but these datasets are also increasingly the most important ones, socially and economically. They are important for undertaking governance, creating public and social value, and for economic enterprises and development. The sooner this fact is recognized and factored in by policymakers, especially while presenting concepts like India Datasets, the better.

In sum, privately held data is not just needed for developing meaningful and effective India Datasets – that can be highly useful for creating digital value for governance, social development, and the economy -- **such privately held data is increasingly far more important for developing India Datasets than data generated by governments.** Therefore, while we appreciate the focus on deriving the maximum value for society from data in government hands, which is the primary focus of the policy, **any meaningful and useful India Datasets policy and program needs to centrally address the issue of accessing important data that is held privately.** We hope and suggest that the NDGFP makes more explicit reference to this part and comes up with a relevant strategy and implementation plan.

## Data is Hoarded, not Shared Voluntarily

The second, connected issue is the naive **belief of the draft NDGFP that private holders of the all-important data will share it voluntarily for developing India Datasets.** It is not that large digital

corporations holding the most extensive and socially important data do not share any data at all. Data sharing initiatives of mega digital corporations like Google, Facebook, Uber, etc., are well known. But, given that exclusive hoarding of socially most important data is currently considered by them as their key comparative business advantage, these corporations do not, and will not, share the more important data that they hold. Such exclusive hoarding of data is precisely what gives them the opportunity to derive rents from it. (We must point out that it is not necessary that exclusive hoarding of data be the main comparative business advantage in a digital economy; the comparative advantage can, and should, shift to being able to best derive insights/intelligence from data to provide better services employing shared infrastructure of socially-important data.)

Some genuine corporate social responsibility-driven data sharing by these corporations does take place, but only the kind that does not threaten their dominance/monopolization in their respective areas of business. **Such dominance may also extend to competing with governments in providing key social services (in areas like health, education, agriculture, public transport, MSME support, and so on), which forecloses voluntary sharing of relevant, important data with governments.** Most data sharing by corporations that does happen is motivated by the objective of locking-in consumers, partners, etc., and to stall regulatory interventions in the direction of mandated sharing of socially important data held by them. Basically, the point that is being made here is that while the main holders of society's data, today, namely a few monopolistic digital corporations, may indeed share some of the data they have, such sharing is driven by their own interests. It certainly cannot serve as a sustainable basis for developing and supporting the needed India Dataset, this being the key objective of NDGFP.

**The otherwise laudable objective of the draft policy to set up India Datasets is undermined by two dubious assumptions; (i) today, governments still hold the socially/economically most important and extensive data, and not a few digital corporations, and (ii) the corporations holding this data – that may be of infrastructural nature and value for effective governance, addressing our many social needs, and for a robust and equitable digital economy -- will voluntarily share the data for developing India Datasets.**

Therefore, as the needed corrective, the NDGFP must recognize the imperative of mandated access to socially/economically important data that is held by a few digital corporations, as an unexceptional and key basis for developing the much-needed India Datasets. We understand that a policy by itself cannot mandate data sharing, and it will require the force of an appropriate law. **However, any policy instituting the concept of India Datasets will be considered half-hearted and largely ineffective if it does not address the need for mandating access to important privately held data.** We are of the view that the

NDGFP should explicitly and prominently recognize this context and need, and **develop arguments for moving towards a data sharing law**. We find that the **draft report of the Committee of Experts on Non-Personal Data Governance Framework provides important material in this regard, also making specific recommendations**. The NDGFP may make reference to that draft report and take it towards implementation.

## Evidence and References for the Need for Mandated Data Sharing

Below are a few arguments and facts that should leave no doubt in anyone's mind that **voluntary sharing of the 'most-needed and useful data' held by corporations is not going to happen**, and that any hope in this regard is unrealistic. Such an unfounded, nay demonstrably false presupposition is ever more dangerous because **the time is short for countries like India to assert the right to its data for its collective purposes and benefits -- what has been termed data sovereignty**. As the currently dominant digital society and economy model concretizes – where a few global corporations, almost all foreign, control most of the data about us, there is unlikely to be many possibilities ever of rolling things back. Such is the nature of data and digital intelligence-based dependencies.

