The Dialogue™
INFORM ENGAGE IDEATE

Comments

OUR RESPONSE TO DRAFT INDIA DATA ACCESSIBILITY AND USE POLICY

Authors: Kamesh Shekar & Karthik Venkatesh

ABOUT THE DIALOGUE

The Dialogue is a public-policy think-tank with a vision to drive a progressive narrative in India's policy discourse. Founded in 2017, we believe in facilitating well-researched policy debates at various levels to help develop a more informed citizenry, on areas around technology and development issues. The Dialogue was ranked amongst the Top-Ten think-tanks in the world to watch out for by the Think-Tank Civil Societies Programme, Lauder Institute, University of Pennsylvania, in their 2020 and 2021 ranking index.

✉ info@thedialogue.co

🌐 www.thedialogue.co

SUMMARY

It is a good initiative by the government to acknowledge the value of data that is collected by government departments and to make it accessible for use for the larger ecosystem has the potential of driving economic and social benefits. This will help startups, the research community, and enterprises unlock the value of data. We welcome that the government has invited comments on the draft policy. Below is our response and recommendations to some of the provisions in the policy.

Provisions of the policy	Recommendations
Principles	
The principles on which data sharing and governance need to be based	The below principles and guidelines must be incorporated within the policy. These principles must be held and promoted by the data council. **Purpose specifications:** The purpose of data collection must be specified within the request. The data collected must be used only for the stipulated purpose, nothing incompatible with the specified purpose. Besides, in case of a change in purpose, the respective government department must be notified. **Fair and lawful processing:** The data processing must be consumer-centric. The results or actions of processing should not translate into exclusion and unfair discrimination directly or as one of the unintended consequences. **Data deletion:** In addition to data retention, the policy must also provide guidance on data deletion. The data shared with the government or other stakeholders must be deleted when the retention policy expires.
Institutional Framework	
6.1. India Data Office (IDO) shall be set up by MeitY with an objective to streamline and consolidate data access and sharing of non personal data repositories across the government and other stakeholders.	Ensure that Data Protection Laws and surveillance reforms are in place prior to adopting this policy. The data access and stipulated purpose check mechanism must be differentiated between the government and other stakeholders. Accordingly, the policy must institute mechanisms, rules and use cases to weed out misuse by various stakeholders.

6.3. India Data Council – comprising India Data Officer and Chief Data Officers of Departments of Government of India and State Governments – shall be constituted with an objective of undertaking tasks that requires deliberations across Ministries, Departments and State Governments. This shall include but not limited to defining frameworks for defining important Datasets, finalizing Data standards and Metadata standards as also reviewing the implementation of the policy. The nomination of Departments and state governments in India Data Council shall be on rotation with a tenure of 2 years for one Department.	The council must be notified through statute after deliberated by the parliament. The council must work closely with the Data Protection Authority (as provisions on non-personal data are expanded in DPB) and evolve its policies and SOPs in congruence with the regulations surrounding data. An expert committee on non-personal Data (NPD) governance also suggests forming a Non-personal Data Protection Authority (NDPA). Thus council must work in tandem with NDPA (if instituted) as well. Besides, other India Data Council should work in synergy with other sectoral regulators as well. As technological developments are fast-paced, it is important to incorporate technical experts within the council as advisors or part-time members of the council.
6.6. Through the IDO, stakeholders including researchers, start-ups, enterprises, individuals and government departments will be able to access non-personal data through mechanisms for data licensing, sharing, and valuation within the overall framework of data security and privacy.	The datasets must be priced according to the contextual and subjective value. The policy must clarify what is considered as "high value" data sets. It is not sufficient to have broad terms such as "based on importance in the market", "socio-economic benefits", etc., which provide for wide powers to the incoming institutional body (Data council) to decide on these things. The guard rails for checks and balances in exercising these powers need to be established, and it's important we decide on the acceptable framework to view/gauge value. Datasets that are to be shared freely when they are "minimally processed". The value addition through labour or transformation is to be priced appropriately.
Government-to-Government Data Sharing	
8.2. Approved inventories will be federated into a government-wide searchable database for government-to-government data sharing. This will minimize duplication of data processing efforts and enable better delivery of citizen centric services.	The policy must bring harmonisation by weeding out the overlapping and conflicting scopes and bringing them to congruence with other existing policies. High-level coordination between the data council and sectoral regulators be instituted to enforce data sharing policies in a coordinated way.

