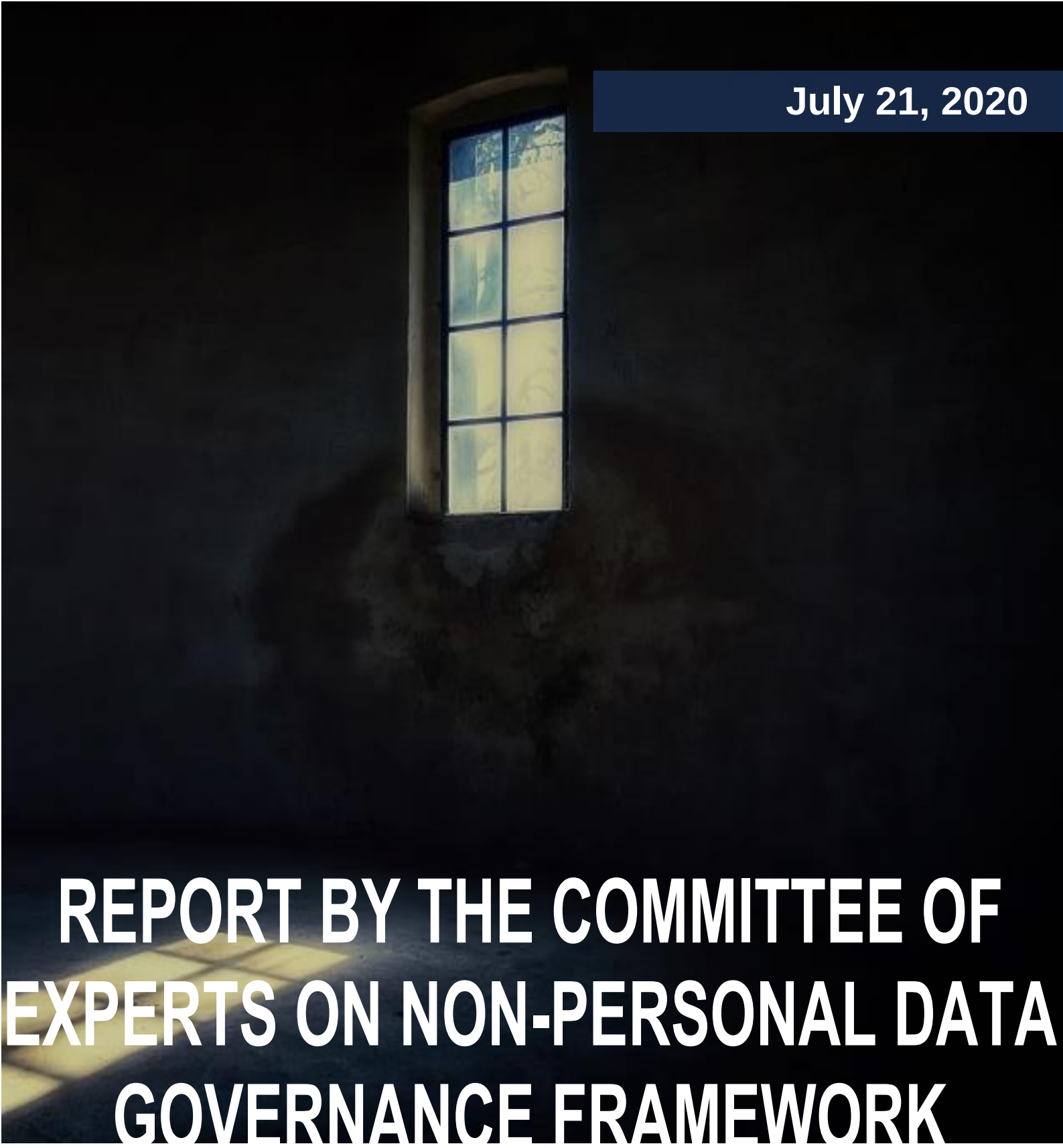

July 21, 2020

REPORT BY THE COMMITTEE OF EXPERTS ON NON-PERSONAL DATA GOVERNANCE FRAMEWORK

- AN ANALYSIS

argus partners
SOLICITORS AND ADVOCATES
MUMBAI I DELHI I BENGALURU I KOLKATA I AHMEDABAD

Introduction

Just as there has been an exponential spurt in awareness regarding the sanctity of personal data and the need to safeguard its privacy, there has also been the development, in parallel, of a philosophy that data is a national resource that should be exploited for the welfare of the society. Even as the Ministry of Electronics & Information Technology ("**MeitY**") has focussed on the regulation and protection of personal data, it has also been pushing ahead with plans to devise a framework for the regulation of non-personal data ("**NPD**"), which will be in parallel to the proposed Personal Data Protection Bill, 2019 ("**PDP Bill**"). Towards this end, in September 2019, MeitY constituted a committee of experts ("**CoE**") which on July 12, 2020 released its report ("**Report**") for public consultation, seeking feedback from the public on a proposed statute ("**NPD Law**") that will regulate the use of NPD for public welfare and the greater good and that will create an authority ("**NPD Authority**") for overseeing such regulation. The Report also lays down key principles to be incorporated in the NPD Law. We have analysed the key proposals mooted by the CoE below.

Why Regulate NPD?

The Report points out that the world is awash with data. The proliferation of big data, analytics and Artificial Intelligence (AI) has led to the creation of many new information intensive services and also the transformation of existing businesses. Organisations have been discovering newer ways to generate value from data. For a few companies that dominate the digital and data business, the network effects lead to outsized benefits and create a certain imbalance in the data and digital industry. Allowing the possibility of data monopolies, in a large consumer market such as India, could lead to the creation of imbalances in bargaining power vis-à-vis a few companies with access to large data sets accumulated in a largely unregulated environment, on one hand, and Indian citizens, Indian businesses including start-ups, micro small and medium enterprises (MSMEs) and even the government, on the other. Therefore, the government's role is to catalyse the data businesses in a manner that maximizes overall welfare. Given the increasing importance and value of data, governments around the world have realised the need to regulate all types of data, including both personal data and NPD.

