



Policy Opinion

Comments on the Revised Report by the Committee of Experts on Non-Personal Data Governance Framework

BY

CENTRE FOR THE DIGITAL FUTURE

FEBURARY 2021



The Centre for The Digital Future (CDF) <https://cdfresearch.org>, is an independent research institution within the not-for-profit India Development Foundation (www.idfresearch.org). Our diverse team possesses knowledge across sciences, economics, policy, law and technology and rich experience in policymaking, regulation, industry, associations, think tanks and academia.

We at CDF aim to positively impact public policy and practices in the realm of digital and data ecosystem through actionable recommendations with evidence-based research and insights so that business related policies become robust, predictable and deliver the best outcomes. We believe that our approach will help India and Indian citizens with a greater value accrual from the digital economy.

Please find below our considered views on the subject matter captioned herein above:

1. The issue regarding extraction of value from data generated by Indians should not be mixed up with discussions on who owns, accesses or controls such data or, with issues regarding privacy, or security. Ownership of an asset is not a pre-requisite for value extraction though it is often a pre-requisite for value creation. Ownership is a means to create value, not necessarily a means of extracting value. For instance, when a scientist develops a patentable idea while working in a private R&D lab, the granted patent is owned by the employing company though the scientist is able to extract a part of the patent value both as tangible remuneration for her work either as salary and/or as a bonus and intangibles such as internal and external recognition awards. Similarly, a worker working in a factory does not own the product she produces. Nevertheless, part of the value generated by the factory's activity is extracted by the workers through the labour contract. The amount extracted by labour is determined (partly) by the labour laws and by competing employers in the market. In much the same way, it is not necessary to make Indian citizens own their data for them to be able to extract a part of the value their data generate. This is applicable to both consumer personal data and community data both of which may be residing with or in possession of profit-making companies, not-for-profit entities and even the government agencies.

2. With ownership comes control regarding how and where the data can be used to create value. For that, one needs specific expertise and specific knowledge and this is what businesses do. Businesses create value by organizing production and sales with the help of suppliers, workers and consumers. And most of these stake-holders extract value from businesses in various ways --- suppliers get paid, workers get wages and consumers get the product. A business, however, is owned by its investors or partners. Even when it comes to the owners of businesses, not all investors have control over the business assets. Granting ownership of data to

individuals destroys the innate ability of businesses to generate value in much the same way as granting decision making power to individual share-holders in a listed company will severely disrupt the process of value creation in businesses. Getting into a discussion on data ownership by individuals may seem conceptually esoteric and simplistic, it could actually turn out to be practically counterproductive.

3. Of course, because data is non-rivalrous (the same data can be used by more than one entity and that too at the same time), the correct discussion should be on how this asset is to be shared among all those who can generate value from it. For this, we need to ensure a mechanism such that anyone who can generate value from the data get access to it. Here it is important to keep two things in mind: (a) data capture and storage are costly and (b) it is datasets, rather than raw data, that generate value and this too is a costly process involving both upfront and continuous investments. Start-ups need to differentiate and succeed on merit and innovation rather than on a clutch of free data access to corpus of their predecessors and competitors; after all, the latter did invest to gather the NPD and may not invest further if they have to share it with others and serve the same literally on a platter. For example, while a vaccine manufacturer may have to provide clinical data to regulators, it would be unfair to mandate it to share such intellectual property with its existing or potential competitors. All the same, sharing of such data with each other can still happen under a commercial contract as long as it does not run afoul of the Competition Act. Thus, a mechanism that allows access to those not generating the data or the datasets, must ensure that those who are generating them have the right incentives to do so. It postulates that data sharing would lead to competition but provides hardly any empirical evidence supporting this theory. In fact, sharing certain type of information could also be violative of the Competition Act! There must be more space, nudge, flexibility and right incentives for voluntary data sharing.

4. Obviously, the foremost instance of such data sharing is what the policy draft refers to as "high value" which may be residing with non-government agencies but is essential for delivery of public service. Sometimes this data may need to be anonymized and, at other times, it need not. For instance, to determine whether there is the possibility of an outbreak (of a communicable disease) hospitals and healthcare providers need to give the total number of infections they have observed in a region within a pre-determined time period. This data could be anonymized. On the other hand, for many of the DBT payments, or the delivery of targeted and subsidized health-care services, the data need not be anonymized to ensure transparency and accountability. We then need specific laws/ regulations made for these special cases that make explicit not only how the data need to be transferred but also for the specific purpose for which this data can be used and how it is to be stored and secured by the receiving agency. Recall this is exactly

what we have done for disaster management or in the case of the Essential Commodities Act.

