## **RESPONSE TO THE**

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

January 2021 | Issue No. 105





### **ABOUT THE ESYA CENTRE**

The Esya Centre's mission is to generate empirical research and inform thought leadership to catalyse new policy constructs for the future. It simultaneously aims to build institutional capacities for generating ideas which enjoin the triad of people, innovation, and value, consequently helping reimagine the public policy discourse in India and building decision-making capacities within government.

Esya invests in ideas and encourages thought leadership through collaboration. This involves curation of niche and cutting-edge research, and partnerships with people, networks, and platforms. Moreover, it prioritises multi-disciplinary research to engender "research clusters", through which practitioners and researchers collaborate.



#### **RESPONSE TO THE**

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

We at the Esya Centre are grateful for the opportunity given to us by the Ministry of Electronics and Information Technology (MeitY) to respond to the Committee of Expert's Report on Non-Personal Data Governance Framework (Report). We appreciate that the Committee has attempted to set out a broad framework that seeks to regulate several facets of the use of Non-Personal Data (NPD) while identifying possible areas of concern.

In our suggestions, we engage with the Committee's key recommendations and identify areas which require greater clarity. We recommend actions that assist in creating effective regulation geared towards achieving defined goals and outcomes.

**Part I** contains the summary of recommendations, and **Part II** contains a detailed analysis of substantive elements of the Report. We have structured our responses into 4 themes namely: i) the definition of Non-Personal Data, ii) overlaps with existing proprietary frameworks, iii) community data and HVDs and iv) the regulatory architecture.



#### **PARTI**

#### SUMMARY OF RECOMMENDATIONS

We appreciate that the Committee has sought to clarify and revise its framework on non-personal data ('NPD'), originally proposed in its first report on the Non-Personal Data Governance Framework ('Report'). However, we find that the second version of the Committee's report ('Revised Report') still has issues that must be resolved.

At the outset, we urge the Committee to reconsider the need for non-personal data regulation at this time, and to first commission studies to understand the existing landscape. There is a need for enhanced understanding of underlying concepts such as data and its intersections with various legal frameworks, and more clarity on the kinds of data collected and analysed by various players, and the economic value of such different forms of data.

Conversely, if the Committee decides to continue to develop its NPD framework, we have identified key issues and areas for reform below, in addition to the recommendations highlighted in our response to the first Report.<sup>1</sup> We frame our analysis and recommendations here under 4 broad themes:

#### 1. Definition of Non-Personal Data

Operationalising the definition of NPD will prove challenging in practice and does not provide sufficient clarity to stakeholders which include consumers, researchers, public authorities, companies and entrepreneurs. The proposed framework's treatment of anonymised personal data is also likely to lead to overlaps with the mandate of the proposed Data Protection Authority and creates avenues for forum shopping and regulatory arbitrage. We recommend that anonymised personal data is dealt with under the personal data protection legislation, and by the proposed Data Protection Authority.

#### 2. Overlaps with existing Proprietary Rights

Although the Committee does not anticipate any overlaps between its data-sharing mechanisms and existing protections such as copyright and trade secrets, there is no clarity on the boundaries of protection offered by these rights in the context of data. Without sufficient clarification, we could see two clear market risks emerge - the erosion of competitive incentive structures which are crucial for private-sector led innovation, and harm to domestic

startups unable to make use of data to scale both domestically and abroad.

We recommend that the Committee advocates that the proprietary rights relating to data are clarified in relevant legislation before the NPD framework is built on unclear concepts.

#### 3. Community Data and High Value Datasets ('HVDs')

The Committee's definition of a community is overbroad and can lead to conflicts of interest. While the intention to protect community rights over data is appreciated, the focus should be on safeguarding the interests of communities which are particularly vulnerable to harm. The process of creating HVDs and appointing Data Trustees outlined in the Report can lead to practical constraints and conflicts of interest. We recommend that any HVDs focus on public data as defined by the Committee, and on incentivising rather than mandating data sharing from private players.

#### 4. Regulatory Architecture

The Committee recommends that a new NPD regulator is set up for various functions ranging from encouraging innovation in startups, processing and storing metadata, adjudicating data sharing issues, addressing privacy and reidentification risks, and addressing 'the negative externalities' of data sharing, among others. As highlighted in our response to the Committee's first report, such a broad mandate (with possible conflicts in goals) would create overlaps with sectoral regulators, make it difficult to formulate accountability mechanisms, increase difficulty in measuring performance and provide opportunities for "mission creep" where the regulatory powers exercised can exceed the intended ambit. We therefore do not believe that the case for a regulatory authority is made out, and urge the Committee to instead focus on creating a forum (which could be based on parallels available in other sectors such as electricity, or similar bodies in other jurisdictions) or other formal mechanisms (such as binding memoranda of understanding) to ensure regulatory cooperation between the various relevant authorities who will need to be consulted in deliberations on NPD.



#### **PART II**

#### **DETAILED ANALYSIS**

The Revised Report builds on the framework that was suggested in the first iteration. Notably, unlocking the economic value of non-personal data in a manner that benefits India and its citizens has emerged as one of the key objectives of the Non-Personal Data Governance Framework. To this end, the Committee has suggested the following major changes from its first Report:

- Amendments to the Personal Data Protection Bill ('PDP Bill') and the introduction of an 'opt-out' mechanism for anonymization of data.
- Introduction of high-value datasets which help operationalise the proposed data sharing framework.
- Removal of 'business to business' data sharing as one of the purposes under the data sharing framework.
- Creation of the Non-Personal Data Authority ('NPDA') as a statutory body under a separate legislation.

While these changes help envision a more effective regulatory framework, the Revised Report has not clarified concepts in important areas, as we highlight below. Suboptimal regulatory design can stifle innovation and inhibit growth in the relevant market,<sup>2</sup> and is likely to stifle the ability of the Government, firms, and the public at large to effectively unlock the economic value present in data. Hence, we identify the key areas where this lack of clarity and suggest alternative approaches to resolve potential confusion and overlap.

#### 1. DEFINITION OF NON-PERSONAL DATA

The Revised Report adopts the same definition of nonpersonal data as the first.<sup>3</sup> This definition differentiates between personal data (**'PD')** and NPD based on whether the dataset in question contains any personal information that can be used to identify an individual. If such information is present, then the dataset is considered to constitute PD. As per this definition, which is based on the EU's Framework for free flow of Non-Personal Data,<sup>4</sup> NPD can broadly consist of two different kinds of data: (i) data that initially related to an individual but has been anonymised to remove any identifiers, such as anonymised health and financial data; (ii) data that never related to an individual and stems from other natural and physical phenomena, such as climate data.

While the definition may appear straightforward, operationalising it will prove to be a challenge for the following reasons:

- i. The PDP Bill requires that anonymization be irreversible. However, research suggests that no existing techniques of anonymization are truly irreversible, in so far as most anonymised datasets can be linked to existing public databases to reveal personal information about individuals. Hence, it is unclear how any dataset can be truly anonymised and, therefore, fall within the ambit of the NPDA unless the definition of anonymisation in the PDP Bill is changed.
- ii. Even assuming that the standard for anonymisation is not complete irreversibility, the regulatory framework proposed in the report creates significant regulatory overlaps that can lead to jurisdictional confusion. For instance, the Report states that if an NPD dataset is re-anonymised, then the jurisdiction to redress any harm caused rests with the Data Protection Authority ("DPA") under the PDP Bill. However, there is no indication as to which authority will decide whether re-anonymisation has occurred with respect to a dataset. In the absence of an institutionalised forum for coordination, such overlaps may undermine the protection of privacy as well as the ability to unlock economic value.

<sup>2</sup> Knut Blind et al., The impact of standards and regulation on innovation in uncertain markets, Research Policy Vol 46 Issue 1, February 2017, accessible at: <a href="https://www.sciencedirect.com/science/article/pii/Soo48733316301743">https://www.sciencedirect.com/science/article/pii/Soo48733316301743</a>.

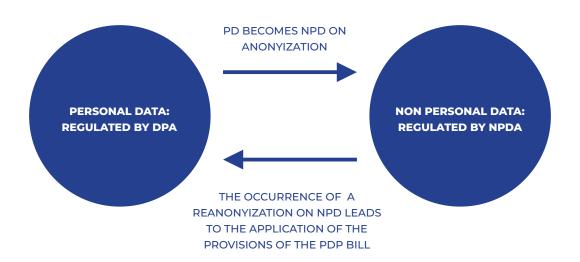
<sup>3</sup>S. 4, Pg. 7, Revised Report.

<sup>4</sup> Regulation for the free flow of Non-Personal Data in the European Union, accessible at: <a href="https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN>">https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807e/from=EN

<sup>5</sup> S. 3(2), Personal Data Protection Bill, 2019.

<sup>6</sup> Luc Rocher et al., Estimating the success of re-identifications in incomplete datasets using generative models, Nat Commun 10 3069 (2019), accessible at: <a href="https://www.nature.com/articles/s41467-019-10933-3">https://www.nature.com/articles/s41467-019-10933-3</a>.





[Figure 1: The fluid relationship between personal and non-personal data under the proposed framework]

- iii. The fluid nature of the definition also does not provide firms and entrepreneurs with requisite clarity as to whether a particular dataset is personal, nonpersonal or mixed. For instance, the Report states that datasets in which personal and non-personal are 'inextricably linked' will be a mixed dataset governed by the provisions of the PDP Bill. However, no guidance is provided on what this term means and how it will be determined. In the EU, where a similar definition of a mixed dataset has been adopted, the regulation itself provides some guidance on when data would be inextricably linked.<sup>7</sup> However, this too has been criticised as vague.<sup>8</sup>
- iv. The above definition also creates avenues for forum shopping and regulatory arbitrage. For instance, the Report states that data custodians will be required to provide an 'opt out' mechanism, with regard to anonymisation of data, at the time of collecting any personal data or information.9 A company which wants to avoid the data sharing requirements under the NPDA Framework can make the optout a default option for all its data collection and, therefore, avoid undertaking anonymisation to remain within the ambit of the PDP Bill.

We recommend that that anonymised personal data is dealt with under the PDP Bill. Not only would this resolve most of the issues illustrated above, but it is also in line with the DPA's mandate to protect privacy.<sup>10</sup>

#### 2. OVERLAPS WITH EXISTING PROPRIETARY RIGHTS

The Committee requires that companies share certain specific subsets of their raw datasets with trustees to form High Value Datasets ('HVD'), which are to be used for public good. Such mandatory data sharing can conflict with existing legal protections, and we appreciate that the Revised Report specifically addresses this potential conflict. We also commend the Committee for examining overlaps with copyright, trade secret law, the Information Technology Act, 2000, and the Competition Act, 2002 based on feedback received on its first Report. However, per the Committee's analysis, there are no overlaps or conflict between the proposed NPD data-sharing mechanism and existing legal frameworks.<sup>11</sup> While we appreciate that the Committee has undertaken this analysis, it is not clear that the ambit of the proposed NPD framework and existing legal protections are

<sup>7</sup> Guidance on the Regulation on a framework for the free flow of non-personal data in

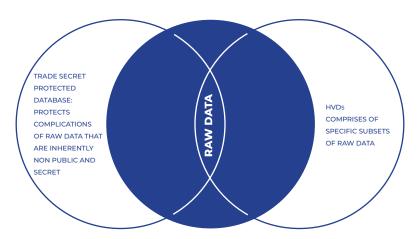
the European Union, accessible at: < https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019DC0250&from=EN>.

<sup>8</sup> Inge Graef at al., Towards a Holistic Regulatory Approach for the European Data Economy: Why the Illusive Notion of Non-Personal Data is Counterproductive to Data Innovation, TILEC Discussion Paper, September 2018, accessible at: <a href="http://ssrn.com/abstract=3256189">http://ssrn.com/abstract=3256189</a>>.

<sup>9</sup> S. 5.4, Pg. 11, Revised Report.

<sup>10</sup> For additional reasons as to why the DPA is better suited to privacy protections, see: P. 6, Response to the Report by the Committee of Experts on Non-Personal Data Governance Framework, Esya Centre, accessible at: <a href="https://bit.ly/2NE7bjP">https://bit.ly/2NE7bjP</a>. 11 S. 9, Pp.32-34, Revised Report.





[Figure 2: Lack of clarity on the differences between and limits of what is protected under trade secrets, what the Committee would require to create HVDs, and 'raw data']

separate. A fundamental issue that also remains unresolved is the lack of clarity on protections offered to data, and on the taxonomy of the various kinds of data.

For instance, a core aspect of the Committee's analysis on copyright and trade secrets rests on the assumption that underlying 'raw' data in both original as well as nonoriginal databases would not receive protection under these frameworks. It states that since copyright only protects 'original' compilations of data, requiring access to the underlying data would not violate copyright law. <sup>12</sup> Under the trade secret jurisprudence, the Committee recognises that there is conflicting precedent on whether the underlying data can be protected, and that a majority of the existing case law is in the context of parties with an existing relationship of confidence. It concludes that existing protection is 'unlikely' to prevent the government from exercising eminent domain over the relevant data.

However, it is not clear that the Committee can rely on the concept of eminent domain as a justification for companies and private parties to share data. Very broadly, eminent domain is the right of the State, as part of its sovereign powers, to take private property for public use on the payment of compensation.<sup>13</sup> 'Property' is at the core of this concept, and it has been interpreted to mean property in all its forms, including exclusive rights such as copyright.<sup>14</sup> However, the Committee notes that legislations in India have not created a property right over data,<sup>15</sup> and copyright is unlikely to protect the underlying 'raw' data. It is therefore unclear how the Committee

relies on this concept to assert its rights. Using the sovereignty of the State to assert rights over data can also create other issues - for example, how would companies be required to treat datasets containing information based on the data of those who are not Indian citizens? Requiring companies to differentiate between data arising from Indian citizens and otherwise can be a time-consuming and expensive process, and may also not always be possible.

In the context of copyright and trade secrets, there is limited case law dealing with databases of the kind that the Committee seeks to regulate. Whether or not a dataset is protected as a trade secret would usually be an ex-post determination made by a Court of law. There is limited clarity on how copyright law would apply in this context - for example, since copyright law protects originality in the arrangement of databases, would requiring companies to share 'metadata' (taken to mean the fields in the database, per the Revised Report) infringe on the copyright of the database owner? The trade secret landscape is more confusing since there is no legislative framework in place, making the scope of the right extremely unclear. In this context, while the Committee acknowledges that there is no database right in non-original databases in India, it is nevertheless important to consider the rights that are to be provided to those that compile/maintain even non-original databases, given the investment of finances and labour in such activities.

<sup>12</sup> S. 9, Pg. 32, Revised Report.

<sup>13</sup> DD Basu, 'Commentary on the Constitution of India (Articles 233 to 307)', 8th Ed, Vol 9, p. 17; available at <a href="https://advance.lexis.com/api/permalink/c65cea7a-ef6b-4d3f-9ae2-923b531cfa48/">https://advance.lexis.com/api/permalink/c65cea7a-ef6b-4d3f-9ae2-923b531cfa48/</a>/'context=1523890>.

<sup>14</sup> DD Basu, 'Commentary on the Constitution of India (Articles 233 to 307)', 8th Ed, Vol 9, p. 8; available at <a href="https://advance.lexis.com/api/permalink/c65cea7a-ef6b-4d3f-9ae2-923b531cfa48/">https://advance.lexis.com/api/permalink/c65cea7a-ef6b-4d3f-9ae2-923b531cfa48/</a>?context=1523890>.

<sup>15</sup> S. 9, Pg. 32, Revised Report.



The Committee recommends the creation of a sui generis framework for data sharing. While clarity on data sharing could be useful, it is premature at this stage. A necessary first step would be to clarify the proprietary rights in data before access can be regulated. Instead of proposing a new framework based on unclear legal protections, the aim should first be to clarify rights in this domain. Without sufficient clarification, we could see two market risks emerge. First, there is a risk that it may lead to an erosion of competitive incentive structures which are crucial for private-sector led innovation. Indeed, the most recent Economic Survey lamented India's poor showing when it comes to investments in R&D, especially those which flow from the private sector, as compared to other advanced countries.16 It is important for India to construct policies and regulation which offers sufficient certainty, which is a key determinant of a jurisdiction's investment climate.<sup>17</sup> Given the global architecture of the Internet, we believe that sufficient policy and regulatory certainty on this front could encourage large global tech players to innovate in India.

Second, for domestic startups, proprietary rights over data (with suitable consumer-side protections) are important from the perspective of growth, scale, and access to capital. Data is often an important underlying asset which investors analyse before committing to a particular organisation. Dower protections or insufficient understanding of a business' legitimate exclusivity over the asset may inadvertently become a hindrance for Indian tech startups to scale both within India and abroad. Already we observe that due to the COVID-19 pandemic, in 2020 Indian tech startups raised less than USD 10 billion for the first time since 2016.

To mitigate such risks, we recommend that the Committee pushes for clarifying or erecting appropriate ecosystem-friendly laws pertaining to data and concomitant proprietary rights before the NPD framework is built on unclear concepts. In this context, the Committee could recommend modernising copyright law, creating a legislative framework for trade secrets, and contemplating standalone protection for non-original databases. The Committee would also do well to focus on ways to incentivise, rather than mandate data sharing in this context. <sup>21</sup> Further, policy must be nimble and in particular aligned with the different private, public and mixed characteristics of data which are highly context-specific determinations.

<sup>16</sup> Read Generally, Innovation: Trending Up but needs thrust, especially from the Private Sector, Economic Survey of India 2020-21, Chapter 8, p. 237-283 available at <a href="https://www.indiabudget.gov.in/economicsurvey/doc/vol1chapter/echapo8\_vol1.pdf">https://www.indiabudget.gov.in/economicsurvey/doc/vol1chapter/echapo8\_vol1.pdf</a>.

<sup>17</sup> Regulatory Reform and Innovation, Organisation for Economic Cooperation and Development (OECD), P 12, <a href="https://www.oecd.org/sti/inno/2102514.pdf">https://www.oecd.org/sti/inno/2102514.pdf</a>; Bastian Schwark, Influence of regulatory uncertainty on capacity investments — Are investments in new technologies a risk mitigation measure?, <a href="https://infoscience.epfl.ch/record/153004/files/15d\_schwark\_paper.pdf">https://infoscience.epfl.ch/record/153004/files/15d\_schwark\_paper.pdf</a>.

<sup>18</sup> Sadowski J. When data is capital: Datafication, accumulation, and extraction, Big Data & Society, January 2019. available at <a href="https://journals.sagepub.com/doi/pdf/10.1177/2053951718820549">https://journals.sagepub.com/doi/pdf/10.1177/2053951718820549</a>

<sup>19</sup> MIT Technology Review Custom + Oracle, The Rise of Data Capital, March 2016, available at <a href="https://files.technologyreview.com/whitepapers/MIT\_Oracle+Report-The\_Rise\_of\_Data\_Capital.pdf">https://www.technologyreview.com/2016/03/21/161487/the-rise-of\_data-capital/</a>; Alex Lazarov, Venture Capital Has a Lot to Learn From Fintech: Data-Driven, New Products, And More Access to a Broader Set of Companies, Forbes, May 2020, available at <a href="https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/venture-capital-has-a-lot-to-learn-from-fintech/?sh=1749f8014fa2>"https://www.forbes.com/sites/alexlazarow/2020/05/13/vent

<sup>20</sup> Tech Crunch, Manish Singh, 'Indian Startups raised \$9.3 Billion in 2020', 27 December 2020, available at <a href="https://techcrunch.com/2020/12/27/indian-startups-raised-9-3-billion-in-2020/">https://techcrunch.com/2020/12/27/indian-startups-raised-9-3-billion-in-2020/</a>.

<sup>21</sup> Annexure 1, Response to the Report by the Committee of Experts on Non-Personal Data Governance Framework, Esya Centre, accessible at: <a href="https://bit.ly/2NE7bjP">https://bit.ly/2NE7bjP</a>>.



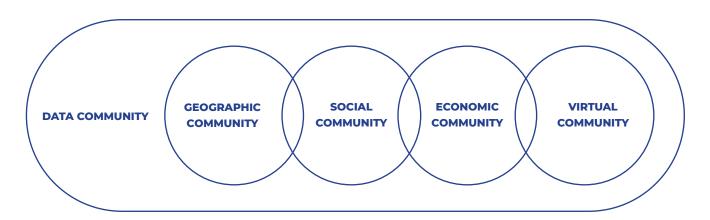
#### 3. OWNERSHIP AND ACCESS

We appreciate the Committee's clarification on the mechanism for the creation of community rights in data. This draft persists with the same definition of community but clarifies how data trustees can be formed. It has also introduced the concept of a 'high value dataset' which is to be a compilation of metadata that is important to a community. Despite these clarifications, the Committee's conception of community rights is fraught with the following problems.

#### i. The Definition of Community

The Committee defines a community as "any group of people that are bound by common interests and purposes and involved in social and/or economic interactions. It could be a geographic community, a community by life, livelihood, economic interactions or other social interests and objectives, and/or an entirely virtual community."22 As illustrated below, this covers a wide range of possible communities which are themselves not clearly defined. The definition also assumes that communities are homogenous and can be represented by a single Government body or non-profit organisation. Other concerns that pertain to the identification of members and knowledge of community members, which were highlighted in the previous response, have also yet to be addressed.<sup>23</sup>

The intention to protect community rights over data is laudable. However, the intention is diluted by defining community in such broad terms. Instead, the focus should be on safeguarding interests of communities which are particularly vulnerable, for instance tribal and indigenous communities. Organisations, such as the Global Indigenous Data Alliance, argue that traditional frameworks of data ownership and sharing do not account for existing power differentials in society.24 A sole focus on greater sharing of data disregards the value of traditional forms of knowledge gathered by communities over significant periods of time. Hence, there is a need to evolve principles and frameworks through which such communities can benefit from the sharing of data that pertains to them and their practices. Other legal frameworks, such as the Forest Rights Acts and the Biodiversity Act, have also established community management over various kinds of shared resources.<sup>25</sup> The Committee can take a leaf out of the systems established under these Acts to a) narrow the definition of a community and b) develop a pragmatic and robust system of community data management.



[Figure 3: The broad definition of community data is based on groups and communities which are undefined and overlap]

<sup>22</sup> S. 7, Pg. 16, Revised Report.

<sup>23</sup> P. 6, Response to the Report by the Committee of Experts on Non-Personal Data Governance Framework, Esya Centre, accessible at: <a href="https://bit.ly/2NE7bjP-.24">https://bit.ly/2NE7bjP-.24</a> Global Indeigenous Data Alliance, Care Principles of Indigenous Governance, accessible at: <a href="https://www.gida-global.org/care">https://www.gida-global.org/care</a>.

<sup>25</sup> Puneeth Nagaraj, Varsha Rao and anr., Community Rights over Non-Personal Data: Perspectives from Jursiprudence on Natural Resources, Data Governance Network, accessible at: <a href="https://datagovernance.org/report/community-rights-over-non-personal-data-perspectives-from-jurisprudence-on-natural-resources-">https://datagovernance.org/report/community-rights-over-non-personal-data-perspectives-from-jurisprudence-on-natural-resources-</a>.



#### ii. Data Trustees and HVDs

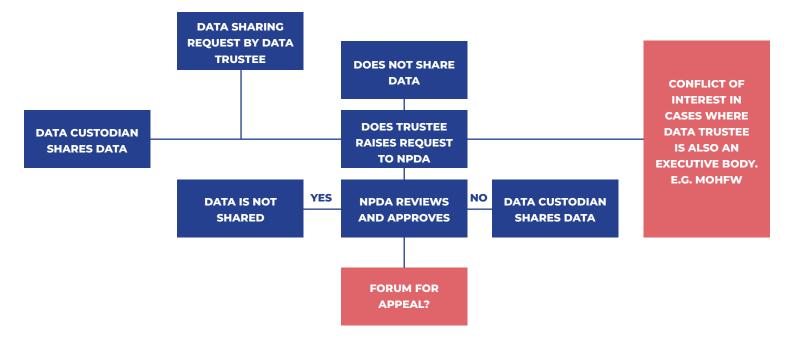
The Committee has clearly outlined how community rights will be enforced and managed. In addition to the role of the data trustee, which can be a Government or private non-profit organisation, that manages community data the concept of a High Value Dataset has also been introduced. HVDs are datasets that are "beneficial to community at large and shared as a public good". To create an HVD, a data trustee needs to approach the NPDA. The NPDA will determine the appropriateness of both the HVD and data trustee as per its guidelines. This could include a minimum expression of interest by the community as well as a public consultation process to map the contours of the HVD.<sup>26</sup>

Despite the safeguards above, the process of creation of HVDs and appointment of data trustees suffers from practical constraints and possible conflicts of interest. The foremost concern is that the possible grounds for the creation of an HVD are very wide and cover potentially all kinds of data that companies may generate.<sup>27</sup> A second set of concerns relates to the potential misuse of data trustees and HVDs. As the Report itself acknowledges, data trustees may themselves be controlled by vested interests and appropriate multiple HVDs. Further, large organisations may register proxy organisations in India just to access the HVDs. Despite

acknowledging these concerns, the Report has not provided any framework through which community members can check the functioning of the trustee. Additionally, the data trustee's mandate to act 'in the interest of the community' does not provide the necessary clarity to ensure its accountability.

Finally, the Report does not discuss any mechanisms to ensure that the NPDA and data trustees perform their duties in a transparent and accountable manner when requiring data custodians to share metadata. Given that both the NPDA and data trustees can be administrative/executive bodies, it is important to provide an independent forum of appeal from the NPDA's decisions pertaining to the creation of HVDs and requests for data.

Other nations, the EU and Australia for example, are also creating high value datasets. However, the primary objective of these HVDs is to make public data easily available and accessible to all. <sup>28</sup> We recommend that the scope of high value datasets be initially limited to the availability and accessibility of public data. The Government is already working towards this objective through its National Open Digital Ecosystems Policy. This can be coupled with a framework that incentivises firms, both large and small, to share their data more openly through fiscal, monetary and contractual incentives. <sup>29</sup>



[Figure 4: The process of data sharing under the framework is riddled with practical concerns and conflicts of interest]

<sup>26</sup> S. 7.8, Pg. 19, Revised Report.

<sup>27</sup> S. 7.6, Pg. 18, Revised Report.

<sup>28</sup> Australia's Open Government National Action Plan 2016-18, accessible at: <a href="https://ogpau.pmc.gov.au/national-action-plans/australias-first-open-government-national-action-plan-2016-18/21-release-high">https://ogpau.pmc.gov.au/national-action-plans/australias-first-open-government-national-action-plan-2016-18/21-release-high">https://oce.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information>.

<sup>29</sup> Annexure 1, Response to the Report by the Committee of Experts on Non-Personal Data Governance Framework, Esya Centre, accessible at: <a href="https://bit.ly/2NE7bjP>">https://bit.ly/2NE7bjP></a>.



#### **4. REGULATORY ARCHITECTURE**

We had pointed out in our response to the Committee's first Report that the proposed NPDA potentially overlapped with sectoral regulators, and we appreciate that the Committee has sought to address these overlaps. However, no major changes have been proposed to the regulatory structure other than highlighting the need for legislation on NPD. The Committee suggests that the role of the proposed NPDA is different from sectoral regulators and that the best way to regulate this emerging area is to set up a new regulator, but does not clarify why this is the case.

As such, the concerns highlighted in our response to the Committee's original Report remain:<sup>30</sup>

- Lack of clarity in definitions and terms: How concepts such as community data will be operationalised is still unclear, as are issues of what a community is, how communities with differing aims must be reconciled, and how overlapping communities with different "best interests" would be addressed.
- Overbroad and unclear mandate: The proposed functions remain the same, and encompass a wide range of actions from providing support to startups, to addressing the negative externalities of data sharing, processing and storing metadata, adjudicating data sharing issues, and addressing privacy and reidentification risks, among others. As we had highlighted previously, such a broad mandate would create regulatory overlaps, make it difficult to formulate accountability mechanisms, and provide opportunities for "mission creep" where the regulatory powers exercised can exceed the intended ambit. 32

- Some functions assigned to the NPDA can also lead to conflicts of interest for example, as discussed in Section 3 (ii) above, the NPDA is put in charge of both appointing the data trustees for HVDs<sup>33</sup> as well as adjudicating their requests for data,<sup>34</sup> making the NPDA akin to both a licensor and a regulator. The role of the NPDA must be better thought out and more narrowly defined to avoid such conflicts.
- Regulatory overlaps: While the Committee states
  that the ambits of the proposed NPDA and
  other regulators such as the DPA, CCI, and legal
  frameworks such as copyright and trade secrets are
  separate, the proposed functions of the NPDA do
  in fact seem to contradict sectoral regulators and
  existing laws.<sup>35</sup> Additionally, it does not discuss any
  mechanisms to ensure regulatory cooperation.
- Lack of safeguards and ethical codes of procedure to protect the independence and autonomy of the regulator:
   The Revised Report does not discuss or specify any mechanisms to ensure transparency or accountability, or assess the performance of the NPDA. It also does not discuss safeguards against overbroad access to NPD by the Government.<sup>36</sup>

We therefore do not agree that the case for having a separate regulatory authority is made out. Instead of a regulator, focusing on creating a forum or other specific mechanism for various Ministries and regulators to coordinate their actions is essential.<sup>37</sup> This is especially important in an area such as data regulation, where there are multiple overlaps with critical sectors such as healthcare, transport, etc. Such frameworks already exist in other sectors in this regard, and they can provide useful case studies to learn from.<sup>38</sup>

<sup>30</sup> Pp. 10-12, Response to the Report by the Committee of Experts on non-personal data Governance framework, Esya Centre. accessible at: <a href="https://bit.ly/2NE7bjP">https://bit.ly/2NE7bjP</a>. 31 Pp. 14, 35, 20, Revised Report.

<sup>32</sup> P. 10, Response to the Report by the Committee of Experts on non-personal data Governance framework, Esya Centre. accessible at: <a href="https://bit.ly/2NE7bjP>.33">https://bit.ly/2NE7bjP>.33</a> S. 7.8, Pg. 19, Revised Report.

<sup>34</sup> S. 7.10, Pg. 20, Revised Report.

<sup>35</sup> P. 11, Response to the Report by the Committee of Experts on non-personal data Governance framework, Esya Centre. accessible at: <a href="https://bit.ly/2NE7bjP-.36">https://bit.ly/2NE7bjP-.36</a> S. 8, Pg. 23, Revised Report.

<sup>37</sup> See, for example, the UKRN, which brings together regulators in transport, utility and financial sectors. It acts as a forum for coordination and resolution of jurisdiction overlaps. More information available at <https://www.ukrn.org.uk/about/>. See also MoUs signed domestically by Australian authorities such as the Australian Prudential Regulation Authority and the Australian Energy Regulator, available at <https://www.apra.gov.au/memoranda-of-understanding-and-letters-of-arrangement> and <https://www.aer.gov.au/about-us/agreements-mous>.

<sup>38</sup> For example, the Financial Stability and Development Council, a non-statutory apex body created to maintain financial stability, develop the financial sector, and improve inter-regulatory cooperation. More information at <a href="https://dea.gov.in/sites/default/files/Gazette%20Notification%20December%2010.pdf">https://dea.gov.in/sites/default/files/Gazette%20Notification%20December%2010.pdf</a>. See also the Forum of Regulators, set up in the power sector to harmonise regulation in the sector, laying standards for performance of licensees, coordination, undertaking research, etc. More information available at <a href="http://www.forumofregulators.gov.in/About\_FOR.aspx">http://www.forumofregulators.gov.in/About\_FOR.aspx</a>.

This document has been prepared by the legal research team at the Esya Centre. For any further contact, please get in touch with us at:

Esya Centre A-281, Bhishma Pitamah Marg, Block A, Defence Colony, New Delhi -110049

www.esyacentre.org