Data Quality & Meta-Data Standards	
11.1 Each Central Ministry / Department shall adopt and publish its domain-specific metadata and data standards. These standards should be compliant with the interoperability framework, policy on open standards, Institutional Mechanism for Formulation of Domain-specific Metadata and Data Standards and other relevant guidelines published on the e-govstandards portal. 11.2 Data standards that cut across domains shall be finalised by India Data Council and once finalised adopted by all concerned government ministries/ departments.	It is also essential to have a mechanism to cross-check and evaluate the integrity, cleanliness of the data, as state and non-state actors would use this for real-life interventions. For instance, mechanising periodic audits for both data collection methods and data could help in cross-Checking. Besides, compar-ing the data with an alternative database can also help determine gaps and mistakes in data points within the coordinated dataset.
Data Sharing Toolkit	
15.2 Data shall remain the property of the agen-cy/department/ ministry/ entity which generated/ collected it. Access to data under this policy shall be strictly in accordance with any act and rules of the government of India in force. Legal framework of this policy shall be aligned with various acts and rules covering the data.	Rights like the right to data correction, the right to be forgotten etc., must be vested in the hands of the individuals. The government departments and ministries must take consent from the individuals for sharing their non-personal data with other government depart-ments and other stakeholders. While various government departments have dif-ferent data correction and update mechanisms, the data council must have a single overarching mech-anism both analogously and digitally to correct and update information. Establishing a single point of contact for correction and updates across registries will make it easier for individuals to navigate the sys-tem. In many cases, navigating the grievance manage-ment system for both individuals and entities is ar-duous, making it difficult to reach the designated portal. Therefore, we suggest incorporating an In-teractive Voice Response (IVR) wing as part of the data council. This automated voice response sys-tem (through call) should navigate individuals (or en-tities) to reach the appropriate grievance portal, i.e., service providers' grievance redressal mechanism or dispute resolution wing the data council.

1. LICENSING AND PRICING

Licensing and pricing of the data and value addition through labour are to be priced appropriately. Pricing of data sets and the value of a particular dataset will depend on the context and umbrella pricing; like in previous cases,[1] will be sub-optimal. Value of data comes from relevance, quality, granularity, and moreover, the linkage to the use case. We need to create strong processes to minimise hurdles in standardising prices.

It is important to understand that datasets can not be priced uniformly, and the value of a particular dataset varies depending on the context in which it is solved. For example, health data from healthcare providers within the government setting is valued differently by a public health research institution and by a pharmaceutical company. Moreover, we need to create streamlined processes by different government departments at all levels to minimise hurdles in the standardisation of prices.

Recommendations

- The datasets must be priced according to the contextual and subjective value.

- Linked to the point of licensing and assessing value, the policy must clarify what is considered as "high value" data sets. It is not sufficient to have broad terms such as "based on importance in the market", "socio-economic benefits", etc., which provide for wide powers to the incoming institutional body (Data council) to decide on these things.

- The guard rails for checks and balances in exercising these powers need to be established, and it's important we decide on the acceptable framework to view/gauge value.

- Datasets that are to be shared freely when they are "minimally processed". The value addition through labour or transformation is to be priced appropriately.

2. NEED FOR A ROBUST INSTITUTIONAL FRAMEWORK

The policy mulls for creating an Indian Data Council that comprises of Indian Data Officer, Chief Data Officers of GoI and various states. The council will then be tasked with defining high-value datasets, facilitating sharing, and behaving as an institutional body that fosters data sharing and creates policies. While this is a welcome move, two points are to be noted here.

- Creating a regulator that is not defined by statute and is not subject to parliamentary/ judicial oversight mechanisms could create problems for how India's data protection landscape is viewed, which in turn can affect India's bilateral data-sharing agreements, adequacy findings etc.

- There are various upcoming data protection regulators and sectoral regulators who govern and regulate data in India. As Indian Data Council will be providing data sharing policies, it might clash with the jurisprudence of the other upcoming and existing regulators.

- The composition of the council is executive driven; there is no provision of incorporating technical experts as part of the council.

[1]*Bulk Data Sharing Policy March 2019.* (n.d.). Parivahan. Retrieved March 17, 2022, from https://parivahan.gov.in/parivahan/sites/default/files/NOTIFICATION%26ADVISORY/8March%202019.pdf

---

<div style="border:1px solid; display:inline-block; padding:4px;">**Recommendations**</div>

- The council must be notified through statute after deliberated by the parliament.

