

Artha Global's response to the proposed National Data Governance Framework Policy, May 2022

Overview

Artha Global appreciates the opportunity to submit feedback on the National Data Governance Framework Policy, May 2022. Based on our previous published research and experience in supporting government policy and implementation initiatives at both the central and state levels, we are pleased to share our recommendations on the proposed policy.

About Us

Artha Global is a **research, consulting and network facilitating organisation** that assists governments in the developing world to design, implement and institutionalise policy frameworks that promote freedom, prosperity, and stability for citizens. We provide actionable **research** and support for policy **implementation** and **institutionalisation**. We also create **convening platforms** to help conceptualise new development agendas, build consensus and cross-sectoral coalitions amongst our partners in government, business, academia, philanthropy and civil society.

Our work focuses on the structural transitions that developing countries must undergo in order to achieve sustainable and inclusive development. These economic, spatial, social and technological transitions can be broadly summarised as the move from *rural to urban*, *informal to formal* and *analogue to digital*, all taking place against the backdrop of imperatives such as the need to address climate change, gender equity, data privacy and future shocks like the pandemic. **We help governments develop the requisite capacity to manage the transitions, and the inevitable dislocations they cause, in order to secure long-term prosperity and social stability for their citizens.**

Context

*“Encourage data access and sharing arrangements that ensure that data are **as open as possible** to maximise their benefits and **as closed as necessary** to protect the legitimate public and private interests.”*

-OECD Council ¹

A national data governance framework policy is a much-needed step in the direction of unlocking public data for various purposes: from improving service delivery, spurring innovation, to even assisting in the fight against misinformation. The World Health Organisation (WHO) recently highlighted the lack of trustworthy information as a major source of risk, impacting the implementation of crucial policy decisions made during the

¹ [Recommendation of the Council on Enhancing Access to and Sharing of Data, 2021. OECD/LEGAL/0463](#)

pandemic.² Enhancing data collection, storage, processing, dissemination and use of data in decision-making, while safeguarding privacy and security, is a key step.

However, three key aspects need to be considered for the effective implementation of the policy. First, there has to be an assessment of the implementation capacity of the proposed institutions. For instance, the policy envisions that the proposed India Data Management Office (IDMO) will perform multiple strategic and operational responsibilities. The IDMO is expected to design and execute all aspects of the framework, right from identification, storage and management of datasets to ensuring accessibility, availability, responsible usage and governance of these datasets. The policy also puts the onus of capacity building and enforcement on the IDMO. Given the expansive list of roles and responsibilities allocated to the IDMO, it seems that the policy has positioned the unit as a silver bullet to tackle the issue of government departments operating in silos, which has plagued effective data sharing and accessibility in India.

Second, the learnings and experience of implementing previous policies should inform the roadmap for the current one. For instance, the National Data Sharing and Accessibility Policy (NDSAP),³ 2012 which also aimed to improve the accessibility of government data, has faced several hiccups in its decade-long run that deserve particular attention. This is discussed in further detail in the note.

Finally, the policy will have to align with other existing and proposed data regulations and policies in the country that attempt to govern data. These include existing open government data policies,⁴ future data protection legislation⁵ and plans for improving data sharing by private companies⁶ to build a coherent data governance ecosystem in the country.

Overall, the present draft of the policy seems to be a high-level overview of the envisioned National Data Governance Framework. The policy needs to provide more detail to make the public consultation process more effective. In this note, we highlight some of the gaps in the policy and provide recommendations that can inform its final draft and implementation. These include detailed recommendations on building a publicly accessible catalogue for the proposed datasets program, charting the jurisdictional clarity of the proposed policy and institutions, and establishing a grievance redressal mechanism independent of the proposed institutions.

² [Infodemic, WHO statement](#)

³ [National Data Sharing and Accessibility Policy \(NDSAP\), 2012](#)

⁴ Ibid.