1. Even as we write this response, India is leading the struggle of developing countries at the WTO seeking that Big Pharma shares its Covid-related IP (intellectual property) as an exceptional measure to help mitigate the biggest public health crisis in a century. We are not making much headway in this regard. So, when big corporations, with their eye solely on maximizing profits for their shareholders, will not share their IP even to address the biggest health crisis of our times, one is unable to understand the basis of this draft policy's hope that digital corporations will voluntarily share the socially important data they hold (when its exclusive access is so profitable to them).

2. When the European Union (EU) brought in the General Data Protection Regulation (GDPR) it sought a self-regulation framework for interoperability between cloud services. A few years later, the EU now explicitly recognizes that such a 'voluntary' approach has failed. Its draft Data Act now makes cloud interoperability mandatory. EU's draft Digital Markets Act mandates interoperability between instant messaging services, and such a mandatory requirement will in time be extended to social media platforms, and then also to other kinds of platforms. So, if digital corporations are not willing to even interoperate voluntarily, despite being urged by law (GDPR), (meaning they are not ready to voluntarily share their 'network resources'), what is the basis of hoping that they will voluntarily share their most valuable data resources, as hoped for in the NDGFP?

3. The EU is increasingly convinced about the need to introduce mandatory sharing of important data. Its draft Digital Markets Act mandates such sharing of data by dominant search engines with their competitors, and by all dominant platforms with business users who may have been instrumental in the generation of the relevant data. The draft Data Act takes mandated data sharing even further, mandating sharing of some privately held data with public agencies, and sharing of data back with users that generate data arising from IoT devices, etc. The recent draft Health Data Space Regulation brings up extensive requirements for mandatory sharing of many kinds of health-related NPD. Similar legislations are expected in other sectors as well, like transport, education, health, energy, and so on. The EU is facing the prospect of digital colonization by US (and China) -based digital corporations, and these are its responses in trying to assert digital and data sovereignty. The EU's situation in this regard is increasingly similar to that of developing countries like India, which face the likelihood of digital colonization even more. In this background, the NDGPF's silence on mandated data sharing is inexplicable.

4. The EU's 2019 *Report on Competition Policy for the Digital Era*, on viewing data as an important input for online services, production processes, logistics, smart products, and AI, recognizes that there will be cases "where duties to ensure data access – and possibly data interoperability – may need to be imposed". Similarly, the 2019 report of the UK's Digital Competition Expert Panel takes cognizance of the fact that "large data holdings are at the heart of the potential for some platform markets to be dominated by single players", and argues against relying on a voluntary approach. Noting that a voluntary approach would be insufficient because "businesses' natural incentives do not line up with delivering these functions", it recommends adopting data openness as a tool to promote competition in the digital market.

5. To these official documents we may add (as mentioned below) a few illustrative examples from fast-expanding academic work pressing for the urgent need to institute mandated data sharing regimes.

5.1. In the book *Access Rules: Freeing Data from Big Tech*, Thomas Ramge and Viktor Mayer-Schonberger argue that freeing of data hoarded by large digital corporations like Big Tech can be a tool for empowering the larger economy and society, for it facilitates information access, disrupts monopoly power, and creates a level playing field. They also make a case for mandatory data sharing, especially in situations where denying access to data would be equivalent to an abuse of market position, in their book *Reinventing Capitalism in the Age of Big Data*.

5.2. Graef and Prufer<sup>2</sup> employ the example of online search engines to establish the need for mandatory data sharing. They point out that in data-driven markets, the data that arises as a byproduct of a user's search becomes an input for innovation for the service provider and allows them to adapt to better suit users' preferences and gives them competitive advantage. Since this is collected as a byproduct of the search and transaction processes, it is virtually free. And in such a market, mandatory data sharing is the only way out because a new entrant in the search engine market can only access good quality data by actually serving customers, data about whom only incumbents have access to.

5.3. In his paper on *Data and Digital Intelligence Commons*<sup>3</sup>, Singh observes that it is fallacious to believe in any kind of meaningful voluntary sharing of key sector data, other than that at the margin which big players may find harmless to share.

