

# **Framing India's Non-Personal Data Framework: Comments on the Draft Report by the Committee of Experts on Non-Personal Data Governance Framework**

*Smitha Francis\**

The Draft Report by the Committee of Experts on Non-Personal Data Governance Framework (henceforth, the Report) has covered a broad spectrum of economic, socio-political and ethical issues related to rights over Non-Personal Data (NPD). An attempt has been sought to achieve a balance between the use of data for innovation and access to data for increasing competitiveness of Indian start-ups and other firms.

Data and data-based intelligence have become the prime drivers of many of the new technology systems.<sup>1</sup> The emergence of advanced data storage, computing power and networking abilities, and the associated emergence of the inter-related technology systems driven by cloud computing, automation, digital platforms, the internet of things, artificial intelligence (AI), etc. are enabling the reorganisation and transformation of not only services, but also the manufacturing and agricultural sectors.<sup>2</sup>

While India is yet to formulate a coherent national digital development strategy, non-personal data protection along with personal data protection (PDP) are fundamental to the normative basis for an inclusive and equitable 'digital development trajectory' for India. Large amounts of evidence from successful technological catch-up by developing countries during earlier technological

---

\* The author is a Consultant with the Institute for studies in Industrial Development (ISID), New Delhi.

<sup>1</sup> Francis, Smitha (2019), 'Catching-up in the Digital Economy: The need for rapid policy evolution', ISID Discussion Note No. 2019/01, Institute for studies in Industrial Development (ISID), New Delhi.

<sup>2</sup> See Francis, Smitha (2018) 'Evolution of Technology in the Digital Arena: Theories, Firm-level Strategies and State Policies', *Working Paper No. CWS/WP/200/47*, Centre for WTO Studies, New Delhi.

revolutions point to the critical role played by regulatory frameworks and institutions. The latter went beyond just correcting market failures, but envisioned and paved the path for synergetic societal transformations and inclusive growth paths. At the present juncture, data protection, competition policy and tax policy are fundamental regulatory interventions required for meeting State's duties towards protecting citizens' rights.

While 'data' has intrinsic value,<sup>3</sup> currently it is largely discussed in the context of its 'economic value'. Access to data and its availability for wider use/re-use are among the most important considerations in the current era of digital intelligitisation. Needless to say that the need for the protection of privacy rights and cyber security/national security remain non-excludable fundamental priorities and state duties.<sup>4</sup> Additionally, the debate on sovereignty over the data that 'gets generated within' or 'exists/is inherent to' a nation's territory cannot be left in a vacuum for a number of reasons related to an inclusive and sustainable development trajectory, as briefly discussed below.

As recognised by now by regulatory authorities globally, the technological (and therefore, capital) asymmetries between leading global digital corporations such as Google, Apple, Facebook, etc. and follower firms are getting widened through the formers' control over extracted data and derived intelligence. The pre-dominant business strategies that the lead firms engage in to consolidate their monopoly positions and for erecting entry barriers for new entrants have been data-based for some years now. In particular, their patented algorithmic designs for data mining and analytics, and monetisation of the extracted data and digital intelligence lead to an entrenchment of their first movers' advantage.<sup>5</sup> This has several adverse implications not only in terms of the ability of Indian start-ups and other firms across sectors to

---

<sup>3</sup> The author is grateful to Rajeswari Raina for highlighting this point.

<sup>4</sup> See Report of the One-Day National Workshop On Catching-up in the Digital Economy: Identifying India's Policy Gaps and Challenges, 31 January, Institute for Studies in Industrial Development (ISID), New Delhi, jointly organised by ISID and Shiv Nadar University, available at [http://isid.org.in/wp-content/uploads/2020/03/ISID\\_Digital\\_Economy\\_Workshop\\_Report.pdf](http://isid.org.in/wp-content/uploads/2020/03/ISID_Digital_Economy_Workshop_Report.pdf)

<sup>5</sup> Francis (2018).

emerge and excel in new technologically important activities, but also for the ability of existing firms across sectors to survive digitalisation.

Foreign digital corporations (and other foreign corporations) are buying up rival Indian players in their respective sectors as well as firms in other sectors that have developed innovative products and services. This is not restricted to e-commerce, media, marketing, etc., but is also happening in health, education, transport, electronics, energy, etc. For example, Amazon has been not just into e-commerce and cloud computing services, but also into pharmaceuticals and e-governance services, among many others. Google has moved into urban planning, and not just autonomous vehicles. In the absence of a national data protection framework, *de facto* control over data is being exercised/enjoyed by those who first extract it digitally. That is, foreign ownership means that they exercise *de facto ownership* over humungous volumes of Indian data, which is being used to build digital intelligence in whichever sector they move into.