- The council must work closely with the Data Protection Authority (as provisions on non-personal data are expanded in DPB) and evolve its policies and SOPs in congruence with the regulations surrounding data.

- An expert committee on non-personal Data (NPD) governance also suggests forming a Non-personal Data Protection Authority (NDPA). Thus council must work in tandem with NDPA (if instituted) as well. Besides, other India Data Council should work in synergy with other sectoral regulators as well.

- As technological developments are fast-paced, it is important to incorporate technical experts within the council as advisors or part-time members of the council.

3. DATA SHARING: A NUANCED APPROACH

While this is a great move, consolidating databases and creating access for the government to easily search and use data is a slippery slope in the absence of strict mechanisms of transparency and accountability. Now, it remains to be seen how access to such pan-government databases will be limited and how it will interact with the protections under *Puttaswamy judgement* I[2] and the proposed government exemptions clauses.

Additionally, as MeitY is yet to finalise the data protection regime for the country to secure informational privacy, G2G should acknowledge that the interlinked non-personal data can potentially create an exploitative relationship. For instance, interlinkage of non-personal data (anonymised) set can reveal the personal identity of the individuals due to triangulation[3] and also cause collective privacy harms.[4] Besides, this confederation of non-personal databases can lead to unintended consequences of surveillance by whoever has access to the system.

While the policy talks about streamlining and consolidating data access and sharing with other stakeholders (not only governments and departments), there is no adequate safeguard in terms of use cases, who can access data and the stipulated purpose for other stakeholders.

Therefore, this shows that the currently envisioned consolidation of data access might serve to cause unintended consequences like data privacy and security breach, and increase surveillance if appropriate mechanisms to prevent the same are not established before the start of this project. The policy should ensure that these exploitative relationships are weeded by imposing optimal liability on the stakeholders, including the government.

<div style="border:1px solid; display:inline-block; padding:4px;">**Recommendations**</div>

- Ensure that Data Protection Laws and surveillance reforms are in place prior to adopting this policy.

---

[2]Puttaswamy Jusgement I, (2017) 10 SCC 1
[3]*MoRTH scraps its Bulk Data Sharing Policy but THIS IS NOT IT!* (2020, June 30). Internet Freedom Foundation. Retrieved March 17, 2022, from https://internetfreedom.in/morth-bulk-data-sharing-policy-scrapped/
[4]Joshi, D. (2020, July 30). *Non-Personal Data Regulation: Interrogating 'Group Privacy'*. Centre for Law & Policy Research. Retrieved March 17, 2022, from https://clpr.org.in/blog/non-personal-data-regulation-interrogating-group-privacy/

- The data access and stipulated purpose check mechanism must be differentiated between the government and other stakeholders. Accordingly, the policy must institute mechanisms, rules and use cases to weed out misuse by various stakeholders.

4. HARMONISATION OF DATA SHARING POLICIES AND SECTORAL REGULATIONS

We caution that the existence of various sectoral level data-sharing policies and regulatory frameworks could cause complexity in the implementation of this policy. Some of the existing key data sharing policies are - **Draft Non-Personal Data (NPD) Governance Framework**[5] (enables data sharing to unlock the economic benefits from NPD), **Geospatial datasets liberalisation**[6] [democratised the existing geospatial datasets (including government)], **Open Data initiative of the Government of India**[7] (provides access to Government-owned shareable data and its usage information in open/machine-readable format), **India Digital Ecosystem of Agriculture (IDEA) framework**[8] (links various publicly available data from various schemes to digitalised land records). We also want to mark that this would cause confusion with the InDEA 2.0 (especially point 8.2) envisioned by MeiTy.

Besides, the multiplicity of regulations and rules on the same functions i.e., data sharing causes various supply-side, demand-side issues and regulatory issues. For instance, the multiplicity of laws can cause regulatory arbitrage. In addition to this, a lack of uniformity would lead to contradictory interpretations in case of a dispute.

> **Recommendations**

- The framework must bring harmonisation by weeding out the overlapping and conflicting scopes and bringing them to congruence with other existing policies.

- High-level coordination between the data council and sectoral regulators be instituted to enforce data sharing policies in a coordinated way.

