Recalibrating Non Personal Data

This editorial is based on "**Data marketplaces: the next frontier"** which was published in The Hindu on 21/03/2024. The article examines various facets of non-personal data (NPD) in today's context, emphasising the need to leverage its benefits by creating scalable solutions and incorporating NPD into the provision of public services.

For Prelims: NASSCOM Report, Data, **Artificial Intelligence (AI)**, **Non-Personal Data (NPD)**, Ministry of Electronics and Information Technology (MeiTY), Kris Gopalakrishnan, National Data Governance Framework Policy (NPD Framework)**,** India Urban Data Exchange.

For Mains: Relevance of Non Personal Data in the Era of Artificial Intelligence.

The importance of digitisation in achieving India's goal of becoming a USD 5 trillion economy cannot be emphasised enough. According to a NASSCOM report, data and **Artificial Intelligence (AI)** could contribute around USD 450-500 billion to India's **GDP** by 2025.

However, the rapid digitisation of government operations is leading to a significant increase in the amount of citizen data being generated. This data typically falls into two categories - **Personal Data** which includes information that can identify individuals, and **Non-Personal Data (NPD)** which excludes personal information.

Application of high value advanced analytics and AI to NPD across key sectors of the economy can help predict socially and economically sound outcomes. Junctures where such data-driven insights can better inform governance and public functions are meteorological and disaster forecasts, infrastructure capacity and citizen use-patterns, mobility and housing patterns, and employment trends, to name a few.
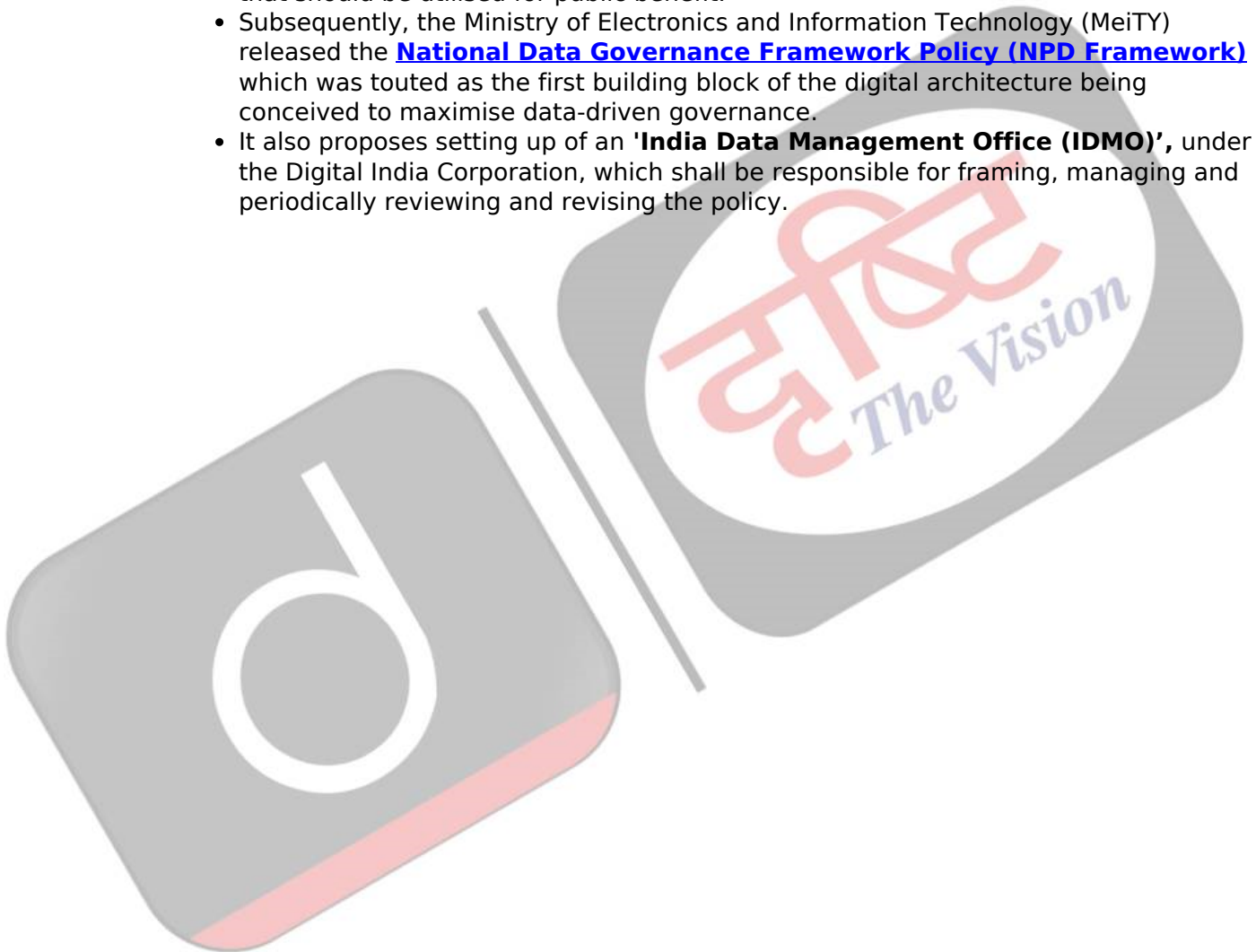
What is Non Personal Data?