Definition and Types of NPD

The Report defines NPD as data which is not 'personal data' (as defined under the PDP Bill), or which does not have any 'personally identifiable information'. Therefore, NPD could be data which is not related to an identified or identifiable natural person or data which was initially personal data, but was later made anonymous through a process of aggregation and anonymisation. Examples of the former include data on weather conditions, data from sensors installed on industrial machines, data from public infrastructure etc. In the latter type of data, individual-specific events should no longer be identifiable.

The Report provides for 3 (three) categories of NPD: (i) public NPD; (ii) community NPD; and (iii) private NPD.

Public NPD is NPD which is collected or generated by the government, or by any agency of the government, and includes data collected or generated in the course of execution of all publicly funded works. NPD which is collected or generated by the government where such data is explicitly afforded confidential treatment under law, shall not constitute public NPD. Anonymised data of land records, public health information, vehicle registration details, details of pollution levels in a city collected for a publicly funded project etc. are examples of public NPD.

Community NPD means NPD, including anonymised personal data, and NPD about inanimate and animate things or phenomena, whether natural, social or artefactual, whose source or subject pertains to a community of natural persons. A 'community' is defined as any group of people who are bound by common interests and purposes, and involved in social and/or economic interactions.

A community could be defined by geography, life, livelihood, economic interactions or other social interests and objectives. It could also be a virtual community.

Thus, 'raw / factual data', without any processing / derived insights, collected by the municipal corporations and public electric utilities, telecom, e-commerce, ride-hailing companies etc., would be community NPD. Community NPD shall exclude private NPD.