Greater the data for analytics and predictive modelling, greater is the revenue potential for the owners of data from monetising it. Greater the volumes of data, greater is also the innovation that follows it for future revenue generation. Given that advancements in AI, network technologies, automation, cloud robotics, blockchain, etc. are all also contingent on data-based intelligence and machine learning, etc. this *de facto* control over data gives digital corporations immense advantages in these emerging technologies also. This status quo has huge adverse consequences not only for competition in these sectors, but also for innovation opportunities and citizens' access to public services in health, transport, etc.

Any such *de facto* control over extracted data means that the immense opportunities to small businesses and other entities that build new products and services utilising the platforms owned by global digital corporations do not usually lead to the same wealth generation capabilities for them as the owners of the platforms or cloud (or Blockchain or AI).<sup>6</sup> The value share is maximum for the latter.

---

<sup>6</sup> Francis (2018)

The highly concentrated structure which we see in many a digital markets consisting of a monopoly or a duopoly (Google in internet search, Amazon in e-commerce, Facebook in social media, Google and Apple in mobile operating systems, and so on), will tend to get mirrored across economic sectors. This will have anti-competitive implications for not just the respective sectors, but also for the 'not-yet-digital' sectors. It therefore has adverse impact on innovative potential in the economy as a whole (and therefore, on national development).

So far, the Competition Commission of India has not considered the excludability conditions imposed by platform and other digital corporations to leverage their network effects and control over digital intelligence. Both of the latter (i.e., network effects and control over digital intelligence) generate significant increasing returns to scale for these digital corporations.<sup>7</sup> Therefore, mandatory data sharing is the only manner in which start-ups and other firms or any other entity will be able to overcome the entry barriers created by network effects to enter existing and new digital markets, let alone be able to compete with the leading corporations.

Some analysts argue that 'free flows of data' are fundamental to increasing the competitiveness and efficiency of India's firms, whether in agriculture, manufacturing or services. They argue that mandatory data sharing and data localisation will stand in the way of India benefitting from digital innovations.

But we must recognise that as production processes<sup>8</sup> as well as the machinery and equipment are being transformed by advances in digital technologies, many supply chains are getting transformed into digital value chains. Data and digital intelligence are driving the digital transformations in the manufacturing and agricultural sectors.<sup>9</sup> The increased use of sensors in devices and application-driven

---

<sup>7</sup> See Francis (Forthcoming), 'Digital Transformations and Structural Exclusion Risks: Towards policy coherence for enabling inclusive trajectories' in Das, Keshab et al *The Digitalization Conundrum in India: Applications, Access and Aberrations*, Springer, Singapore, ISBN 978-981-15-6906-7.

<sup>8</sup> Direct ICT application areas include control technologies, advanced visual and physical human-machine interfaces, navigation and perception technologies, monitoring and diagnostics devices, locomotion technologies and integrated product-process-production system design and simulation techniques, etc. (Alcorto 2014).

<sup>9</sup> See Francis (2018) and Sampath, Padmashree Gehl (2019) 'Regulating the Digital Economy: Dilemmas, trade-offs and potential options', Research Paper No. 93, South Centre, Geneva.

machines, and equally importantly, the growth in networked devices are continuously increasing the scale and scope of real time data extraction. All kinds of public data - whether of utilities usage, traffic, domestic and international trade and financial sector transactions, health, farming practices or the weather, environment or ecosystems - are also the raw material for analytics-based innovation.<sup>10</sup>

Given the data-centricity of AI/deep-learning-driven design and new product/process developments in manufacturing, precision agriculture, etc., the largest share of value creation ('efficiency gains') will accrue to the owners of the software/intelligence based on data (including those collected through devices and machines, of machines, people and matter in production and non-production spaces).

That is, data and data analytics are the downstream segments of digital value chains, on top of which come design and product development. Companies controlling these downstream segments of value chains within manufacturing, agriculture or services will extract the maximum rents/value within those digital value chains. Furthermore, setting technical standards for emerging products and services in the digital sphere will also be influenced by such control over and access to data.

Thus the manner in which access to data is managed by public policy will have critical implications for India's digital development trajectory. At a society-wide level, this has serious adverse implications for any entity in every sphere of life as every activity (or matter) becomes connected digitally.