- **About:**
  - Any data which is not personal data is categorised as non-personal data. In terms of origin, non-personal data can be data which is never related to natural persons (such as data on weather or supply chains), or data which was initially personal data, but has been anonymised (through use of certain techniques to ensure that individuals to whom the data relates to cannot be identified).
- **Types:**
  - **Public Non-Personal Data:** Data collected or generated by the government in the course of publicly funded works. For example, anonymised data of land records or vehicle registration can be considered as public non-personal data.
  - **Community Non-Personal Data:** Raw or factual data (without any processing) which is sourced from a community of natural persons. For example, datasets collected by municipal corporations or public electric utilities.

- **Private Non-Personal Data:** Data which is collected or generated by private entities through privately owned processes (derived insights, algorithms or proprietary knowledge).
- **Scope**:
    - NPD constitutes the primary kind of citizen data obtained by the government, which possesses the potential of serving as a 'public good'. To create synergies and devise scalable solutions, integration of NPD in the dispensation of public services is generally being advocated for.
- **Indian Context:**
    - The **National Strategy for Artificial Intelligence**, for instance, contemplates making some types of government data available for the 'public good' and mandating corporations to share aggregated data, as a means of overcoming the hurdle of limited data access within India's AI ecosystem.
        - Elsewhere, the **2018-2019 Economic Survey** of India likened data to a natural resource and stated that personal data, once anonymised, becomes a 'public good' that should be utilised for public benefit.
        - Subsequently, the Ministry of Electronics and Information Technology (MeiTY) released the **National Data Governance Framework Policy (NPD Framework)** which was touted as the first building block of the digital architecture being conceived to maximise data-driven governance.
        - It also proposes setting up of an **'India Data Management Office (IDMO)',** under the Digital India Corporation, which shall be responsible for framing, managing and periodically reviewing and revising the policy.