Private NPD, means NPD collected or produced by persons or entities other than the governments, the source or subject of which relates to assets and processes that are privately-owned by such person or entity, and includes those aspects of derived and observed data that result from private effort. Private NPD includes inferred or derived data and insights involving application of algorithms, proprietary knowledge. It may also include such data in a global dataset that pertains to non-Indians and which is collected in a foreign jurisdiction.

Sensitive NPD

The CoE has recognised that NPD could be sensitive data if: (i) it relates to national security or strategic interests; (ii) it contains business sensitive or confidential information; or (iii) it is anonymised data, that bears a risk of re-identification. The CoE is of the view that NPD shall inherit the sensitivity of the underlying personal data from which the NPD is derived. Also, even if the underlying personal data is not sensitive, or even when there is no underlying personal data, the NPD may still be sensitive if it relates to vital infrastructure, which is sensitive from a national security perspective.

Management of Anonymised Data

The CoE is convinced that large collections of anonymised data can be de-anonymised, especially when using multiple NPD sets. Therefore, anonymised personal data should continue to be treated as the NPD of the relevant data principal. If any subsequent harm arises from re-identification, or from processing such NPD otherwise, the data principal shall be able to take suitable recourse. The data principal's consent should be obtained for anonymisation and usage of such anonymised data. Such consent should be taken upfront at the time of collection of his/her personal data.

Key Stakeholders in respect of NPD

The data principal, data custodian, data trustee and data trusts are key stakeholders in respect of NPD.

Just as in the case of personal data, the data principal in case of any NPD is the natural person with respect to whom such NPD relates. In case of community NPD, the community, that is the source and/or subject of community data, may be treated as the data principal for such data.

The data custodian undertakes collection, storage, processing and use of NPD. Data custodians may be public entities such as such as government ministries, telecom companies or private sector entities such as e-commerce entities. Data custodians must adopt prescribed anonymisation standards and use NPD in a manner that is in the 'best interest' of the data principal. Data custodians have a 'duty of care' to the individual or community from which NPD has been collected. The data custodian may be equated to the data fiduciary under the PDP Bill. The CoE expects that the NPD Law, while providing community data rights will also lay down principles and guidelines for data custodians.

Each data principal community is expected to exercise its data rights through an appropriate community data trustee. Principles and guidelines about who can become a trustee of community

data is expected to be laid out in the NPD Law. The CoE has hinted that for a majority of community data, the relevant government entity or community body may act as the data trustee.

It would be possible to create institutional structures in the form of trusts, for containing and sharing a given set of data. Such a trust would be bound by specific rules and protocols. Data custodians may voluntarily share data in such data trusts. When governments or data trustees seek mandatory sharing of important data for a sector for specific purposes, such data would be managed by data trusts. It is expected that the forthcoming NPD rules and regulations will flesh out in detail how data trusts may be constituted and how they should function.

The CoE has not mandated that the data trustee needs to be located in India or that the trusts should be established in India.

Rights over NPD

As mentioned above, in case of NPD derived from the personal data of an individual, the data principal in respect of the personal data will continue to be the data principal in respect of the NPD. Such NPD has to be utilised in the 'best interests' of the data principal. In the case of community NPD collected in India, the rights over such data shall vest with the data trustee of that community, with the community being the beneficial owner of such data. Benefits accruing from the processing of community NPD, should accrue not only to the organisations that collects such data, but also equally to the community that typically produces the raw or factual data that is being captured. Such NPD has to be utilised in the 'best interests' of the community. At present, it is unclear what would amount to 'best interests', though it may be expected that the forthcoming NPD rules and regulations will flesh this out in detail.

The CoE also believes that since public NPD is derived from public efforts, it should be considered to be a national resource.

NPD relating to individuals is often not just a proprietary or personal asset, but may be considered a collective or shared asset because many parties have overlapping legitimate contributions to, and interests in, it. Therefore, a community may have rights over data relating to the community that may be collected by private data custodians or public organisations. Based on this principle, a private data custodian's drones taking pictures of agriculture farms of local farmers, with or without standing crops, and using it to analyse soil types, health of crops etc. may be considered to be community data.