⁵ [Data Protection Bill, 2021](#)

⁶ [Report by the Committee of Experts on Non-Personal Data Governance Framework, 2020](#)

Comments and Recommendations

Our comments and recommendations on the policy are divided into the following sections:

- A. Definitional and jurisdictional clarity
- B. Design of institutions
- C. Data sharing protocols
- D. Policy implementation and enforcement

A. Definitional and jurisdictional clarity

2. Objectives & Purpose of Policy

2.1. To transform and modernize Governments data collection and management processes and systems through standardised guidelines, rules and standards for the collection, processing, storage, access, and use of Government data – with the objective of improving governance through a whole-of government approach towards data-led governance.

2.2. To enable and catalyze vibrant AI and Data led research and Start-up ecosystem, by creating a large repository of India datasets. This will be achieved by establishing guidelines, rules and standards to build and access to anonymised non personal data to ensure the growth of Indian datasets. This will be the catalyst for Artificial Intelligence and analytics ecosystem, which in turn would be kinetic enablers of India's digital economy.

2.3. Objectives of the NDGFP are:

- a. To accelerate Digital Governance.*
- b. To have standardized data management and security standards across whole of Government;*
- c. To accelerate creation of common standard based public digital platforms while ensuring privacy, safety, and trust;*
- d. To have standard APIs and other tech standards for Whole of Government Data management and access.*
- e. To promote transparency, accountability, and ownership in Non-personal data and Datasets access. For purposes of safety and trust, any non-personal data sharing by any entity can be only via platforms designated and authorised by IDMO.*
- f. To build a platform that will allow Dataset requests to be received and processed.*
- g. To build Digital Government goals and capacity, knowledge and competency in Government departments and entities.*
- h. To set quality standards and promote expansion of India Datasets program and overall non-personal Datasets Ecosystem.*
- i. To ensure greater citizen awareness, participation, and engagement.*

1. Citizens, as individuals or groups, should also be able to actively participate in their country's data governance ecosystem that the policy envisages. Hence, we recommend that sections 2.2 and 2.3 recognise citizens as active stakeholders and this should reflect in other aspects of the policy, such as in Section 5.3, by allowing

citizens access to the India Datasets program. This will further the policy's objective of ensuring citizen participation and engagement.

2. The National Data Governance Framework Policy (NDGFP), 2022 succeeds other national data policies such as the National Data Sharing and Accessibility Policy (NDSAP), 2012 and National Geospatial Policy (NGP),⁷ 2021. The NDSAP enabled better access to the data generated by ministries and departments at the national level. The NGP, on the other hand, is aimed at providing citizens and enterprises access to the geospatial data generated at the central and state level and freeing them from licensing requirements for the generation and use of spatial data. Since the NDGFP is not the first attempt at establishing data governance frameworks for public data, it needs to distinguish its goals from previous efforts. It also needs to demonstrate more direction and planning in identifying and fixing shortcomings in the implementation of previous policies.
3. The policy needs to clarify how it will work in tandem with previous and proposed data policies and regulations. This will reduce siloed and duplicated efforts while enabling clear delineation on aspects such as:
 - a. Jurisdictions and responsibilities of the related **institutions and actors**:
 - i. The functioning of the Chief Data Officers (CDOs) under NDSAP and those proposed under the draft NDGFP,
 - ii. The functioning of the proposed IDMO (under NDGFP) in conjunction with other relevant data bodies such as the proposed DPA (Data Protection Bill, 2021), Geospatial Data Promotion and Development Committee (under NGP, 2021), and the existing NDSAP Project Management Unit (under NDSAP, 2012), and
 - iii. The functioning of the proposed IDMO in conjunction with other sector-specific regulators in the country.
 - b. Criteria for publication of any datasets on a particular **data platform** vis-a-vis other platforms (such as data.gov.in under NDSAP, the proposed Geo-Platform under NGP, NITI Aayog's NDAP⁸ and the proposed India Datasets platform under NDGFP) and,
 - c. Criteria for the **type of access** (open, registered, restricted, request-based, etc) for different kinds of datasets across policies.
 - d. Approach to **identification and classification of datasets** prescribed across policies (for instance, creation of negative lists as prescribed by NDSAP).