Datasets falling
between clearly
personal data
and clearly
anonymized data
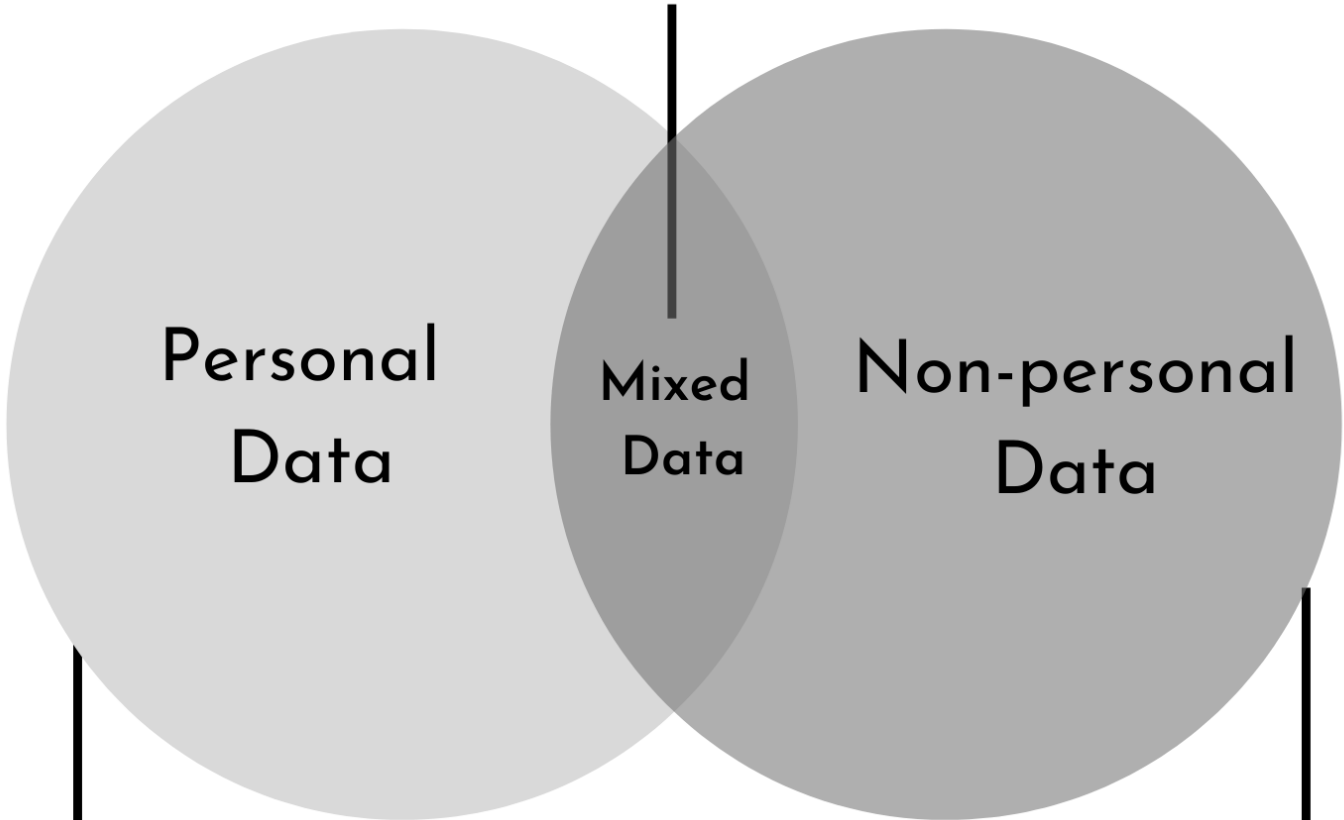
Personal Data

Mixed Data

Non-personal Data

Contains characteristics of a person used to identify them

Data related to finance, health, political beliefs, sexual orientation, caste-tribe, biometric, intersex, transgender, religious affiliations

'Non-personal data' is usually any set of data that does not contain personally identifiable information

It also includes dataset which was personal data earlier but now are "Anonymized"

What are the Different Concerns Associated with Non Personal Data?