Unless private NPD collected by a private organisation relates to a community, it would not have to be shared with any third party without remuneration. Further, algorithms and proprietary knowledge need not be shared with any third party.

Data Businesses

Every commercial enterprise has the potential or ability to derive new or additional economic value from data collected by it in the course of its business activity. Data business is not, therefore, an independent industry sector. It is a horizontal classification cutting across different industry sectors. Any business enterprise, whether public or private, which collects data in the course of its business, must register as a 'data business' once it reaches a certain data-related threshold. Below such threshold, registration as a data business shall be voluntary. Threshold requirements may vary with time, context and need and will be fixed and intimated by the NPD Authority, if needed in consultation with sector regulators.

Registration is expected to be a one-time activity and there shall be no need to obtain a license to be a data business. Initial registration shall require a business ID (or country code and country business ID), name(s) of the digital platform or business, associated brand names, rough data

traffic and cumulative data collected in terms of number of users, records and data. The nature of data business, and kinds of data collection, aggregation, processing, uses, selling, data-based services developed etc. should also be stated.

Once the data traffic or collection exceeds set limits, the data business shall be required to submit meta-data about the data users and community from which data is collected, with details such as classification, closest schema, volume, etc. This will be as per a directory of data classification and schema published by the NPD Authority.

Data businesses shall be required to disclose data elements collected, stored and processed, and data-based services offered. The report can be made in digital format. Every data business must declare what they do and what data they collect, process and use, in which manner, and for what purposes (like disclosure of data elements collected, where data is stored, standards adopted to store and secure data, nature of data processing and data services provided).

The CoE calls for harmonisation of data-related directories and disclosures required for personal data and NPD, so that businesses supply the same information only once. The meta-data about data being collected, stored and processed by the data business shall be stored digitally in meta-data directories in India. Open access shall be provided within India to these meta-data directories to Indian citizens and India-based organisations. By analysing the meta-data, potential users shall be able to identify opportunities for combining data from multiple data businesses and/or governments to develop innovative solutions, products and services. Subsequently, data requests may be made for the detailed underlying data.

The compliance requirements imposed on a data business shall be irrespective of whether the business is regulated by another sectoral regulator or not. The sectoral regulators can ride on top of the data business compliance requirements i.e. use these compliance requirements as a base and add any sector specific data disclosure requirements as they may deem fit.

Data Sharing

The CoE strongly believes that providing individuals and organisations with open-access to metadata and regulated access to the underlying data of data businesses will spur innovation and digital economy growth at an unprecedented scale in India. Such access will necessitate the establishment of mechanisms to support data requests and data sharing.

Data may be requested for purposes of national security, law enforcement or regulatory purposes. This would cover mapping security vulnerabilities and challenges, crime mapping, pandemic mapping, devising anticipation and preventive measures, and for investigations and law enforcement. Data may also be requested for community benefits or public good, research and innovation, policy making or for better delivery of public services. It may be requested to encourage competition and provide a level playing field or encourage innovation through start-up activities (economic welfare purpose), or for a fair monetary consideration as part of a well-regulated data market.

The CoE calls for appropriate data sharing mechanisms to be established for sharing public, community and private data. The government should improve on existing open government data initiatives, and should ensure that high-quality public NPD sets are available. With respect to private data, only the raw / factual data pertaining to community data that is collected by a private organisation needs to be shared, subject to well-defined grounds at no remuneration. At points or levels where processing value-add is non-trivial with respect to the value or collective contribution of the original community data and collective community resources used, (or otherwise for reasons of overriding public interest) data sharing may still be mandated but for a fair, reasonable and non-discriminatory remuneration.

A data request may be made to the relevant data business for the detailed underlying data. Alternatively, they may be made to a data custodian based on the meta-data of the data custodian. If the data custodian services the request, the transaction is complete. If not, the request can be escalated to the NPD Authority which will evaluate the request from social, public and economic benefit perspective. If the request is genuine and can result in such benefits, the NPD Authority will request the data custodian to share the raw or factual data.

The principles for storage of data as enunciated in the PDP Bill are recommended for NPD as well. Sensitive NPD may be transferred outside India, but shall continue to be stored within India. Critical NPD (which will follow the definition of critical personal data which is to be notified by the central government) can only be stored and processed in India. General NPD may be stored and processed anywhere in the world.

NPD Authority

The CoE has recommended the creation of an authority for the collection, processing, storage and sharing of NPD. Before making such a recommendation, the CoE considered whether: (i) the sharing of NPD could be self-regulated by businesses and other stakeholders; (ii) various sectoral regulators could address issues that are related to NPD; and (iii) the Data Protection Authority (DPA), proposed to be set up under the PDP Bill, can deal with NPD as well in coordination with the Competition Commission of India (CCI) and other sectoral regulators?

The proposed NPD Authority shall be tasked with enabling legitimate sharing requests and requirements, and with regulating and supervising corresponding data sharing arrangements involving data businesses, data trustees and data trusts. The NPD Authority shall also be responsible for addressing market failures and supervising the market for NPD.

The burden of administering the NPD Law shall be on the NPD Authority. This shall include, exercising various powers for regulating 'data businesses', defining and updating threshold values for registration as a data business, supervising data porting and sharing mandates and requests, managing the meta-data directories, adjudicating on data-sharing disputes etc. The NPD Authority would be expected to certify rules and technology frameworks for various kinds of data sharing, data safety, anonymisation etc. and set standards in this regard.

Technology Architecture

The CoE has penned the following guiding principles for setting up a technology architecture for the implementation of the NPD Law:

- All sharable NPD and data sets created or maintained by government agencies, companies, start-ups, universities, research labs, non-government organisations, etc. should have a Representational State Transfer (REST) application programming interface ("**API**") for accessing the data.
- Data sandboxes can be created where experiments can be run, algorithms can be deployed and only output be shared, without sharing the data.
- Data storage should be in a distributed format so that there is no single point of leakage. Data sharing should be undertaken using APIs only, such that all requests can be tracked and logged. All requests for data must be operated after registering with the company for data access. Even when data is stored in a distributed or federated form, as appropriate, there could be coordinated management as would be required for data trusts and data infrastructures for important NPD in different sectors.
- The data exchange approach must be standardised, regardless of data type, exchange method or platform. Data that is collated should be available appropriately on a data exchange for stakeholders to use and make inferences. The data exchange should be able

to take-in any form of data and produce output that is standardised and usable to all stakeholders.

- De-anonymisation of anonymised data should be prevented. This may be achieved through the best of breed differential privacy algorithms.

This paper has been written by Vinod Joseph (Partner) and Deeya Ray (Associate).

DISCLAIMER

This document is merely intended as an update and is merely for informational purposes. This document should not be construed as a legal opinion. No person should rely on the contents of this document without first obtaining advice from a qualified professional person. This document is contributed on the understanding that the Firm, its employees and consultants are not responsible for the results of any actions taken on the basis of information in this document, or for any error in or omission from this document. Further, the Firm, its employees and consultants, expressly disclaim all and any liability and responsibility to any person who reads this document in respect of anything, and of the consequences of anything, done or omitted to be done by such person in reliance, whether wholly or partially, upon the whole or any part of the content of this document. Without limiting the generality of the above, no author, consultant or the Firm shall have any responsibility for any act or omission of any other author, consultant or the Firm. This document does not and is not intended to constitute solicitation, invitation, advertisement or inducement of any sort whatsoever from us or any of our members to solicit any work, in any manner, whether directly or indirectly.

You can send us your comments at:
argusknowledgecentre@argus-p.com

Mumbai I Delhi I Bengaluru I Kolkata I Ahmedabad

www.argus-p.com