Sharing non-personal data collected by public and private entities will not only lead to increased transparency, distribution of gains and more innovation, indeed it becomes essential to ensure that Indian economy and Indian citizens will derive the maximum value share from the digital transformations across sectors. A balanced NPD framework is needed for India also to harness the value from its strength in data analytics and design capabilities.

---

<sup>10</sup> See Singh, Parminder Jeet (2018) "Digital Industrialisation in Developing Countries — A Review of the Business and Policy Landscape", Report prepared for the Commonwealth Secretariat, London.

Within a national digital transformation framework, regulatory oversight of non-personal data will also enable the government to ensure that AI and automation happen in a regulated pace (matching with digital skill development) and in a more equitable manner in employment-intensive activities and sectors. This is fundamental for India to address her problems related to employment in the face of digitalisation.

### **Specific comments on the Report**

Chapter 3 discussing 'Value generation from data and the Case for Regulating Data' mentions that "Data therefore is increasingly taking the centre-stage in core-technological businesses, all economic sectors around the world and in addressing various social and public administration issues." However, the kind of digital transformations that are considered are exclusively those in the services sector. This section could incorporate some of the aspects discussed above.

A number of specific issues that need to be clarified in the Report are given below:

- As several analysts have pointed out that the risks of re-identification of 'anonymised' personal data are rather high, and as the Report itself has acknowledged it, there should be some clear benchmark to define non-identifiable personal data. This is essential to clearly earmark whether re-identified personal data will come under the NPD framework or under the purview of the PDP bill (which now includes "inferred data") currently under debate in the Parliamentary Committee.
- Under 'Defining non-personal data', the distinction between public non-personal data and community non-personal data must be made clearer. For instance, if "non-personal data collected or generated by the governments, or by any agency of the governments, and includes data collected or generated in the course of execution of all publicly funded works", then how is it that "datasets collected by the municipal corporations and public electric utilities" are given as an example under Community non-personal data?

- The Report maintains that raw/factual datasets comprising anonymised user-information data collected by private data custodians (such as telecom, e-commerce, ride-hailing companies, etc.), may be considered Community Non-Personal Data. As larger proportion of economic activities have moved online with accelerated digitalisation post-COVID-19 and as digitalisation of agriculture and manufacturing increases alongside that of services, it may be emphasised that such Community non-Personal Data sets will feed into sectoral and macroeconomic data sets. It must be noted that e-commerce data and payments for imports through e-commerce firms will become an integral part of trade data.<sup>11</sup> (Such data may also become part of price inflation data.)
- That is, apart from the 'Data Source' and 'Data Subject' logic used in the Report for articulating the legal basis for rights over Community NPD, the mandate for community and public rights over NPD collected by private data custodians must be made in terms of its requirement for national statistics database purposes fundamental to policymaking, for taxation purposes, as well as for guiding public provisioning to meet state's constitutional duties towards citizens. Thus the formulation and prescription of meta-data directories by the Report is critical.
- In light of the discussion above, it is evident that at this stage in India's development, incentives for innovation are adequately provided for in 5.4. (iii). (Algorithms / proprietary knowledge have not been considered for mandatory data sharing.)
- Open source algorithms/application architectures/cloud architectures, public digital infrastructural provision, etc. are among the ideal options to chart equitable digital transformation in a highly unequal society like India in order to enable the large mass of SMEs and other entities to benefit from digital innovations. While working towards the same, it may also be essential to mandate algorithm impact assessments for data businesses

---

<sup>11</sup> Kallummal, Murali (2020) 'Global Digital Trade: Emerging Challenges from the Transformation of Trade Flows', Presentation at One-Day National Workshop on 'Catching-up in the Digital Economy: Identifying India's Policy Gaps and Challenges', ISID, New Delhi.

above a certain level, which assess the outcomes of different algorithms on groups as well as any collective data-driven harms.<sup>12</sup>

- The NPD Authority must have the powers to enforce sanctions for non-compliance.
- Section 5.1 (i) Data sovereignty: There is a conflict between the first sentence and how the Report defines NPD. There is also a conflict between the first and the second sentence. These must be corrected. Data sovereignty should be about all Non-Personal Data (as defined) collected in/from/on India or by Indian or foreign entities.
- Section 7. I (Data sharing for sovereign purposes) does not define the “necessary and proportionate” standard for government access to private NPD and community NPD for national security and law enforcement purposes. This is necessary to ensure privacy rights.

September 2020

---

<sup>12</sup> Martin Tisne 'The Data Delusion: Protecting Individual Data isn't Enough When the Harm is Collective', FSI White Paper, Cyber Policy Centre, Stanford University.