- **Sensitivity Issues Involved:**
    - Unlike personal data, which contains explicit information about a person's name, age, gender, sexual orientation, biometrics and other genetic details, non-personal data is more likely to be in an anonymised form.
        - However, in certain categories such as data related to national security or strategic interests such as locations of government laboratories or research facilities, even if provided in anonymised form can be dangerous.
        - Similarly, even if the data is about the health of a community or a group of communities, though it may be in anonymised form, it can still be dangerous.
- **Lack of Effective Regulation:**
    - Unfortunately, unlike Personal Data, there is a stark absence of regulation for NPD. As of date, less efforts have been made at the executive level to construct governance policies for the same.
    - 'Report by the Committee of Experts on Non-Personal Data Governance Framework' highlighted the lack of effective regulation and stressed upon the urgency of effectively regulating the non-personal data in India along the lines of personal data.
        - Experts also believe that the final draft of the non-personal data governance framework must clearly define the roles for all participants, such as the data principal, the data custodian, and data trustees.
- **Undue Advantage to Big Techs:**
    - A government committee headed by Infosys co-founder Kris Gopalakrishnan in 2020 has suggested that non-personal data generated in the country be allowed to be harnessed by various domestic companies and entities, which raises significant concerns:
        - The data sets will heavily favour big tech companies. Only big tech companies possess the capital and infrastructure to create such large volumes of data. Others will find it difficult to match the capabilities of these technology giants.
- **Issues of Mixed Datasets – Containing Both Personal and Non-Personal Data:**
    - The reality of mixed datasets and the inevitable overlap between the two means that a clear demarcation is not tenable.
        - The language of the **Digital Personal Data Protection (DPDP) Act, 2023** alludes to mixed datasets falling under the ambit of personal data protection standards, much like the application of the **General Data Protection Regulation (GDPR)** in Europe.
    - While it may be possible for data to be non-human and non-personal, the distinction becomes murky when data is derived from an individual, especially considering the challenges of anonymisation.
        - This issue is a point of contention even within the GDPR framework but seems to be overlooked in the proposed legal framework, which is concerning given the mandatory data sharing envisioned in the DPDP Act 2023.
- **Lack of Effective Utilisation of NFD:**
    - **None of the above-mentioned laws** (DPDP Act, 2023 and NPD Framework) **provide for an enforceable regime for NPD in India.** For this reason, vast reservoirs of NPD stand unregulated and are supported only by limited guidance in dissemination, use, or exchange thereof.
        - Such a de-siloed accumulation results in sub-optimal legal and policy decisions, and engenders sub-par strategies at sectoral and national levels.
- **Unprotected Interflow Across Departments:**
    - The unprotected inter-flow of NPD across government departments, third-parties, and citizens can make sensitive aspects of NPD vulnerable due to privacy breaches. This can unfairly benefit capacity-carrying actors like Big Tech.
        - The imperfect analysis of crucial public trends can result in faulty decision-making. Such exchange of data is also inefficient as it fails to unlock the power of interdisciplinary legislative and policy-making.
- **Issues in NPD Framework:**
    - The NPD Framework, being a pioneering step, also exhibits several gaps. It formulates

abstract high-level principles and objectives for NPD governance but lacks tangible, actionable guidance to achieve them.
- While legislation is expected, practical operationalisation is overlooked, leaving questions unanswered regarding stakeholder rights and obligations across sectors.
    - Additionally, mechanisms for pricing of data and appropriate legal structures for data exchange are not addressed. The absence of standardised governance tools compounds these challenges.

What Measures Need to be Adopted to Effectively Harness NFD?