5. QUALITY OF DATA

While the policy talks about data standards, as we move towards opening the data, it is also essential that India Data Council notify standards that would ensure the integrity and cleanliness of the data. This becomes a crucial aspect as state and non state actors will be using the data for real-life interventions and applications. Besides, it is also essential that the India Data Council develop mechanisms through which government departments can test the data and periodically audit the same for quality in terms of integrity and cleanliness.

---

[5]*Report by the Committee of Experts on Non-Personal Data Governance Framework*. (2020, December 24). MeitY. Retrieved March 8, 2022, from https://ourgovdotin.files.wordpress.com/2020/12/revised-report-kris-gopalakrishnan-committee-report-on-non-personal-data-governance-framework.pdf

[6]*Guidelines for acquiring and producing Geospatial Data and Geospatial Data Services including Maps* (2021, February 15). Department Of Science & Technology. Retrieved March 8, 2022, from https://dst.gov.in/sites/default/files/Final%20Approved%20Guidelines%20on%20Geospatial%20Data.pdf

[7]*Open Government Data (OGD) Platform India – An Overview.* (n.d.). MeitY. Retrieved March 8, 2022, from https://www.meity.gov.in/writereaddata/files/OGD_Overview%20v_2.pdf

[8]*Consultation Paper on IDEA.* (2021, June 1). Department of Agriculture, Cooperation & Farmer Welfare Government of India 1st June 2021. Retrieved March 8, 2022, from https://agricoop.nic.in/sites/default/files/IDEA%20Concept%20Paper_mod31052021_2.pdf

Recommendations

- It is also essential to have a mechanism to cross-check and evaluate the integrity, cleanliness of the data, as state and non-state actors would use this for real-life interventions. For instance, mechanising periodic audits for both data collection methods and data could help in cross-checking. Besides, comparing the data with an alternative database can also help determine gaps and mistakes in data points within the coordinated dataset.

6. DATA PRINCIPLES

While the policy adopts privacy and security by design as a principle in addition to data retention guidelines, Some key principles with respect to technology and data protection need to be identified and recognised. The policy must flush out some of the key principles discussed below in the recommendations.

Recommendations

The below principles and guidelines must be incorporated within the policy. These principles must be held and promoted by the data council.

- **Purpose specifications:** The purpose of data collection must be specified within the request. The data collected must be used only for the stipulated purpose, nothing incompatible with the specified purpose. Besides, in case of a change in purpose, the respective government department must be notified.

- **Fair and lawful processing:** The data processing must be consumer-centric. The results or actions of processing should not translate into exclusion and unfair discrimination directly or as one of the unintended consequences.

- **Data deletion:** In addition to data retention, the policy must also provide guidance on data deletion. The data shared with the government or other stakeholders must be deleted when the retention policy expires.

7. CITIZEN'S OWNERSHIP OF DATA

The policy makes the public non-personal data the property of the respective government department or ministries. This is problematic as the data generated, collected and archived by the ministries, departments etc., is of the citizens. Therefore, the ownership must be transferred to the citizens, and the below recommendations must be considered accordingly.

Recommendations

- Rights like the right to data correction, the right to be forgotten etc., must be vested in the hands of the individuals.

- The government departments and ministries must take consent from the individuals for sharing their non-personal data with other government departments and other stakeholders.

- While various government departments have different data correction and update mechanisms, the data council must have a single overarching mechanism both analogously and digitally to correct and update information. Establishing a single point of contact for correction and updates across registries will make it easier for individuals to navigate the system.

- In many cases, navigating the grievance management system for both individuals and entities is arduous, making it difficult to reach the designated portal. Therefore, we suggest incorporating an Interactive Voice Response (IVR) wing as part of the data council. This automated voice response system (through call) should navigate individuals (or entities) to reach the appropriate grievance portal, i.e., service providers' grievance redressal mechanism or dispute resolution wing the data council.

The Dialogue™

INFORM ENGAGE IDEATE

The Dialogue™ is a public-policy think-tank with a vision to drive a progressive narrative in India's policy discourse. Founded in 2017, we believe in facilitating well-researched policy debates at various levels to help develop a more informed citizenry, on areas around technology and devel-opment issues. The Dialogue™ has been ranked as the world's Top 10 think-tanks to watch out for, by the Think-Tank and Civil Societies Programme (TTCSP), University of Pennsylvania in their 2020 and 2021 rankings.

https://thedialogue.co