CHANAKYA LAW REVIEW (CLR)

Vol VI, Issue I January- June, 2025 pp. 112-129

DECODING NON-PERSONAL DATA AND ITS GOVERNANCE: LEGAL PERSPECTIVES IN INDIA'S EVOLVING DIGITAL ECONOMY

Dr. Kuldeep Singh Panwar[1]

Jaishree Gaur[2]

ABSTRACT

In the digital era, data generation and analysis form the backbone of economic innovation and policymaking. A significant portion of this data particularly that used in Big Data analytics is machine-generated and categorized as Non-Personal Data (NPD). This article explores the legal definition, characteristics, and implications of Non-Personal Data within the context of India's digital ecosystem. It further examines the key elements of the proposed Non-Personal Data Governance Framework developed by the Government of India. By comparing the two versions of the draft reports, the article provides a critical analysis of the evolving regulatory approach and its potential impact on data governance and economic development.

Key Words- Data Governance, Non-Personal Data, Anonymized Data, Data Principal, Expert Committee

Introduction

'The Data can serve as the stones that enable one to cross the river.'- Deng Xiaoping

The buzz words for various industrialists, governments, academicians, researchers, policy makers, enterprises etc. of the 21st century can be said to be 'data', an intangible stuff which acts as a fuel for driving the wagons of the 4th Industrial Revolution. Some call data as the 'new

[1] Associate Professor, HOD, Department of Law, Nagaland University, Lumami
[2] Research Scholar, Department of Law, Nagaland University, Lumami

oil'[3] whiles others term the flow of data across the world as 'the lifeblood of digital economy'[4]. But in its simplest form data can be understood to mean 'transmissible and storable computer information.'[5] Also, in the Indian legal context, the definition of data has been provided under the Information Technology Act, 2000, as,

"*a representation of information, knowledge, facts, concepts or instruction which are being prepared or have been prepared in a formalized manner, and is intended to be processed, is being processed or has been processed in a computer system or computer network and may be in any form (including computer printouts magnetic or optical storage media, punched cards, punched tapes) or stored internally in the memory of the computer*."[6]

Data is gaining so much limelight due to the benefits it offers for those who want to utilize it, for example, companies collect and analyze data of an individual's online habits to target advertisements to those individuals and in turn gain economic benefits. As a person's online habits may directly or indirectly expose a lot about that person's likes and dislikes, ideology, food habits etc., so now even the governments and political parties are also gaining interest in these datasets for policy making and election strategies. Hence a new 'data arms race'[7] has already begun and so governments of various countries are coming up with regulations to protect their national interest by securing the data of their citizens by enacting privacy laws, regulating cross border flow of data, some even consider data as national asset and hence facilitating the social and economic benefits that might by utilizing it as a public good and ensure overall economic growth. Huge amounts of data are being generated every day, as is evident from an estimate which predicted that the size of the digital universe is expected to reach around 44 zettabytes by 2020.[8] To bring this into perspective this number can be thought

[3]Kiran Bhageshpur, "Data is the New Oil — and That's a Good Thing", (2019) Forbes, available at: https://www.forbes.com/sites/forbestechcouncil/2019/11/15/data-is-the-new-oil-and-thats-a-good-thing/?sh=62a78fbb7304 (visited on March 12, 2025).

[4] World Economic Forum, "A New Paradigm for Business of Data: Briefing Paper", (2020), available at: https://www.weforum.org/reports/new-paradigm-for-business-of-data/ (visited on April 19, 2025).

[5]D. Harshaavardhan, "Origin of Data", (2019) Medium, available at: https://medium.com/@15euee043/origin-of-data-b06fa66ab56f (visited on March 20, 2025).

[6]Information Technology Act, 2000, s. 2(o).

[7] Atlantic Council & Thomson Reuters, "Is Data the 21st Century Arms Race?", (2017) Atlantic Council, available at: https://www.atlanticcouncil.org/in-depth-research-reports/report/big-data-a-twenty-first-century-arms-race-2/ (visited on March 29, 2025).

[8] Jeff Desjardins, "How Much Data is Generated Each Day?", (2019) World Economic Forum, available at: https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/ (visited on April 21, 2025).

to be 40 times more bytes than the observable stars.[9] Some of the statistics that highlight the amount of data generated are given below:

1. 294 billion emails and 500 million tweets are sent out every day
2. Daily data created on Facebook amounts to around 4 petabytes and the number of searches done every day is around 5 billion.

Based on these figures it was estimated that by 2025, 463 exabytes of data would be created every day.[10] This is being made easy everyday with better technologies available on the smartphones itself. According to an article published on Statista, the number of smartphone users in the world nearly amounted to 2.7 billion in 2019 which is expected to exceed 3.8 billion in 2021. In a similar vein, it was predicted that 748 million people in India would own smartphones by 2020, and 1.5 billion by 2040.[11]

This massive data generation leads to various social and economic benefits. Primarily the value creation out of the data happens as illustrated in the figure below:

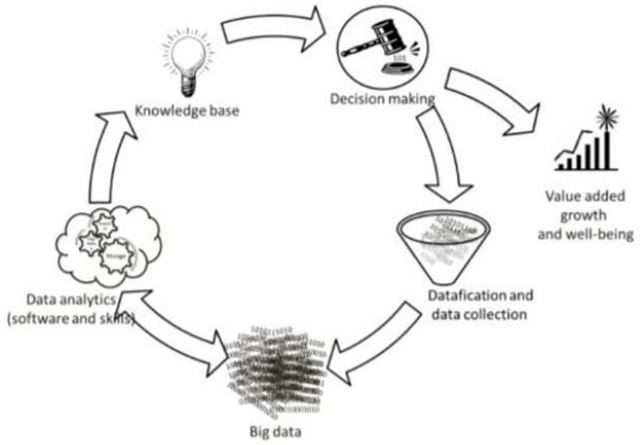

Figure 1 Data value creation[12]

It is evident from the figure that this process of value creation out of data is not linear and involves, at different phases, feedback loops. The OECD (Organisation for Economic Co-

[9] *Id.*
[10] *Id.*
[11] Sophia Sun, "Number of Smartphone Users in India 2010 to 2020 with Estimates until 2040", (2023) Statista, available at: https://www.statista.com/statistics/467163/forecast-of-smartphone-users-in-india/ (visited on April 20, 2025).
[12] Organisation for Economic Co-operation and Development, "Mapping Approaches to Data and Data Flows: Report for the G20 Digital Economy Task Force", (2019), available at: https://www.oecd.org/sti/mapping-approaches-to-data-and-data-flows.pdf (visited on March 29, 2025).

operation and Development) report for "the G20 Digital Economy Taskforce"[13] describes the 2 stages by which social and economic value of data is extracted:

1. Through analysis, data is transformed into knowledge which can be used for gaining insights.

2. Decision making is one of the most important aspects of any business; hence insights from data can help in taking action and making decisions.

Among the data generated in the digital economy major parts comprise of the machine generated data. The data is directly or indirectly collected through different sources using different means and technologies. Data analytics is the most important function of Big Data application; this function is achieved using this machine generated data. The data is mainly utilized by the big industries and firms for development and monetary benefits. The data analytics successfully provides the right view of the market for what the demand is high and what are new requirements. This further helps in maintaining the conventional ideas which are highly being used and also gives way to new innovation and ideas for the needs which are not fulfilled. In this way machines generate data helps in increasing the economy as well as promoting new innovations and ideas. This machine generated data is referred to as Non-Personal Data.

After analyzing the importance of Non-Personal Data in the digital economy it becomes necessary that these data are properly regulated and inclusive legislation must be made for the same.

Non - Personal Data and the Need for its Governance in the era of Digital Economy

Non-Personal Data is not uniformly defined by any legal statute. It is a new concept which is coming forward with the enlightenment of personal data and privacy. Generally, Non- Personal Data are described as "what is not personal data"[14] but not specifically defined. In the most fundamental form of Non-Personal Data are any dataset which does not have personally identifiable information. Data which has no individual identifiable information or anything by which the person to whom that data belongs can be identified or the data which does not belong to any individual is called as Non-Personal Data.

[13] *Id.*

[14] Personal Data Protection Bill, 2019, No. 373, as introduced in the Lok Sabha.

"The European Data Protection Framework acknowledges two categories of data: personal and Non-Personal Data. There is data that is always non-personal (because it never related to an identified or identifiable natural person) and there is also data that once was personal but no longer is (as linkage to a natural person has been removed)."[15]

The legal test to differentiate between Personal and Non-Personal Data is embodied in "**Recital 26 GDPR"** [16]according to which:

"Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information, should be considered to be information on an identifiable natural person. To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments."[17]

Data not caught by this test falls outside the scope of European data protection law. Indeed, Recital 26 GDPR goes on to state that:

"The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable."[18]

The term information used has broader perspective. It includes both objective information as well as subjective analysis.[19]

[15] Michèle Finck & Frank Pallas, "They Who Must Not Be Identified: Distinguishing Personal from Non-Personal Data under the GDPR", (2020) 10 International Data Privacy Law 11.
[16] *Id.*
[17] Regulation (EU) 2016/679 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation), [2016] OJ L 119/1.

[18] *Id.*
[19] Article 29 Data Protection Working Party, "Opinion 04/2007 on the Concept of Personal Data (WP 136)", (2007), available at: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf (visited on March 29, 2025).

In the case[20] the European Court of Justice (ECJ) made it clear that while the information in a residence permit application and the data used in legal analysis both qualify as personal data, the assessment-related legal analysis does not. Various forms of data can be videos, photos, alphabetical or numerical data. The fact that term "information is used rather than "data" is used in Article 4(1) GDPR suggests that the data must have some informational value.

In a case[21] the Court ruled that 'the main objective of the EU legislative is not to restrict the scope of the concept, the information is not limited to the sensitive or private but includes all the different types of objective and subjective information' can be can be interpreted from the use of term 'any information.'[22]

The identification of information as personal data is based on the perspective. It is important to see from whose perspective data is identified and assessed.

Breyer's Case is the primary case relating to this issue in the European Courts.[23] Mr. Breyer had viewed a number of German federal government websites that kept log files of information about access operations.[24] This data includes the visitor's dynamic IP address, which is an IP address that varies with each new Internet connection. Breyer claimed that the storage of the IP address was against his constitutional rights. Static IP addresses "are protected personal data because they enable those users to be precisely identified," the European Court of Justice (ECJ) said in the *Scarlet Extended Case*.[25] The recital 30 GDPR, IP addresses are regarded as online identifiers as well. Although a dynamic IP address does not contain information on an identifiable natural person, it can be used to create log entries that can be used to trace the personal information, if the ISP has the requisite supplementary information.[26] Despite the fact that the information needed to identify Mr. Breyer was held by the ISP rather than German authorities, the dynamic IP address was nonetheless classified as personal data in this way.[27]

[20] YS v. Minister voor Immigratie, Integratie en Asiel and M and S, [2014] EU:C:2014:2081.
[21] Peter Nowak v. Data Protection Commissioner, [2017] EU:C:2017:582.
[22] *Id.*. at para. 34.
[23]Breyer v. Germany, [2016] EU:C:2016:779.
[24] *Id.*
[25]Scarlet Extended SA v. SABAM, [2011] EU:C:2011:771.
[26] *Id.* at para 39.
[27] *Id.* at para. 49.

There are several actors in the society who have the ability to control Non-Personal Data and prevent others from doing so this is called as "data ownership".[28] It is based on the technological measures which strengthen the position of these handfuls of people who have ownership of the Non-Personal Data over others who cannot access it. Data ownership in the hands of few people provides power concentration in the private actors resulting in the economic power and socio-political power which leads to the very harmful effects on the society as whole.[29]

The Non-Personal Data holds the utmost importance in the digital economy and have the tendency of playing major role in the future development of Nation as whole. So, it becomes very important to properly regulate the Non-Personal Data. There is a dire need of legal rules enhancing access to Non-Personal Data in large scale processing environments.

Non-Personal Data Governance in India

The *Digital Personal Data Protection Act, 2023* (DPDP Act, 2023), enacted on August 11, 2023, represents India's first comprehensive legislation for personal data protection, but it does not regulate non-personal data (NPD). This aligns with the approach of the Digital Personal Data Protection Bill, 2022, which similarly excluded NPD from its scope, as noted in earlier analyses. In contrast, the *Data Protection Bill, 2021*, had proposed including NPD within the ambit of the Data Protection Act, with the Joint Parliamentary Committee (JPC) arguing that NPD, often derived from anonymized personal data, could not be easily distinguished from personal data during processing. The JPC recommended a single Data Protection Authority (DPA) to oversee both personal and non-personal data, citing the need to avoid grey areas in anonymization and re-identification. However, this recommendation was not adopted in the DPDP Act, 2023, which focuses solely on personal data, leaving NPD governance unaddressed by statutory law as of May 2025.

Earlier versions of data protection legislation, such as the *Personal Data Protection Bill, 2019* and *2018*, also did not regulate NPD, though the 2019 bill included provisions allowing the Central Government to direct data fiduciaries to provide anonymized personal data or other NPD for policymaking and service delivery. Section 17 of the DPDP Act, 2023, retains a similar provision, empowering the Central Government to request anonymized data for

[28] Wolfgang Kerber, "A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis", (2016) Gewerblicher Rechtsschutz und Urheberrecht, Internationaler Teil (GRUR Int) 989.

[29] *Id.*

evidence-based policy formulation. This continuity suggests a limited role for NPD in enabling public good objectives, but without a comprehensive regulatory framework.

The absence of NPD-specific legislation as of May 2025 contrasts with the recommendations of the Committee of Experts on Non-Personal Data Governance Framework, constituted by the Ministry of Electronics and Information Technology (MeitY) in September 2019 under Kris Gopalakrishnan. The committee's draft reports (July 2020 and December 2020) and final report (November 2021) proposed a dedicated Non-Personal Data Authority (NPDA) to regulate NPD, classifying it into public, community, and private categories. These recommendations, however, have not been implemented, and the JPC's preference for a single DPA was not reflected in the DPDP Act, 2023, which establishes a DPA solely for personal data. Recent consultations by MeitY, such as the 2024 discussion paper on data governance, indicate ongoing interest in NPD regulation, emphasizing data as a public good to support economic growth and innovation. These developments suggest that a future NPD framework may emerge, but no concrete legislation exists as of May 2025, leaving a regulatory gap in India's digital ecosystem.

Non-Personal Data expert committee

September 2019: The Ministry of Electronics and Information Technology (MeitY) first constituted a Committee of Experts for Non-Personal Data Governance Framework (NPDG) to come up with a data governance framework in 2019. Infosys co-founder Kris Gopalakrishnan was asked to lead the committee.

July 2020: The expert committee released its draft report to the public for consultation and feedback. It defined non-personal data as any data that is not related to an identified or identifiable natural person or is personal data that has been anonymised.[30] It proposed that NPD should be regulated by a new regulatory body, the Non-Personal Data Authority (NPDA). This data, the committee recommended, should be further classified into three categories— public NPD, community NPD, and private NPD.[31]

[30] PRS Legislative Research, "Non-Personal Data Governance Framework", (2023) *PRS India*, available at: https://prsindia.org/policy/report-summaries/non-personal-data-governance-framework (visited on March 4, 2025).
[31] *Id.*

December 2020: The committee then released a revised report addressing several concerns raised in the 1500 submissions received by them.[32] The submissions are not public. The committee wrote that the framework should become the basis of new legislation for regulating NPD.

November 2021: The committee submitted the final report to MeitY which has not been made public yet. The final report contains consultations and feedback received on the revised report. The final draft reaffirmed the committee's position on a national Non-Personal Data Protection Authority reporting to the Indian government.[33] It appears the committee's recommendations were not heeded by the Joint Parliamentary Committee working on the Data Protection Bill.

Salient Features and Analysis of the Proposed Non-Personal Data Governance Framework in India

The salient features of Non-Personal Data Governance framework are analyzed in the discussions below.

Definition: Non-Personal Data has been defined by the committee in consonance to the definition available under section 91 (2) of the PDP bill, 2019 which states that "Non-Personal Data means the data other than personal data."[34]

The report also takes into account the definition mentioned in the "Guidance on the regulation on a framework for the free flow of Non-Personal Data in the European Union, 2019" document to further explain the two types of Non-Personal Data:

(a) Data not related to natural persons such as weather data, data from sensors installed on machines etc. (b) Anonymised data, which was initially personal data but after applying some transformation techniques the data is no longer identifiable to any person.

The Non-Personal Data was categorized, in the version 1 report, in 3 different ways:

---

[32] Ministry of Electronics and Information Technology, "Revised Report by the Committee of Experts on Non-Personal Data Governance Framework", (2020) *OurGovDotIn*, available at: https://ourgovdotin.files.wordpress.com/2020/12/revised-report-kris-gopalakrishnan-committee-report-on-non-personal-data-governance-framework.pdf (visited on April 4, 2025).

[33] Money Control, "Government Can Have Just One Authority for Personal, Non-Personal Data: Kris Gopalakrishnan", (2023) *Money Control*, available at: https://www.moneycontrol.com/news/business/government-can-have-just-one-authority-for-personal-non-personal-data-kris-gopalakrishnan-7731491.html (visited on April 5, 2025).

[34] *Supra* note 31.

Public Non-Personal Data[35]

Private Non-Personal Data[36]

Community Non-Personal Data[37]

The revised report, while doing away with such categorization, still retains the underlying concept of public, private and community data.

While explaining the private data in Appendix 3[38], the report says that private data may additionally contain, 'a global dataset containing data about non-residents gathered in foreign jurisdictions (other than India)'[39] Further clarification is required whether sharing of this data for High Value Datasets (HVDs) is exempt or not as it might affect outsourcing into India.

Re-identification of Non-Personal Data: The version 1 report while identifying the risks related to re-identification recommended appropriate measures to be put in place to prevent or minimize such risks.[40] However, the version 2 report, in order to further harmonize both the data governance frameworks, introduced several clarifications.[41]

The committee reflects the opinion of various researchers that there are inherent risks associated with the anonymization of personal data, by acknowledging that the risk is a valid one and data principals (individuals) might need more protection in this regard.

Consent for Anonymized Data: It identifies that a separate consent is needed in case of Non-Personal Data since its nature is different from that of PD, so consent provided for PD can be construed as consent for Non-Personal Data. Hence it recommends for providing the data principals, prospectively, with the option to "opt out" of data anonymization at the time of collecting personal data.[42] It also recommends providing the option of revocation of consent if data is not already anonymized.

Although the opting out option provided in the revised version of the report appears to be a move in the right direction, as it recognizes the right of an individual to have informed choices and provides flexibility, rather than rigidity, for technological mobility which is similar to social mobility in the physical world.

---

[35] *Supra* note 29, at para. 4.2.
[36] *Id.* at para 4.3.
[37] *Id.* at para 4.4.
[38] *Supra* note 31.
[39] *Id.*
[40] *Supra* note 29, para. 3.7.
[41] *Supra* note 31, para. 5.1, 5.2.
[42] *Id.* at para. 5.4.

There might be difficulty in practical implementation of opting out system as has been explained here. Consider a scenario where, once data has been anonymized and compiled into a dataset, then at a later stage of time there might difficulty in identifying an individual who is choosing to opt out. As it would defeat the whole purpose of providing anonymity in the first place itself, if it could be re-identified at a later stage.

However, it is important to provide this kind of flexibility, because in the cyber world technology keeps on improving and so do the policies of the companies provide the services. Of at a later stage of time, say, a company processing Non-Personal Data changes its policy which might have an unintended consequence which some users might not find comfortable to deal with. At that point, this option can come handy and would weaponize the users in their act of cyber self-defense.

This will benefit data principals in making informed choices because any personal data which is being anonymized runs the risk of re-identification. This will in turn increase the trust of Data Principals in data businesses collecting information.

Streamlining Sensitivity and Localization norms of Non-Personal Data: The initial report classified the Non-Personal Data as 'Critical Non-Personal Data' and 'Sensitive Non-Personal Data'.[43]

The revised report does not explicitly make such classification. But in order to provide certain checks and balances to the HVD creation and sharing process, the revised report recommends that those data which are derived from personal data would attract the localization requirement of the personal data from which it has been derived.[44] So, if Non-Personal Data is an anonymized sensitive personal data, then it would inherit the underlying localization requirement of sensitive personal data.

Defining Data Business: The initial report suggested for a new classification of business which would collect, process, store, manage both personal and Non-Personal Data.[45] It can either be a government entity or a private organization.[46] The revised report retained this concept and provided further clarity on the threshold parameters that could be used to classify data business that would be required to register in India.

[43] *Supra* note 29, para. 4.5.
[44] *Supra* note 31, para 8.15.
[45] *Supra* note 14.
[46] *Supra* note 31.

The report also considers threshold parameters as provided by the "California Digital Privacy Law" which might include the following parameters – "gross revenue, number of consumers/households/devices handled, and percentage of revenues from consumer information."[47]

The data businesses would be required to share meta-data (data about data).

For achieving ease of doing business and doing away with the possibilities of license raj, the data business would be required to do one time registration. Report clarified that the existing businesses in various sectors can be data businesses.

Further the report provides that data trustees, data processors, data custodians can be data businesses.[48]

Various data entities proposed by the report for Non-Personal Data ecosystem:

In order to establish rights over Non-Personal Data, the committee envisages number of entities, their roles and responsibilities and their interrelations. These entities are as follows:

Data custodian: Data custodian can either be a private entity of a government organization which would collect, store, process, and use etc., the Non-Personal Data.[49] They will have to share Non-Personal Data as and when the request is made to them. Data Custodian stays in touch with consumers and also collaborates with Data Trustee and Non-Personal Data Authority.

The data custodians have 'duty of care', while handling the Non-Personal Data of a community, towards that community.[50] Therefore, it is their responsibility that no harm must occur to the Non-Personal Data owing to risks of re-identification. This would ensure that data custodians would follow best practices for safe data sharing.

Data Processor: An entity who is responsible for processing data on the behalf of data custodians. Regarding sharing mandate of data processors, the committee clarified that when a data processor is processing Non-Personal Data on behalf of a data custodian, in this case data processor is not expected to share Non-Personal Data. While on the other hand when the data processor is processing its own data, the data processor can be considered as data custodian and would be required to share Non-Personal Data. Examples of data processors

[47] *Id.*
[48] *Id.*
[49] *Id.*
[50] *Supra* note 29, para. 4.8.

include IT/ IT/ ITeS companies, cloud service providers, software-as-a-service provider etc. The Initial Report introduced the concept and the version 2 further clarified and elaborated the concept providing better clarity and understanding.

The committee has made an exemplary move in the direction of providing these new concepts which specifically deal with Data governance and covering the cases of B2C model like E-commerce. Furthermore and more people are indulging in B2B model, for the better functioning cases of B2B can also be included.

Data Trustee: The committee has recommended data trustee for creation, maintenance, sharing of High Value Datasets (HVDs). These trustees can either be government organization or non-profit organization which includes society, trust, and company.[51]

These data trustees also have a 'duty of care, along with that they must also establish grievance redressal mechanisms.

The committee has mentioned term "community can raise grievances", the strict interpretation of this will result in the exclusion of individual grievances but the liberal interpretation will include the same. To avoid conflict in the interpretation further clarification in this regard would be appreciated.

Data trustees have been empowered to levy 'nominal charge' to the data requesters which can be used by them for the maintenance of their data infrastructures.

Moreover, data trustees can also be recognized as data businesses.

Non-Personal Data Authority: Non-Personal Data Authority is proposed to be created as a specialized separate authority to govern the Non-Personal Data ecosystem.[52] This is a newly emerging area of regulation. As a result, it was necessary for the authority to possess expertise and to keep up with the constantly shifting technical environment. The responsibilities and priorities that fall under this jurisdiction are very different from those that are already in existence like the data protection authority (DPA) and CCI and another sectoral regulator.

Some of the functions of Non-Personal Data Authority as proposed in the draft report are addressing privacy issues relating to re-identification; creating a data sharing framework; act as adjudicating authority in case data custodian refuses to share HVD.

Committee states that Non-Personal Data Authority would work in its own sphere and is unlikely to function in areas of other authorities such as Competition Commission of India and Data Protection Authority and the functioning of all the Authorities should be harmonized.

[51] *Supra* note 31.
[52] *Supra* note 29, para. 8.

For better protection and the risk owing to Non-Personal Data sharing the committee also says that the sectoral regulators can create additional data regulations over those created by Non-Personal Data Authority in a horizontal fashion. This might lead to overregulation of the data regime and instead of facilitating the process, might in the long run, hamper ease of doing business.

Furthermore, the committee also recommends establishment of an "Academia-Industry Innovation Body" under Non-Personal Data Authority to develop, enhance, innovate on aspects like data governance, data sharing, interoperability etc. this body will comprise experts from academia, government, industry and society.

High Value Datasets (HVDs): The committee defines the HVDs as the datasets that are beneficial to 'community' at large and can be useful for various purposes such as policy making, creating high-quality jobs, businesses, research and education, financial inclusion, health, energy etc.[53] it considers HVDs as public good.

Community is defined by the committee as 'any group of individuals participating in social and/or commercial activities and united by common interests and goals. It could be a geographical group, a group of people with similar interests and aspirations in terms of their careers or economic well-being, or it could be entirely virtual.'[54]

HVDs are supposed to be created, maintained and shared by the data trustees. But since any group of individuals can organically create data trustee, hence this leaves room for uncertainty and creates various issues. As the question arises, who is to be considered as a good data trustee. There are conflicts and other internal dynamics within a community which the committee seems to ignore. There is risk for underrepresentation/ misrepresentation within the communities themselves, where there might be some groups which already have been in a disadvantaged position and a data trustee which is not correctly identified. It might also happen that the data trustee doesn't have its incentives aligned with all the subgroups of a given community.

Generally, laws in data protection regimes have mostly focused on individual agencies and the rights that are provide over information. The rights are provided as long as they relate to an identifiable individual. The report mentions possibilities of collective harm by inappropriate handling and exposure of this kind of data. So, recognition of group privacy is a new and

[53] *Supra* note 31.
[54] *Id.*

definitely a welcome step, because it brings the consideration of group agency to the forefront of this Non-Personal Data debate.

Data Sharing: The committee refers data sharing to 'restricted access to Non-Personal Data for specified uses with adequate protections.'[55]

The committee clearly lays down the 3 specific purposes for which Non-Personal Data must be shared – sovereign purposes, public good purposes and business purpose.

The committee mentions under what circumstances data should be requested and shared and that the granularity or level of data requested would be determined by the government on a case-to-case basis for data request related to sovereign purposes. Further it bars Non-Personal Data from adjudicating the same.

The committee does not make any recommendation regarding data sharing for business purposes as such mechanism already exists.

Meta-data collected by the data businesses with respect to the Non-Personal Data collected by them must be made available for open access through a directory managed by Non-Personal Data. Access to this open directory is made available to organizations registered India. Moreover, to create certain datasets for public good, data trustees may refer to the meta-data to make requests for relevant sub-sets of data.

The revised report takes a similar stance to the initial report on the issue of mandating data sharing by private entities for public good purposes. This move can prove to be bottleneck for overseas players and is likely to disincentivize domestic players as well.

Also, the amount of data that any big tech or startup has collected after many years of efforts and making decisions to come out with products that created a customer base for them. This is the reason they have abundant Non-Personal Data due to network effect. Now if government mandates data sharing this might defeat the purpose of competition. Moreover, this would disincentivize new startups as well, to actually work towards reaching this goal of coming up with a product that would create their own customer base.

Data sharing, even if done for the benefit of public, may create IPR related issues and privacy issues. The report mentions these issues but do not provide a concrete framework to overcome such issues.

[55] *Supra* note 29, para. 7.

Analysis of Legal Issues Relating to Non-Personal Data Framework

The committee analyzed Non-Personal Data framework from the perspective of 6 laws including the Constitution of India. The following observations have been made by the author in this regard:

Constitution of India: the committee defends its stance of providing the mandatory data sharing framework with the socialistic directive principles mentioned in the constitution in Article 39 (b) and (c). The word "material resources" used in Article 39 (b) can be considered to include informational or data resources. This inference is drawn from case[56] where the term 'material resources' were taken to mean 'everything of value or use in the material world'. The committee also refers to Article 39 (a) and (b), but it would have been sufficient for the purpose of Non-Personal Data sharing mechanism to rely only on Article 39 (b) and (c) as legislations enacted to give effect to these two directive principles can even override the fundamental rights enshrined in Article 14 and Article 19 of the Indian Constitution.

Trade Secrets Law: The committee acknowledges that there is no law at present that codifies trade secret protection in India. But what is trade secret? According to WIPO (World Intellectual Property Organization), "trade secrets are intellectual property (IP) rights on confidential information which may be sold or licensed. In general, to qualify as a trade secret, the information must be commercially valuable because it is secret, be known only to a limited group of persons, and be subject to reasonable steps taken by the rightful holder of the information to keep it secret, including the use of confidentiality agreements for business partners and employees. The unauthorized acquisition, use or disclosure of such secret information in a manner contrary to honest commercial practices by others is regarded as an unfair practice and a violation of the trade secret protection."

The committee acknowledges that there are precedents which might give the interpretation as raw data pertaining to be protected under trade secrets. The Committee also suggests cases which might suggest that raw data must not be granted trade secret protection, hinting the further scope for elaboration.

Copyright Law: In India Database can fall under the "literary work" category under section 2 (o) of the Copyright Act 1957. Any database can be said to have 2 ore components – structure and content. While structure of the database can be protected under the copyright act 1957. But

[56] State of Karnataka v. Ranganath Reddy, (1978) SC 215 (India).

the content can't be granted similar protection as India does not give sui generis rights which are recognized over the content of database. Moreover, copyright law does not protect ideas or raw data.

The interpretation of paragraph 9.3 of the report which deals with copyright law, it can be deducted that there are basically three doctrines which are taken in consideration by the experts. The sweat of the brow and modicum of creativity are two doctrines which decide the copyright standards/ standard of originality and merger doctrine. Sweat of the brow is the low standard and modicum of creativity is the high standard of copyright. India strongly followed the doctrine of 'sweat of the brow' for a considerably long time. However, the standard of 'originality' followed in India is not as low as the standard followed in England. In a case,[57] the Supreme Court discarded the 'Sweat of the Brow' doctrine and shifted to a 'Modicum of creativity' approach as followed in the US.

The Supreme Court in case[58] shifted to Merger doctrine which states that no one may claim copyright in that single manner of expression or depiction because that would evict everyone else from the right to express or depict that idea.

Both Copyrights and Trade Secrets are part of TRIPS framework and India is a member of WTO and committed to TRIPS agreement. The springboard doctrine, which is also recognized by Indian courts, can hamper the mandatory data sharing regime being created by the Non-Personal Data framework. The doctrine says that "a person who has obtained information in confidence is not allowed to use it as a springboard for activities detrimental to the person who made the confidential communication and springboard it remains even when all the features have been published or can be ascertained by actual inspection by any member of public."[59]

Hence this doctrine if applied in case of Non-Personal Data can prove to be a ban for the data sharing mechanisms proposed by the committee. Hence further deliberations are required on this issue.

Competition Law: Competition law might prove useful in tackling issues of data governance, since it stretches over industries which are not subject to targeted regulation. Competition

[57] Eastern Book Company v. D.B. Modak, (2008) 1 SCC 1 (India).
[58] Emergent Genetics India Pvt. Ltd. v. Shailendra Shivam & Ors., I.A. Nos. 388/2004, 1267/2004 & 1268/2004 in CS (OS) 50/2004 (India).
[59] John Richard Brady & Ors. v. Chemical Process Equipments P. Ltd. & Anr., [1987] Delhi 372 (India).

analysis certainly helps to identify power-related problems surrounding "data ownership". The essential facilities doctrine ("EFD") could be viewed as a solution in this respect.

The EFD refers to cases where a monopolist refuses to grant access to goods or services. If these amount to an essential facility for other market players to create their own products, the dominant undertaking has an obligation to 'share them with everyone asking for access, including competitors'. The doctrine, which evolved in EU case law starting in the early 1990s, has been applied both to material infrastructures and incorporeal assets.

Conclusion

India's journey to regulate non-personal data (NPD) remains a pivotal challenge as it aims to become a $5 trillion digital economy. The Digital Personal Data Protection Act, 2023, enacted on August 11, 2023, established a robust framework for personal data but left NPD unregulated, creating a gap that persists as of May 2025. The Kris Gopalakrishnan Committee's 2019–2021 proposals for a Non-Personal Data Authority and a framework categorizing NPD as public, private, and community data have not been implemented, despite their vision of NPD as a public good driving innovation and policymaking. Recent MeitY consultations in 2024 signal ongoing interest in NPD governance, emphasizing its potential to fuel economic growth and social development, yet no legislation has emerged.

NPD holds immense promise for empowering MSMEs and startups by providing access to anonymized datasets and market insights, addressing structural barriers highlighted in India's Economic Surveys. However, mandatory data-sharing risks disincentivizing tech giants, requiring a balanced approach to foster competition. Privacy concerns, particularly re-identification risks in anonymized data, demand transparent safeguards, while intellectual property issues, including trade secrets and limited data copyright protections, need clearer legal frameworks. Competition law tools like the essential facilities doctrine could curb data monopolies, but their application must avoid stifling innovation.

India's approach to NPD governance, guided by SMART principles Simple, Moral, Accountable, Responsive, and Transparent can set a global benchmark. A simple framework ensures compliance ease, moral accountability protects privacy, and transparency builds trust. As MeitY's discussions continue, India must craft a policy that unlocks NPD's value while addressing risks, fostering an inclusive digital ecosystem. By pioneering NPD regulation, India can harness this digital gold, supporting its vision of a dynamic, innovative economy.