B. Design of Institutions

5. Institutional Framework

5.1. An "India Data Management Office (IDMO)" shall be set up under the Digital India Corporation ("DIC") under MeitY and shall be responsible for framing, managing and periodically reviewing and revising the Policy. The IDMO shall be responsible for

⁷ [National Geospatial Policy \(NGP\), 2021](#)

⁸ [National Data and Analytics Platform](#), NITI Aayog

developing rules, standards, and guidelines under this Policy that shall be published periodically.

5.6. Every Ministry/Department shall have Data Management Units (“DMUs”) headed by a designated CDO who shall work closely with the IDMO for ensuring implementation of the Policy

5.7. The IDMO shall be staffed at DIC by a dedicated government data management and analytics unit

1. The policy should specify the structure of the IDMO, and the composition of the teams that will perform the various functions within its remit. The IDMO will have to build specialised teams with technical and domain expertise given that it is tasked with a wide range of responsibilities cutting across legal, technical and capacity-building activities. The activities will range from creating and maintaining the India Datasets program and building technical capacity across government agencies to drawing up vendor contracts and managing stakeholders and more. For the human resources required on the data and technology front, we recommend that IDMO refer to the structure of the Data and Strategy Unit proposed by the Development Monitoring and Evaluation Office (DMEO)⁹ in order to improve digitization processes across ministries and departments. The ToR proposes the organisation structure of a Data and Strategy Unit to have four sub-units: monitoring, statistical, technology and data analytics. It also details the various roles required at each sub-unit and the skills required for each role.
2. The institutional framework described in the policy requires clarity. The policy needs to clearly state the roles and obligations of the various actors including the IDMO, the Data Management Units (DMUs), the State Level Officers and the Chief Data Officers (CDOs). Further, as mentioned in our recommendation for Section 2, the policy needs foresight on the potential regulatory and functional overlaps resulting in multiple agencies having jurisdiction over similar activities. This will require delineation of the roles of new bodies (IDMO, DMUs) in conjunction with pre-existing bodies¹⁰ at different levels of the government. For instance, the activities being carried out by the Indian Urban Data Exchange (IUDX) or Smart Cities Open Data Portal (SCODP) may overlap with a DMU set up at the Ministry of Housing and Urban Affairs under the proposed policy resulting in internal friction, unnecessary costs and data findability challenges for end-users. We also recommend that the proposed CDOs in the policy be dedicated roles rather than an existing bureaucrat with several responsibilities. The recruitment process for CDOs should also screen them for technical skills required to effectively perform their responsibilities.
3. We recommend that the IDMO be adequately budgeted¹¹ considering the range of resources that will have to be deployed there. For a rough estimate, the policy can

⁹ [Data and Strategy Unit: Detailed Terms of Reference\(ToR\)](#), Department of Development Monitoring and Evaluation Office

¹⁰A potential reference for providing such clarity is [EU Data Act](#) that clarifies the consistency of the proposal with existing rules on processing of data in EU.

¹¹ [Institutions for data governance: Building trust through collective action](#). Data for Better Lives, World Bank

look at the expenditure of national statistics offices in India and elsewhere that also have similar capacity requirements. To provide a comparable estimate, a study conducted in 2019¹² found that it costs countries about \$2.9 billion USD annually to support the development of sound statistical bodies. The budgeting of IDMO will be a crucial factor in determining effective implementation.

4. We also recommend that the IDMO have representation from civil society, academia and industry to ensure inputs from diverse stakeholders throughout the implementation of the policy. A good example of such an engagement can be seen in the implementation of the EU's Infrastructure for spatial information in Europe (INSPIRE) program that builds spatial infrastructure for the purposes of EU environment policies.¹³ The implementation body formed under INSPIRE is supplemented by a number of stakeholder organisations who are invited to give input on INSPIRE's work. Further, the implementing agency forms targeted sub-groups and invites experts from the stakeholder organisations, to address specific implementation issues from time to time.

6.17 Redressal Mechanism: The IDMO shall institute a mechanism for citizens to request datasets, register grievances and establish responsibility of DMUs under the IDMO to respond in a timely manner, to facilitate transparent and accountable data sharing ecosystem

The policy makes IDMO in charge of not only processing data requests but also deciding upon grievances registered by the citizens. This can potentially compromise the independence of the redressal process. In addition to that, the draft policy invests the IDMO with multiple responsibilities, which will create a high administrative burden on the institution, considering the large volume of datasets to be managed on the India Datasets platform. The IDMO may not be able to give sufficient attention to all grievances across departments. In a similar example where the Reserve Bank of India had to deal with grievances related to digital transactions, it set up the *ombudsperson scheme* which could serve as a useful reference.¹⁴ Ombudspersons have become core components of open government initiatives¹⁵ even in other national and sub-national level jurisdictions prioritising citizen concerns. In principle, ombudsperson offices can exercise extensive powers and autonomy to investigate grievances registered by citizens and mediate conflicts in case of maladministration while interacting with a public agency. This characteristic positions them uniquely as a defender of public interest.¹⁶

We urge that the policy establish an independent ombudsperson office with the main purpose of handling stakeholder complaints and representing people's interests in the data sector.

¹² PARIS21 (2019), "[Financing challenges for developing statistical systems: A review of financing options](#)", PARIS21 Discussion Paper, No. 14, Paris.

¹³ [Infrastructure for spatial information in Europe \(INSPIRE\)](#)

¹⁴ Reserve Bank of India. [Ombudsman Scheme for Digital Transactions](#), 2019.

¹⁵ OECD. [The Role of Ombudsmen Institutions in Open Government](#). OECD Working Paper on Public Grievance No. 29.

¹⁶ Ibid.

C. Data Sharing protocols

5.3 IDMO will design and manage the India Datasets platform that will process requests and provide access to the non-personal and/or anonymized datasets to Indian researchers and start-ups.

In line with our recommendations on Section 2 of the policy, we reiterate that besides researchers and start-ups, **citizens**, as individuals or groups, should also be able to place requests to access the datasets on the proposed India Datasets program. This will boost transparency and participatory governance in the country.

6.3 India Datasets program: IDMO will enable and build the India Datasets program, which will consist of Non-personal and anonymized datasets from the Government entities that have collected data from Indian citizens or those in India. Private entities will be encouraged to share such data.

We suggest that metadata access to the catalogue of the India Datasets program should be freely available through APIs without any barriers to ensure better data discovery by all end-users. We also recommend that dissemination standards for non-personal and anonymised data adopted by the NDGFP use technical standards and formats suggested by international best practices. The research and data access policy of the Office of National Statistics, Government of the United Kingdom, provides a useful reference.¹⁷

6.8 Datasets Access and Availability: The IDMO shall notify protocols for sharing of non-personal datasets while ensuring privacy, security and trust. The IDMO will notify rules to provide data on priority/ exclusively to Indian/ India based requesting entities. The IDMO will also judge the genuineness and validity of data usage requests, for datasets other than those already made available on Open Data portal.

6.9 Limits to Data Requests: The IDMO will retain the rights to decide whether requesting entities may be allowed access to full databases/ datasets or combinations thereof, for their use cases

Section 6.9 of the policy says that the IDMO reserves the right to refuse access to certain entities without making it clear if there are any guidelines or recourse mechanisms for this.

1. The *EU Directive on open data and re-use of public sector information*¹⁸ provides a useful reference in this scenario. It states that the public body in charge must provide grounds for refusal of data requested. Based on this, we suggest that an additional clause be added to the policy obligating the IDMO to provide adequate grounds for refusal of access to any dataset.
2. However, if the data requesting entities wish to challenge IDMO's decision on refusal of access to data, our recommendation on Section 6.17 to appoint an independent

¹⁷ [ONS Research and Data Access Policy](#), Office of National Statistics, Government of the United Kingdom

¹⁸ [Directive \(EU\) 2019/1024](#) of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information

ombudsperson's office, as part of the grievance redressal will provide the necessary recourse mechanism to do so.

D. Policy Implementation and enforcement

6.12 Capacity & Skill Building: The IDMO shall support holistic and comprehensive capacity building initiatives for officials in all government agencies to build data and digital literacy, knowledge, and skills. The IDMO shall also assist in setting up DMUs in Ministries and Departments to create dedicated capacity for data management.

6.16 Awareness Building: the IDMO shall ensure adequate awareness building by sharing SOPs, FAQs, Operating Manuals and shall also ensure appropriate branding for quick adoption of the Policy.

6.17 Implementation Manual: The detailed implementation guidelines including the data sharing toolkit, operational manuals, mechanisms for data anonymization and privacy shall be brought.

1. The policy should consider learnings from the implementation of the NDSAP in order to inform its capacity-building initiatives. This is because the roles of the officials outlined in the current policy are similar to the steps mandated for each department by the NDSAP to ensure the release of qualitative data. Our evaluation of the implementation of NDSAP identifies a few notable gaps.¹⁹ These include capacity issues including lack of nominations to CDO posts, delays in curating and communicating data-sharing lists and lack of preparation of action plans that allow tracking of the enforcement of the policy. The current policy needs to address the shortcomings of the NDSAP in order not to fall prey to similar challenges. For instance, lateral hiring of CDOs could be considered given the capacity constraints of existing bureaucrats to take up additional responsibilities. Further, we recommend incentivising ministries and departments, proactive publication of data publishing roadmaps and commitments, and strengthening community engagement.²⁰
2. The resources required for effective implementation of the policy can benefit from contributions made by civil society, academia and industry stakeholders. Hence, we recommend that resources such as the implementation manual, data sharing toolkits, operating manuals, FAQs and training materials be open for continuous public contribution and feedback. The Etalab Guides²¹ provide a good reference for creation of such resources. Etalab Guides provide an online repository of resources and tools targeted at roles similar to that of the CDOs and other officials at the data management units (DMUs). It allows interested external contributors to provide feedback using the website or through the GitHub repository of the page.²² We also recommend a periodic assessment of needs of the DMUs across ministries/departments and subsequent revisions to the data sharing toolkit.^{23 24} This

¹⁹ For a detailed evaluation, refer to this [IDFC Institute report](#)

²⁰ Ibid. For detailed recommendations

²¹ [Etalab Guides](#)

²² [GitHub resource by Etalab Guides](#)

²³ [Manuais e Orientações - Portal Brasileiro de Dados Abertos](#)

²⁴ [Open Data in 60 Seconds \(World Bank toolkit\)](#)

is crucial to provide appropriate and context-relevant policy, legal and technical support to optimally manage risk associated with data sharing over time.

6.18. User Charges: The IDMO may decide to charge User charges/ Fees for its maintenance/ services

The above clause 6.18 suggests that there could be potential user chargers and access to datasets may not be completely free. However, we suggest user charges must not be applied for two reasons: one, the marginal costs of providing this data to an additional user are minimal. Second, open public data can generate numerous positive externalities in the future that could prove greater than the current cost of putting out such data. The Economic Survey 2019²⁵ also argues that data collected by governments should be treated as a public good and placing user charges could flow against this recommendation. Charging users to access government-collected data could limit the various downstream benefits (for example, of innovation through research).²⁶ The tax revenues that can be realised at a later juncture from additional growth could be greater than the revenue generated from user chargers.

²⁵ [Economic Survey 2019-20](#)

²⁶ See paper on [Big open data for development: A vision for India](#)