5. Sometimes sector regulators require data to understand how markets are functioning to determine whether, where and how an intervention, or action, is necessary. This requires legal backing to the regulatory authority to intervene whenever there is any violation of the safeguards the data holder is obliged to provide. Innovation requires experimentation; any attempt to force pre-determined outcomes destroys the very process of experimentation. The data economy in India needs to evolve and, if in the process it generates potentially undesirable outcomes, the sector regulator can step in to mitigate the same. Evolution is a trial-and-error mechanism, it needs to be overseen through regulation but not through the prism of any pre-determination of outcomes. This is especially true for the data economy where the ability to enforce regulation is a function of the technological developments and most of them are rapidly evolving and as yet unknown. Additionally, given the globalized nature of the digital economy and the underlying technologies, it is vital to consider the technological feasibility of enforcement of the policy without recourse to measures that overall may cause more harm than good.

6. In a market-based society, the government enables innovation by creating an ecosystem that encourages entrepreneurial activities. One important way to boost the digital economy is to enable businesses to share in the vast amount of data that the government collects from its citizens, per force and using taxpayer money. In fact, every government agency is obliged to provide free, accurate and timely data in machine-readable format under the Government of India's [National Data Sharing and Accessibility Policy, 2012](#). The government already has set up <https://data.gov.in> for this purpose.

7. It is disappointing that the report on NPD has nothing to say about this. Instead, it tries to design the data market-place with a number of opaque bodies and data trusts and an equally if not more unclear set of rules and regulations regarding how the data economy must evolve. This runs counter to the thinking behind the reforms started in 1991 aimed at removing excessive controls that were the bane of the command economy.

8. In addition, the very concept of 'community data' is extremely problematic. One could simultaneously belong to multiple communities at the same time in the same context; e.g., one could belong to a religion, region, linguistic group, an alumni group and a member of a Resident Welfare Association. There is no clarity on which of these communities would have the right to determine whether and how the NPD could be used or shared. Also, there is neither a practical and expeditious mechanism or forum for dispute or conflict resolution across such

collectives nor one to deal with the scenario of an individual surrendering or exiting from membership of a particular community. In addition, there could be potential conflicts between the NPD custody and other functions of a community. A 'community' may even be temporary or dynamic in its nature or composition, for example, people in a festival fair or attending a rally or co-passengers in a bus or train. Accordingly, unless there is legal certainty on specific 'community' and unless that collective has exclusive rights to NPD, it is extremely difficult to implement the proposition of community data rights, fraught with arbitrariness and dynamic discretion.

9. It is important to have a national strategy for the data economy. The Chinese approach has been to isolate its national data and, thereby, capture value. The European approach has been focussed on data privacy with punitive measures for data breaches. The US approach, since most of the technology giants are based there, has been to create a system that maximizes value for the (data) technology companies. India should also develop its own strategy based on what it hopes to achieve through digitalization.

10. A clear focus on the objective is more important than the focus on pre-determined processes. For example, suppose we decide that retaining maximum value for its citizens is India's objective from digitalization. Also, suppose that there could be two distinct ways we can get value creation from our data. One is to support a regulatory mechanism *A*, the other is to support regulatory mechanism *B*. Suppose *A* produces a total value of 150, while *B* produces a total value of 100 only. Before we decide on the process, we must consider our objective of maximizing the value for Indians. If India gets 30 from process *A* and 50 from process *B*, then India should opt for *B* and not *A*. For the US, given that it houses the technology companies, it gets the most of process *A*; China, on the other hand, opted for something like process *B*. Europe has worked out a compromise by which it allows process *A* but keeps chipping away at the surplus generated in that process through damages for (European) consumer harm or through taxation.

11. Accordingly, while designing our non-personal data regulations, we firstly need to articulate specific policy objectives, and secondly, understand our context and then only endeavour to design the policy ecosystem and the regulatory apparatus that generates the process that best meets our objective. In other words, it is the responsibility of the policymaker to show explicitly how the policy design being proposed will enable us to meet our objective. There are no "best practices" from other countries to follow, only some pointers, at best.