Inputs to MeitY on the Draft India Data Accessibility and Use Policy

Submission from IT for Change

March 2022

Inputs from IT for Change[1] to MeitY on the

Draft India Data Accessibility and Use Policy

Dated 18th March, 2022

With our societies getting increasingly digitalized and datafied, if governments are to keep fulfilling their central societal role, they too must get digitalized and datafied. But this must be undertaken in appropriate ways that protect and enhance people's rights, requiring strong new institutional checks and balances on one hand, and new enabling roles for the state on the other. Both these imperatives — undertaking digitalization and datafication, and basing these processes in fundamental rights — are essential for upholding our 'social contract'.

In this context, a new key role of the state is to help make available as a public infrastructure — all data that is vitally required for various social and economic activities in a digital age. This serves the socio-economic objectives of both productivity and equity. We find the draft Data Accessibility and Use Policy an extremely important and timely initiative in this regard, and congratulate all those involved. We agree with most basic premises of the draft policy. However, as a contribution towards making it even better, and placing it firmly in a forward-looking digital society context, we offer some inputs below. These consist of first laying out a few larger framework level issues, and then presenting five specific actionable principles for MeitY's[2] consideration.

I. Some Larger Framework Issues

1. Background: From RTI to 'Open Data'

The first framework for sharing of government information/data was in terms of a Right to Information (RTI). Aimed at ensuring governmental accountability, it has largely been a pull mechanism where specific information is sought and provided (although push mechanisms like Section 4 of the RTI Act 2002 also exist). As technological advances made possible widespread publishing of information and data, the 'Open

---

[1] IT for Change is an Indian NGO with a global footprint, and in Special Consultative Status with the United Nations ECOSOC. Please visit www.itforchange.net for further information. Please write to ITfC@itforchange.net if you need any clarification or further information.

[2] Ministry of Electronics and Information Technology, Government of India.

Government Data' movement came to the fore. With it, the 'resource nature' of information/data became as important as its 'accountability' aspect. Taking cues from the open source software movement, open data proponents sought open publication of most government data without any constraints regarding the use or the user.[3] Moreover, there were to be (mostly) no special enablements for any specific uses or users. (1) Open, unconstrained publishing, and (2) 'neutrality', or rather indifference, with respect to use/users, have been two key features of 'open data'.

2. The Nature of Data Mutates in a Digital Society

With the emergence of a digital society and digital economy, data (and data-driven 'digital intelligence'[4]) is increasingly becoming 'the' central resource in today's world, and not just an ancillary one as in the previous 'open data' framework. Data, today, largely arises from automated processes, often without predetermined purposes, continually, and in almost incomprehensible quantities. Further, a lot of data may be constrained with regard to 'open sharing' because of reasons of security, privacy (individual and collective), and various economic, social, and cultural rights related to data.[5]

Such features of data in a digital society/economy context completely transform what it means for governments to share their data. An appropriate new conceptual framework is therefore required, going beyond a 'right to information' approach or an 'open data' framework, while subsuming the values, objectives, and requirements of both.

3. 'Data as Infrastructure', Like Roads and Electricity

In a digital society, data is not just an ancillary resource; it has become like an infrastructure that serves as an indispensable common basis for nearly all digital activities. Sharing of data can be considered similar to governments constructing roads or providing electricity to citizens. If the latter were industrial era infrastructures, data is the key digital age infrastructure.[6] As a shared resource, roads are necessary for

---

[3] https://okfn.org/opendata/
[4] UNCTAD's Digital Economy Report, 2019, https://unctad.org/system/files/official-document/der2019_en.pdf
[5] Normally only IP rights are mentioned in relation to economic rights over data. However many other kinds of individual and collective economic rights to data exist, for instance as provided in some recent EU legislation like the draft Digital Markets Act, and the draft Data Act, and envisaged in the report of India's Committee of Experts on Non Personal Data Governance Framework. Such rights also exist in legacy frameworks like for traditional knowledge, UN Convention on Bio Diversity, etc.

almost all aspects and activities of the society and economy. Governments cannot simply notify a strip of land as open for anyone to ply transport over (which would be akin to the 'open data' approach). It has to make proactive investments and arrangements to ensure that private transport can actually ply, and ply meaningfully and appropriately, starting from laying asphalt right up to developing convenience stops. In the larger public interest, it may also be needed to bar certain kinds of vehicles/activities on a road, at all or during some specific times. In a similar way, safe and effective sharing of digital age data very often requires the need to both (1) enforce some constraints, and (2) provide specific enablements. We therefore propose that any data sharing and use policy in the digital age should internalize such a 'data as infrastructure' or 'data as utility' framework. (We prefer the name 'data sharing and use policy' over 'data accessibility and use policy'.)

Sharing of government data in a digital society context needs to be explicitly recognized as going qualitatively (and not just quantitatively) beyond an 'open data' approach, as also the earlier Right to Information framework. Data sharing needs to be seen as an indispensable society wide infrastructure ('data as infrastructure'). However, it should simultaneously (1) maximize the possibilities of information/data enabled accountability of governments ('Right to Information'), and (2) provide as much informational and data resources as possible in an open and unconstrained manner (open data).

4. The Three Contexts of Government Data Sharing

With immeasurable amounts of data getting produced, governments need to make three assessments about what data to proactively share and how: (1) whether and how certain data enhances its accountability to and engagement with citizens, (2) which data can serve as a useful 'open resource', and (3) what kinds of data need to be shared in a manner that involves enforcing some constrains, as also providing special enablements ('data as infrastructure' framework).

The EU covers (2) above under its Open Data Directive and (3) under the new draft Data Governance Act. In India, (1) is covered by the RTI Act and (2) under the National Data Sharing and Accessibility Policy, 2012. The current policy seems addressed mostly to (3) while also mentioning (2). Absence of an RTI imperative, however, is a glaring gap. RTI data may be 'open data' in content but not necessarily in objective

---

[6] This proposition gets often countered by citing the rivalrous versus non-rivalrous nature of different resources. We are unable to to go into detail here, but may just state that the issue is much more complex than admitting of such a simplistic dualism.

('accountability measure' versus 'resource'). As new technical and institutional means for data sharing are developed that are suited to a digital society context, the 'Right to Information' objective and need must also be explicitly foregrounded. This would help government agencies make the right choices about which data to make available and how.

5. A State Fit for the Data Age: Maximizing Citizen Welfare

A key feature of the draft policy is intra-governmental data sharing. This becomes very important as we move towards 'platformization of governance' as implicitly conceptualized in GoI's consultation paper on the NODE (National Open Digital Ecosystems) architecture as the future of governance. As is increasingly true for all institutions and organizations, if governments do not datafy and platformize appropriately, their legitimate role and power will wither, quickly and irretrievably. By default, most social, economic and political power in a digital society will then be left in the hands of a few global digital corporations. We are already seeing such a trend shaping up all around us. The Government of India is taking the right steps towards becoming the central orchestrator of digitally platformized sectors, building and maintaining their core protocols and platforms, such as with UPI in finance and IDEA in agriculture. The vertical top-down regulatory role of the industrial era state needs to be supplemented by a center-to-peripheries platform-ochestrator role in a digital era state. (Partly in a 'code is law, architecture is policy' mode, but backed by law.) As with its current regulatory role, the new digital era role needs also to be institutionalized in appropriate laws, incorporating all the necessary checks and balances arising from the fundamental rights and the directive principles under the Constitution.

Apart from intra-governmental sharing, such 'platformization of governance' would obviously also centrally involve extensive data sharing with actors outside the government, in both directions. This draft policy will take care of inter-governmental data sharing and sharing of government data with outside actors. The other very important element of getting the main holders of society's data today — the big digital platform corporations — to share data back into the national data commons is being dealt with by the GoI's Committee on Non-Personal Data Governance Framework. Its draft report is hereafter referred to as the NPD Report.

6. The Need for Strong Institutional Checks and Balances

This proposed policy is a very significant step in providing a technical and institutional framework for sharing of government data. However, huge (digital age level) collection and holding of data by governments, and its seamless sharing across all government organs, without corresponding development of institutional checks and balances, raises deep concerns about a very likely unaccountable concentration of data power with the state. New laws, and if needed new constitutional provisions, are urgently needed as institutional checks and balances to complete what would be an 'appropriate datafication' of the state towards fulfilling its legitimate objectives. Substantial and rapid progress on that side is much needed today (on which subject we have separately been making contributions). In their absence, state's datafication will keep encountering great resistance, which will be quite legitimate. This would not allow the sustainability of various innovations and transformations that the government, especially MeitY, is undertaking towards rendering our state machinery fit for the digital age.

II. Five Specific Inputs for the Draft Policy

Against the above background, we have the following specific points to make as inputs to the draft policy. These are in the form of five additional principles that may be explicitly adopted by the policy.

1. Rules-based Non-discrimination, and Non-Exclusivity

When the central resource of a digital society is in question, especially in the context of the government's role as the key orchestrator of digitally platformizing sectors, it becomes extraordinarily important that government data is only shared in a fully rules-based non-discriminatory manner. A public resource like government data must only be shared in a rules-based manner, that is transparent, accountable, and non-discriminatory. There can be no exclusivity. However, if special circumstances (which need to be clearly stated and justified) indeed call for an exclusive arrangement, it should be done transparently in a rules-based manner, and regularly revisited, with avenues available for appeals against it. This has always been

important, but becomes even more crucial when data is now a necessary infrastructure, and the government a key provider of such infrastructure.

While the draft policy does mention non-discrimination, it needs to be made much more clear, specific, and binding, with appropriate rules duly framed and publicized. The policy can adopt and apply the provisions of Article 4 of EU's draft Digital Governance Act[7], with some adaptations if necessary.

Many well-meaning and much-required efforts for developing a due public sector role in platformization of different sectors have been marred by successive governments' failure to apply appropriate arrangements for transparency, non-discrimination and non-exclusivity. For instance, the Swasth App (with its Aarogya Setu linkage), and the agriculture ministry entering into exclusive MoUs with various actors (including with global digital corporations like Microsoft) which *inter alia* involve data sharing, have raised major legitimate concerns. If the government is to successfully conduct their digital transformation exercise, it must adopt rules-based, transparent, and accountable approaches of non-discrimination and non-exclusivity. These must be enshrined in law (and as a beginning, in binding administrative procedures), as is done in the EU's draft Data Governance Act. Adequate avenues of independent third party redress, including judicial, should be available to all aggrieved parties.

In special circumstances where some early, and 'strictly temporary', experiments may need to be conducted in partnership with specific parties with a view to develop proof-of-concept etc., it must also be rules-based (EU's draft Data Governance Act enshrines in law the need to follow clear procedures in such cases). Sandbox provisions can be created for experimental phases by employing open and transparent administrative procedures and laws. In default, what happens is that such 'experiments' tend to become permanent, with some privileged private parties shaping the design and form of the new 'digital system' (based on what could be their narrow interests) and entrenching themselves as digital leaders in a sector employing the patronage of government agencies. This has been a major and valid criticism of many digital transformation efforts undertaken by various governments. It must be addressed urgently by developing elaborate rules and corresponding institutional mechanisms.

---

[7] https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020PC0767&from=EN

2. A Positive, Enablement Approach, Which can be Use and User-Specific

The draft policy has taken the right step by including in its title 'use' of data, which is a progress over the earlier 'open data' thinking of the National Data Sharing and Accessibility Policy, 2012. How data can actually be used for appropriate purposes, and by all the intended users, becomes very important for data sharing in a digital society context. The draft policy has rightly created a new institutional forum of the India Data Office as a specific anchor point and mechanism for sharing government data. Equally worthy of appreciation is the accompanying governance framework of the India Data Council.

Proposing a 'data as infrastructure' approach, we discussed how data is becoming key to society-wide activities, and data sharing and use very diverse and often complex in nature. In the circumstances, it is not at all adequate to simply publish data and let people and organizations do what they may with it. As is with other vital infrastructures of society, shared data also needs to be proactively made useful and usable for different kinds of uses as well as users.

Almost every activity in a digital society could be generating digital data. It is not possible to collect, retain and share all such data. An active choice element is therefore involved at all steps. If data minimization is a principle for governance of personal data, in case of non-personal data, too, it may be needed to extinguish and not retain much of the data that gets created. And from the retained data, careful, use and benefit-based selections will need to be made of data that is the most important to be provided as infrastructural facility. The government data-related provisions of EU's draft Data Governance Act speaks of "a register of available data resources containing relevant information describing the nature of available data".

This may be laboring on an obvious point, but some choices will be necessary to make about what data is considered as more important to be shared. Specific investments will be required not just to make such data available, but also meaningfully usable. Often, this may involve employing some degree of judgment about the kinds of intended uses and users. The concept of High Value Datasets (also employed in some EU legislation, and in India's NPD Report) was seen in an earlier version of this draft policy but has since disappeared. We recommend that the concept of High Value Datasets be restored, and guidelines provided on how to strongly invest in specific processes to make such data actually usable for important social and economic uses and users.

We have mentioned that, as governments get datafied, the India Data Office will emerge as a new institutional forum. (It will be a good idea to also have a single point for data requests for every ministry/department/agency, as provided for in the EU's Data Governance Act.) With such a forward-looking understanding, governments will need to invest in building new capacities and new mechanisms, oriented towards sharing data in ways that it can be, and is actually, used for a variety of social and economic purposes. Some data needs may be in real-time, which has its own challenges, requiring complex mechanisms and processes.

Governments the world over, especially in the EU, are undertaking many projects for appropriate data sharing, which are oriented to specific purposes and users (like the GAIA-X project of the EU). Such projects may be cross-sectoral or sector-specific, with some sectors like transport, health and agriculture being early starters. India, too, needs more elaborate guidelines, rules, plans, mechanisms and projects for actual data sharing to happen in ways that are useful to the society and economy in transformational ways. We see the current draft being perhaps too bare-boned and lacking in this regard.

3. Positive Discrimination Towards Goals of Equity

With governments handling some of the most important data — the key resource in a digital society — and also (rightly) aspiring to be the main orchestrator of various digitally-platformizing sectors, it no longer remains a matter of merely making data resources available to everyone, in an open, equal, and indiscriminate manner. The distributional aspects of the state's objectives and policies begin to get implicated in its new digital and data roles. It may be required to go beyond non-discrimination, to undertake protective discrimination in favor of weaker sections, or those otherwise preferred (like domestic industry over foreign industry).

Such protective discrimination is already emerging as a norm in data sharing laws and policies. Both EU's draft Digital Market Act and draft Data Governance Act contain data sharing obligations and data access privileges that operate in such a protective-discriminatory manner. They are asymmetrical between less powerful actors and big digital platforms. India's NPD Report also provides for such asymmetry and protective discrimination in data sharing and data access provisions. (We must remember that such things

have been antithetical to 'open data' thinking.) MeitY's NODE framework also cautions against an unfair advantage that some may obtain from data sharing mechanisms, and the need for protections against it.

Such concepts and provisions are well-established in theory and practice of governing commons. The most well-known framework for governing commons by Elinor Ostrom considers explicit provisions about 'inclusion and exclusion' as key to successfully governing any commons.[8]

It is already well-known that more powerful players (read Big Tech) can partake from a (data) commons in a manner that further skews power equations in their favor, to the detriment of all others. (Data is really not a fully non-rivalrous resource in that sense; its use or over-use by some can be at the cost of others.) Policymakers the world over are very concerned about such possibilities, and are seeking to devise specific checks against them. For a deeply unequal country like India, where, additionally, most key platform companies are all foreign, this principle of protective discrimination in data sharing policies becomes even more salient. We can separately contribute to how this principle can further be developed, instituted, and implemented.

4. Localization of Government Shared Non-Personal Data

EU's draft Data Governance Act is explicit in that at least certain categories of government shared non-personal data may not be sent outside the region, or, in other words, be localized. For other kinds of government shared non-personal data as well, various conditions of 'legal regime adequacy', applicability of jurisdiction, etc., have been placed. India must also adopt similar cautions and restrictions, as necessary.

Indian data going abroad is not only an issue vis-a-vis security and privacy but also a matter of enforcement of various kinds of economic rights. For instance, India, along with other developing countries, is already contending at global forums that gene sequencing information related to its flora and fauna is something that it has sovereign rights over (under the Nagoya Protocol of the UN Convention on Bio Diversity). India's NPD Report proposes collective community rights to its data, including at the national level. EU's draft Digital Markets Act and draft Data Act give various kind of economic rights to data-generating actors as well.

Any policy on sharing non-personal data in the hands of the government should explicitly recognize these important issues and concerns, and provide appropriate guidelines regarding them.

---

[8] https://www.onthecommons.org/magazine/elinor-ostroms-8-principles-managing-commmons

5. Trusteeship Framework as Against Ownership/Property Framework

The draft policy mentions that the department that generates or collects data will be the 'owner' of the data, and that data will remain its 'property'. We understand that these terms 'owner' and 'property' — may have been employed here loosely, similar to how corporations internally designate 'data owners' as those in control of, and responsible for, certain data sets. However, extreme caution must be exercised in adopting such terms in policy documents, when the political economy of data (questions of who owns data etc.) is one of the mostly hotly contested issues globally. The proposition that someone owns data by the virtue of being its collector could be extended to imply that big digital corporations own all the data that they collect and hoard. This is not at all the accepted legal position anywhere in the world. In fact, the default regime of 'data collector as owner' is being strongly pushed back everywhere, including in the EU and in India (through the NPD Report).

The draft policy should therefore avoid using terms like 'owner' and 'property' with regard to departments and agencies that collect or even generate data.

There is a long tradition in law and practice for governments to be the trustees of a society's common resources, or 'resource commons'. Such frameworks exist in India, as elsewhere, with regard to commons like water bodies, forests, minerals, traditional knowledge, etc. Society's data in government's hands is also a kind of 'resource commons'. Governments should adopt a 'trustee' role and designation with respect to all such data, taking from similar roles and conceptual/legal frameworks in many other areas. This should be explicitly recognized in this policy and its implementation mechanisms. The concept and practice of 'date trustee' is actively being developed in many parts of the world, as also in India's NDP Report. This policy should also adopt the 'data trustee' role for government agencies that handle society's data.