- **Critical Assessment of NPD Framework:**
    - A critical evaluation of the NPD Framework to address the existing gaps will be beneficial. This will supplement MeiTY's effort to regulate NPD and will help forge data exchanges as suitable media to make NPD interoperable across sectors.
    - By creating a regulatory design for data exchanges in India, public-welfare functions can be digitised and automated to a large extent. This reduces administrative burden, facilitates inter-sectoral integration, builds the safeguards to using/sharing NPD, and makes digitisation of civic functions more participatory in nature.
- **Formulating Blueprints Governing Data Exchanges:**
    - Data exchanges are scalable ecosystems which galvanise multiple stakeholders. This makes them a fertile ground for deploying advanced analytics for outcome-oriented decision making and helps achieve economies of scale.
    - In India, Telangana has created an agriculture data exchange, and the Ministry of Housing & Urban Affairs, in partnership with the Indian Institute of Science, has established the India Urban Data Exchange.
        - The Department of Science & Technology also plans to set up data exchanges to implement aspects of the **National Geospatial Policy**.
    - With growing interest in data exchange structures, it is important to develop a blueprint for governing them in India. This examination will align with global discussions on regulating data exchanges and support the efforts of MeiTY and other bodies in governing Non-Personal Data (NPD) in India.
- **Lessons From European Union (EU):**
    - In 2019, the EU came out with a regulation framework for the free flow of non-personal data in the European Union, in which it suggested that member states of the union would cooperate with each other when it came to data sharing.
    - Such data, the EU had then ruled, would be shared by member states without any hindrances, and that they must inform the "commission any draft act which introduces a new data localisation requirement or makes changes to an existing data localisation requirement".
- **Recommendations of Expert Committee:**
    - The Expert Committee constituted by the MeiTY to study various issues relating to non-personal data submitted its report in July, 2020. **The Committee made following recommendations:**
        - **Formulating Roles in NPD Governance Framework**: The data principal is the entity to whom the non-personal data relates to. This entity can be an individual, a community, or a company. Data principals may exercise rights over their data through a representative entity, called data trustee.
            - The Committee recommended establishing 'data business' as a new category of business in the country. Entities (including government agencies) which collect, process or store data beyond a threshold (as specified by the regulator) will be classified as data businesses.

- ○ **Non-Personal Data Authority**:
  - • This regulatory authority will be established for putting in place the framework for governance of non-personal data. It will consist of experts in fields such as data governance and technology.
  - • The Authority will be responsible for framing guidelines with respect to data sharing and risks associated with non-personal data.
- ○ **Sharing of Non-Personal Data:**
  - • Any entity may raise a data-sharing request for a: (i) sovereign purpose (such as national security or legal requirements), (ii) public interest purpose (policy making or better delivery of services), or (iii) economic purpose (to provide for a level playing field or for a monetary consideration).
  - • The Committee recommended that public data, community data or private data (limited to raw/factual data collected by a private entity) can be requested at no remuneration.
- ○ **Rights of Community Over NPD:**
  - • The Committee held that a community can exercise rights over non-personal data. It defines community as any group of people that are bound by common interests and purposes and are involved in social or economic interactions.
  - • The community could be a geographical community or an entirely virtual community.
- ○ **Data Custodians and Processors:**
  - • Data custodian is a public or private entity which undertakes collection, storage, processing, and use of data. Data custodian will have a duty of minimising harms to the concerned community.
  - • A data processor is defined as a company that processes non-personal data on behalf of a data custodian. Data processors will not be considered a data custodian under the framework.

Conclusion

While NPD holds promise as a 'public good' and can enhance public services, its unregulated status poses risks, including de-anonymisation and unfair advantages for certain entities. The current governance framework, including the National Data Governance Framework Policy, lacks enforceability and operational clarity, leaving NPD largely unregulated and hindering its potential benefits.

To address these challenges and leverage the potential of NPD, a comprehensive regulatory design for data exchanges is essential. By formulating a blueprint for governing data exchanges, India can enhance the digitisation of public-welfare functions.

> **Drishti Mains Question:**
>
> Define non-personal data and explain its importance in the context of digital economy.
>
> Discuss the challenges associated with the regulation and governance of non-personal data.

UPSC Civil Services Examination, Previous Year Questions (PYQs)

Prelims:

Q1. 'Right to Privacy' is protected under which Article of the Constitution of India? (2021)

(a) Article 15
(b) Article 19

(c) Article 21
(d) Article 29

Ans: (c)

Q2. Right to Privacy is protected as an intrinsic part of Right to Life and Personal Liberty. Which of the following in the Constitution of India correctly and appropriately imply the above statement? (2018)

(a) Article 14 and the provisions under the 42nd Amendment to the Constitution.
(b) Article 17 and the Directive Principles of State Policy in Part IV.
(c) Article 21 and the freedoms guaranteed in Part III.
(d) Article 24 and the provisions under the 44th Amendment to the Constitution.

Ans: (c)

Mains:

Q. Examine the scope of Fundamental Rights in the light of the latest judgement of the Supreme Court on Right to Privacy. **(2017)**