5.4. A report of the University of Chicago's Stigler Committee on digital platforms observes, "Digital platforms generate several concerns across different fields, all linked to the power of data. To address these concerns in a holistic way, there needs to be a single regulator able to impose open standards, to mandate portability of and accessibility to data, to monitor the use of dark patterns and the risks of addiction..."<sup>4</sup>

5.5. In the paper "*Data as a common in the sharing economy: a general policy proposal*", Bruno Carballa Smichowski presents a general policy proposal based on making data as a common in the sharing economy using reciprocity licenses.<sup>5</sup>

5.6. The authors of the paper on *Economic Governance of Data*<sup>6</sup> highlight how academic discourse, "recognizing its limitations to address the central digital society problem of a few global corporations collecting and controlling the bulk of world's data", has moved beyond voluntary sharing to "FRAND (fair, reasonable and non-discriminatory) -based data sharing as happens with

---

<sup>2</sup> Inge Graef & Jens Prufer. (2018). *Mandated data sharing is a necessity in specific sectors*. *Economisch Statistische Berichten*, Vol. 103 (4763).<https://ssrn.com/abstract=3206685>

<sup>3</sup> Parminderjeet Singh. (2019). *Data and Digital Intelligence Commons (Making a Case for their Community Ownership)*. Data Governance Network Working Paper 02. <https://ssrn.com/abstract=3873169>

<sup>4</sup> Stigler Centre for the Study of the Economy and the State. (2019, September). *Final Report of the Stigler Committee on Digital Platforms*. Retrieved from <https://www.chicagobooth.edu/-/media/research/stigler/pdfs/digital-platforms---committee-report---stigler-center.pdf>

<sup>5</sup> Bruno Carballa Smichowski. (2016). *Data as a common in the sharing economy: a general policy proposal*. Centre d'Economie de l'Université Paris Nord. hal-01386644

<sup>6</sup> Parminderjeet Singh & Anita Gurumurthy. (2021, January). *Economic Governance of Data*. Data Governance Network. <https://itforchange.net/sites/default/files/1880/Economic-governance-of-data.pdf>

standards essential patents, and compulsory licensing on the model of health emergency related access to IP protected medicines.”

## The NDGFP Needs to Match the Highest Political Vision

One has been hearing pronouncements from the highest quarters of the Indian political establishment, including from the Prime Minister, about the central importance of data as key to owning and controlling our future. In his address to the World Economic Forum in 2018, **the Prime Minister asserted that data is a huge asset and that “the flow of global data is creating the biggest opportunities and the greatest challenges”**. **“He who controls the data controls the world.”**<sup>7</sup>

In light of such imposing and clearly true assertions, the vision and plans presented in the NDGFP for gathering, reclaiming, sharing, and making available our data to and for our people, governments, and enterprises are extremely weak and lukewarm. They reach nowhere close to what is urgently needed for India and Indians to control our own data, and thereby, our own future. **A meaningful NDGFP, and India Dataset strategy and implementation plan, would be the right direction to take, but it requires recognition of the need for some kind of mandated sharing of data that is currently hoarded exclusively by a few corporations.** Such data is needed to be shared much more widely, in the form of India Datasets, among various social and economic actors both to (i) realize its best value in an equitable manner, and (ii) escape various kinds of imminent economic, social, cultural, and political controls over India, and its people, by a very few, mostly foreign corporations.

Digital economic domination and control can quickly turn into political domination and control, which underpins the increasingly credible digital colonization discourse.

---

<sup>7</sup> Arijit Barman. (2018, January). *India means business in times of protectionism: PM Narendra Modi*. Economic Times. [https://economictimes.indiatimes.com/news/politics-and-nation/india-means-business-in-times-of-protectionism-pm-narendra-modi/articleshow/62627968.cms?utm\\_source=contentofinterest&utm\\_medium=text&utm\\_campaign=cppst](https://economictimes.indiatimes.com/news/politics-and-nation/india-means-business-in-times-of-protectionism-pm-narendra-modi/articleshow/62627968.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst)