# Dimensional Analysis of Future of Non-Personal Data Sharing

*Examining Approaches and Governance Mechanisms*

CUTS International

#2111

# Dimensional Analysis of Future of Non-Personal Data Sharing

## *Examining Approaches and Governance Mechanisms*

Published by

CUTS International

**Consumer Unity & Trust Society**
D-217, Bhaskar Marg, Bani Park, Jaipur 302016, India
Tel: +91.141.2282821, Fax: +91.141.2282485
Email: cuts@cuts.org, Web site: www.cuts-international.org

Citation: (2021), 'Dimensional Analysis of Future of Non-Personal Data Sharing: Examining Approaches and Governance Mechanisms,' CUTS International, Jaipur, India

Authors: Shubhangi Heda, Assistant Policy Analyst and Setu Bandh Upadhyay, Research Associate, CUTS International.

This Research Report has been published as part two of the two-part CUTS project entitled, 'Examining the Rationale, Assumptions and Approaches to Non-Personal Data Sharing'. The Part one report may be accessed here.

#2111

# Contents

**Abbreviations** .......... **4**

**Acknowledgement** .......... **5**

**Introduction** .......... **6**

**Scope of Data** .......... **8**

Should we approach data sharing by focusing on NPD? .......... 9

High-Value Datasets .......... 12

Conclusion and the Way Forward .......... 17

**Stakeholders' Interactions and Governance Mechanisms** .......... **18**

Concerns with Stakeholder Categorisation Proposed by the Report .......... 19

Comparative Analysis of Stakeholder Interactions and Governance Goals .......... 21

Non-Personal Data Authority (NPDA) .......... 30

Conclusion and the Way Forward .......... 33

**Purpose of Sharing Data** .......... **34**

Conclusion and the Way Forward .......... 41

**Data Valuation and Incentive Mechanisms** .......... **43**

Variables at Play .......... 45

Approaches .......... 46

Allied Factors .......... 48

Conclusion and the Way Forward .......... 49

**Accountability and Consumer Rights** .......... **50**

Privacy .......... 51

Checks and Balances .......... 52

Grievance Redressal .......... 54

Conclusion and the Way Forward .......... 56

**Key Takeaways** .......... **58**

**Annexure - I** .......... **61**

# Abbreviations

| | |
|---|---|
| B2B | Business-to-Business |
| CoE | Committee of Experts |
| CCI | Competition Commission of India |
| DPA | Data Protection Authority |
| ESMI | Electicity Supply Monitoring Initiative |
| EU | European Union |
| GDPR | General Data Protection Regulation |
| HVD | High-Value Datasets |
| MOUs | Memorandum of Understandings |
| NPD | Non-Personal Data |
| NPDA | Non-Personal Data Authority |
| PDP Bill | Personal Data Protection Bill 2019 |
| PSD 2 | Public Sector Informations Directive |
| NPDGF | Non-Personal Data Governance Framework |
| ODI | Open Data Institute |
| UK | United Kingdom |
| US | United States |

# Acknowledgement

This study would not be possible without the efforts of various people in different from such as writing, research, encouragement, guidance, feedback, discussions, editing. Without the efforts of these people, this study would never have taken a comprehensive form.

First and foremost, we are thankful to all the stakeholders such as policy experts, industry representatives, academicians and members of the civil society who took out the time to participate in our consultations to enlighten us with their knowledge and insights, helping us create an impactful research product.

We are grateful to Pradeep S Mehta, Secretary General, CUTS International, Bipul Chatterjee, Executive Director, CUTS International, for their support and encouragement. We are immensely grateful to Amol Kulkarni, Director (Research) CUTS International, for his timely guidance and feedback throughout the study.

We would also like to thank Kapil Gupta, Assistant Policy Analyst, CUTS International, for his insights, contribution to research and involvement at the initial stages of the project.

We also appreciate the efforts of Madhuri Vasnani for editing, Rajkumar Trivedi and Mukesh Tyagi for preparing the layout of this Report. Akshay Sharma deserves a special mention for his contribution towards coordinating the information dissemination activities.

We express our sincere gratitude to all such individuals, whether or not named above, without whom the publication of this study would not have been possible and who helped us research this complicated and dynamically evolving area.

Finally, any error that may have remained is solely ours.

**CUTS Centre for Competition, Investment & Economic Regulation**

# Introduction

More and more countries are realising the growing importance of data , making "data sharing" a buzzword within the sphere of  digital economy policies across jurisdictions. The possibility of data sharing can include a multitude of stakeholders, it can be undertaken through varied purposes and can result in unique regulatory and economic implications. As the concept of data sharing develops in practice, it is important to note that 'data' as a resource presents unique challenges, specifically for the Indian digital economy, which has just started to deliberate on primary questions around data protection, data empowerment architecture and consent mechanisms.

The Committee of Experts (CoE) in India has attempted to present a governance and data sharing model for the Indian digital economy in a Non-Personal Data Governance Framework (the Report). The Report proposes data sharing for public interest purposes through data trustees.

In doing so, it has undertaken a challenging task to develop new categorisations of data, new kinds of intermediaries, and a new regulatory authority. Albeit its laudable efforts in giving impetus towards opening discussions and deliberations among stakeholders regarding the data economy, it also presents a range of complications. This necessitates closely examining the components of data sharing models and governance frameworks, so that economically and socially progressive ways of data use within the digital economy could be sustained.

To this end, this study presents an in-depth assessment of the approaches and recommendations stipulated in the Report. This has been done through **conducting an in-depth assessment through secondary research, comparative jurisdictional analysis (please refer to Annexure I) and conducting stakeholder consultations**. The rubric of the analysis to identify the parameters of assessment is inspired by the research agenda as proposed by Rene Abraham et al.[1] and has been used in research in assessing and identifying different data access and governance models.[2] It proposes for assessment of data governance models on the parameters of – governance mechanisms (data ownership, and allocation of decision-making authority), scope of data governance

[1] *Rene Abraham, Johannes Schneider, and Jan vom Brocke, "Data Governance: A Conceptual Framework, Structured Review, and Research Agenda," International Journal of Information Management 49 (December 2019): 424–38, doi:10.1016/j.ijinfomgt.2019.07.008.*

[2] *Marina Micheli et al., "Emerging Models of Data Governance in the Age of Datafication," Big Data & Society, September 1, 2020, doi:10.1177/2053951720948087.*

(application of governance mechanism on stakeholder, the scope of data, domains covered for data sharing), antecedents of data governance (impact of already existing relationships, facets, and practices) and consequences of data sharing (the purpose of data governance).

The above-mentioned **parameters have been modified to analyse the following dimensions -**

1. **Scope of Data** – This dimension covers the assessment of categories of data covered within the data governance and sharing framework proposed by the Report. It assesses the practicality of making categorisations between personal data and non-personal data (NPD) and sheds light on the complications, which may occur due to overlapping and unclear domains of data, and bypassing the proprietary interests in data.

2. **Stakeholder Interactions and Governance Mechanisms** – Within the data sharing chain, the way in which the governance model approaches stakeholder interests presents a crucial dimension in prescribing dynamics between them. This dimension highlights the challenges presented by the proposed framework to maintain synergies between stakeholders, and emphasises developing mechanisms of trust and collaboration through presenting an overview of varied data sharing and governance models adopted in different jurisdictions before legalising any data sharing approach through a regulatory mechanism.

3. **Purpose of Sharing** – This dimension illustrates challenges that may emerge in developing mechanisms to achieve the expected value creation due to - lack of a mechanism to identify problems statement, lack of transparency in the functioning of intermediaries such as data trustees, and lack of data equity in allocating benefits of data sharing.

4. **Data Valuation and Incentive Mechanisms** – Under this dimension, the assumptions and recommendations of the Report are scrutinised with respect to valuating the data as well as proposed incentive mechanisms to encourage data sharing. There is a comparison of approaches to data valuation to determine what approach fits best for data sharing, which are missing in the Report.

5. **Accountability and Consumer Rights** – This dimension explores and discusses various important aspects of accountability and consumer rights like privacy, grievance redressal, and checks and balances, which the report does not deliberate upon. The dimension also presents a comparison of accountability approaches from several jurisdictions, which can inform the building of more community and consumer-oriented approaches to data governance in the Indian context.

# 1

# Scope of Data

The Report specifically focuses on non-personal data (NPD) as the subject of its governance and sharing framework. In an attempt to provide clear distinctions between personal data and NPD, it states that any data, which is devoid of personally identifiable information will be covered within NPD, however, in cases where such data turns into identifiable data at any point it will be covered under the Personal Data Protection Bill 2019 (PDP Bill). It also proposed amendments to the PDP Bill to clearly demarcate the ambit of both frameworks, whilst keeping mixed datasets out of the scope of the proposed framework, wherein the personal data and NPD are intrinsically linked.[3]

Here, the primary concern is to ascertain - whether focusing on NPD data is a good starting point in developing an approach towards data-sharing? Moreover, apart from the concerns regarding blurred categorisation between personal and NPD and over-reliance on anonymisation techniques, there are also nuances to be navigated appropriately to understand the 'life cycle of data'.

This is crucial because different categories of data emerge depending on the way in which it is collected, stored, and used attracting differential treatment in terms of proprietary rights, level of sensitivity, and access and control mechanisms, which are important in understanding the scope of data and eventually in proposing an approach to data sharing.[4] Some of these concerns with respect to the Report are analysed below:

## Should we approach data sharing by focusing on NPD?

The Report, in an attempt to map out the typologies of data, gives examples of data that is collected by public and private entities through different instruments and whether it is available in the private or public domain. In this context, the premise of suggesting NPD as a starting point of data sharing rests on the assumption that there exists a practicable way, in which personal data and NPD can be separated. Apart from the inherent difficulty in creating bifurcation between personal and NPD through anonymisation, there is a further complication when different treatments are conferred to them in separate frameworks.[5]

While these distinctions are easy to understand in theory, due to the lucid nature of the data these categorisations may become overlapping and confusing. Within its entire 'life cycle',[6] data goes through different processes and the points where it stops being personal data and falls into the category of NPD can be difficult to deduce. From the stage of

[3] *Page 7-8 of the Report*

[4] *'Risks and Challenges of Data Access and Sharing | Enhancing Access to and Sharing of Data : Reconciling Risks and Benefits for Data Re-Use across Societies | OECD ILibrary' (OECD, 2019), https://doi.org/10.1787/276aaca8-en.*

[5] *Heda and Upadhyay. 'Navigating the Puzzle of Non-Personal Data Sharing: Three-Pronged Analysis of Rationale and Assumptions,' 2021 CUTS International, Jaipur, India*

[6] *Priyank Jain, Manasi Gyanchandani, and Nilay Khare, 'Big Data Privacy: A Technological Perspective and Review', Journal of Big Data 3, no. 1 (26 November 2016): 25, https://doi.org/10.1186/s40537-016-0059-y.*

gathering data to its analysis and usage, the characteristics of data changes very quickly – it can be anonymised; it can be combined with other datasets; more sensitive information could be attached to the data; the grouping and categorisation of data could also change depending upon the purpose of sharing.

The key point to note here is the way that these changes happen does not give the time to give different legal treatment to different categories of data. Additionally, studies and scholars have also cautioned that if the data is completely anonymised its quality can considerably degrade, which can impact the ultimate purpose of data sharing.[7] Thus, the applicability of anonymisation techniques presents legal and technical complications.

In this regard, an example of the 'data lifecycle' of the European statistical data illustrates that at the time of collecting, data is personal even if it is 'pseudonymised' and it falls under the ambit of General Data Protection Regulation (GDPR). Thereafter, to create meaningful deductions from such data, this pseudonymised data is linked to the different kinds of personal data to create a comprehendible dataset. After this, the key to identifiable components is destroyed. At this stage, it is not that all keys are destroyed, and some data is just anonymised. This still leaves open the issue of identification and also demonstrates the complication in determining the application of both frameworks (GDPR & Free Flow of NPD) in the life cycle of one dataset as there may still exist between pseudonymised and anonymised data.[8]

This gets complicated further when personal data and NPD are intrinsically linked such that they qualify to be a mixed dataset, as this presents an additional category that is subject to different legal treatment, and without specific standards to clearly distinguish this category there can be risks of overlaps.

Along with this, it also presents the risk of strategic behaviour from the industry, in which they would be more inclined to comply with the data protection law by stating they are operating with mixed datasets or non-anonymised datasets, which could hinder the data-sharing objective of the framework. These concerns regarding the scope of data are significant as the rights and liabilities flow from the categorisation of data.

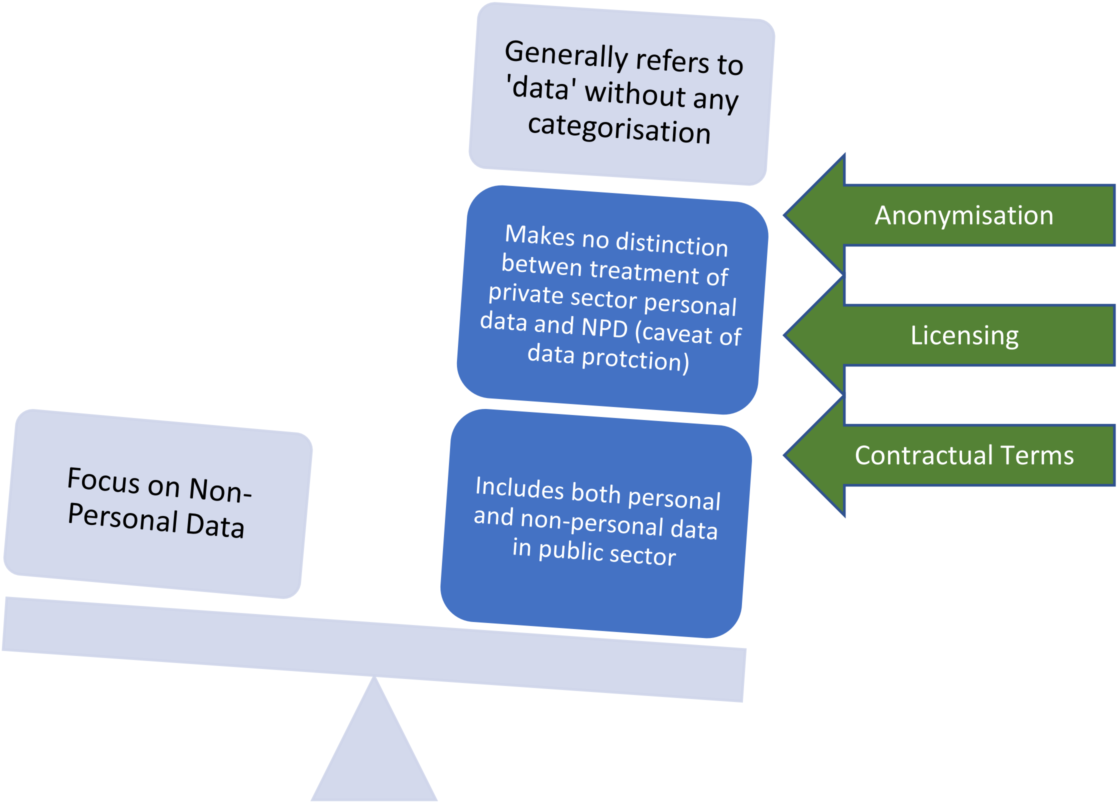
In this regard, our **comparative jurisdictional analysis** gives an interesting perspective in the context of data typologies being covered. Out of the 19 data sharing frameworks analysed (refer to Annexure I), except the European Union (EU) Framework on Free Flow of Non- Personal data and EU Open Data Directive, all other frameworks cover both

[7] *Regulating Non-Personal Data in the Age of Big Data, Health Data Privacy under the GDPR (Routledge, 2020), https://doi.org/10.4324/9780429022241-8.*

[8] *Inge Graef, Raphael Gellert, and Martin Husovec, 'Towards a Holistic Regulatory Approach for the European Data Economy: Why the Illusive Notion of Non-Personal Data Is Counterproductive to Data Innovation', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 27 September 2018), https://doi.org/10.2139/ssrn.3256189.*

personal and non-personal data, with the caveat of applicability of data protection principles. At the same time, it is also important to note that, most of the frameworks analysed already have established data protection laws, thus they rely on data protection principles while prescribing safeguards for sharing personal data. Across the spectrum, many data-sharing frameworks also rely on anonymisation techniques as a tool for safeguarding personal information. However, there has been criticism of the EU and United Kingdom (UK) Data Strategy for not addressing the nuances related to blurred lines between personal and NPD.[9]

**Figure 1: Author's Analysis of Approaches in other Jurisdictions**


Generally refers to 'data' without any categorisation
Makes no distinction betwen treatment of private sector personal data and NPD (caveat of data protction)
Includes both personal and non-personal data in public sector
Focus on Non-Personal Data
Anonymisation
Licensing
Contractual Terms

However, in this context, it is also important to acknowledge that these strategies have aimed to provide a broader approach that can then unpack into more nuanced legislation, and thus providing a holistic basis to move forward.

One of the guiding approaches in this regard is from the EU Framework on Free Flow of Non-Personal Data[10]. which the Indian Report has also attempted to consider. While defining NPD and its interface with personal data and mixed datasets, it highlights concerns regarding the efficacy of anonymisation techniques and states that adducing the level of anonymisation should be done on a case-to-case basis, depending on the kind of

[9] *Eline Chivot, 'EU Data Strategy Has Worthwhile Goal, But Misses the Mark', Center for Data Innovation (blog), 13 August 2020, https://datainnovation.org/2020/08/eu-data-strategy-has-worthwhile-goal-but-misses-the-mark/.*

[10] *https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018R1807*

dataset and the technique of anonymisation. Furthermore, it stipulates that in order to make this determination, all reasonable factors to establish identifiability within the dataset that can be applied by the aggregator, or any data controller should be accounted for in assessing the efficiency of the anonymisation technique.[11]

Thereafter, the EU Data Strategy covers both personal and NPD, however, it relies on GDPR and guidelines towards the treatment of mixed datasets to forming the categorisations where necessary.[12] This observation reflects on the underlying importance of forming proper anonymisation frameworks.

Many scholars and relevant researchers in the field have pointed out that anonymisation is a complicated procedure.[13] Reflecting from the learnings of other jurisdictions such as the EU, which is already facing complications with regard to anonymisation, the Report should prescribe a well-defined approach to anonymisation, rather than moving forward with establishing binaries that cannot function in practice.

Similar concerns were also highlighted within our **stakeholder consultation** with the experts, in which they stated that looking at data sharing through an ecosystem approach and then making categorisation such as high-value datasets and also design framework, can help in rectifying these blurred lines between personal and NPD. This warrants for re-assessing whether India should follow a holistic data-sharing approach, such that it could acknowledge the sensitivity and contexts in which the data exits rather than entrenching binaries of personal and NPD without first prescribing appropriate data protection and anonymisation techniques.

## High-Value Datasets

The kind of data covered within governance and data-sharing frameworks significantly impact the stakeholders which would have an interest in its usage. The Report has attempted to stipulate different categorisations of data through a matrix by giving examples,[14] however, it still envisages similar treatment to all those categories of NPD. No differentiation is made between data collected through overlapping entities such as

[11] *https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019DC0250&from=EN*

[12] *https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019DC0250&from=EN*

[13] *Nadezhda Purtova, 'The Law of Everything. Broad Concept of Personal Data and Future of EU Data Protection Law', Law, Innovation and Technology 10, no. 1 (2 January 2018): 40–81, https://doi.org/10.1080/17579961.2018.1452176, Michèle Finck and Frank Pallas, 'They Who Must Not Be Identified - Distinguishing Personal from Non-Personal Data Under the GDPR', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 1 October 2019), https://doi.org/10.2139/ssrn.3462948, Michèle Finck and Frank Pallas, 'They Who Must Not Be Identified—Distinguishing Personal from Non-Personal Data under the GDPR', International Data Privacy Law 10, no. 1 (1 February 2020): 11–36, https://doi.org/10.1093/idpl/ipz026, Sophie Stalla-Bourdillon and Alison Knight, 'Anonymous Data v. Personal Data — A False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 6 March 2017), https://papers.ssrn.com/abstract=2927945.*

[14] *Page 8-9, of the Report*

public-private partnerships, or data that may have a proprietary interest.[15] In order to bypass these overlapping and conflicting interests, the Report focuses on 'high-value datasets' (HVD), which refers to a part of the NPD dataset, which is of 'public interest'. Such data can be with both the private and public sectors.

Conceptualising the categorisation of HVD is based on 'public interest', which leads to vagueness and uncertainty as the meaning of 'public interest' cannot be made clear and this provides a tool, through which certain proprietary interests could be bypassed.[16] Additionally, the power of data trustees to make a decision on the dataset being HVD may lead to a conflict of interest and accountability issues relating to its usage.

This account of the Sidewalk Lab experiment highlights the complication that may arise in proposing a new category of data without mapping the data ecosystem and determination of proprietary interest in assessing control of such data. Thus, in the Indian context, a comprehensive approach that could identify interaction with the umbrella category of NPD along with the rights and limitations of varied stakeholders is required. At the same time, it will also be important to ascertain other categories of data that could be intrinsically linked with HVD, which may make the separate categorisation of HVD difficult.

[15] *'Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-Use across Societies' (OECD, 2019), ../sti-2019-1215-en/index.html.*

[16] *https://cuts-ccier.org/pdf/comments-on-revised_npd-governance-framework.pdf*

### Box 1: Sidewalk Labs - Toronto Project

One of the examples, which shows the significance of clearly determining the scope of data is the Sidewalk Labs experiment in Toronto. The experiment or pilot was undertaken to implement a Master Innovation Development Plan for a 'smart neighbourhood' in Toronto.[17]

The project created an 'urban data trust' for the governance of a special category of 'urban data'. To apply a 'commons' data governance framework 'urban data' was conceptualised as a commons resource to be available for the public interest. Geography was a critical factor that determined the nature of urban data, conceiving its existence independent of its collectors. It was argued that one of the motivations to do this was to bypass the legal barriers related to public and private ownership and personal and non-personal categorisation.

However, the problems arose when it was difficult to categorise the purely urban data from transaction data such as from utilities or ride-hailing cabs as it may be related to geographical context, but also related to an individual. Some clarifications were given in this regard through anonymising such transaction data, in which relevant or weighing the public interest in categorising it as urban data.

Despite this, it became increasingly difficult to determine where urban data ended, and the transaction data began. The most critical concern that was ignored in defining urban data was the proprietary interest in data and its eventual relation to the sensitivity of the data, this led to an over-inclusive definition of urban data, which ignored the nuances of public and private. Many of such issues eventually led to the closure of the project.[18]

Another framework proposed for a similar categorisation is the EU Open Data Directive[19] (refer to Annexure I). The framework has prescribed some thematic categories[20] for identifying HVD within the public sector to be shared free of cost, available for bulk download, and accessible in a machine-readable format. EU is still to roll out a comprehensive plan for the categorisation of these datasets. Even if thematic categories

[17] *Synced, 'Google's Sidewalk Labs Walks Away from Toronto Smart City Project', Medium, 7 May 2020, https://medium.com/syncedreview/googles-sidewalk-labs-walks-away-from-toronto-smart-city-project-d41393edf232.*

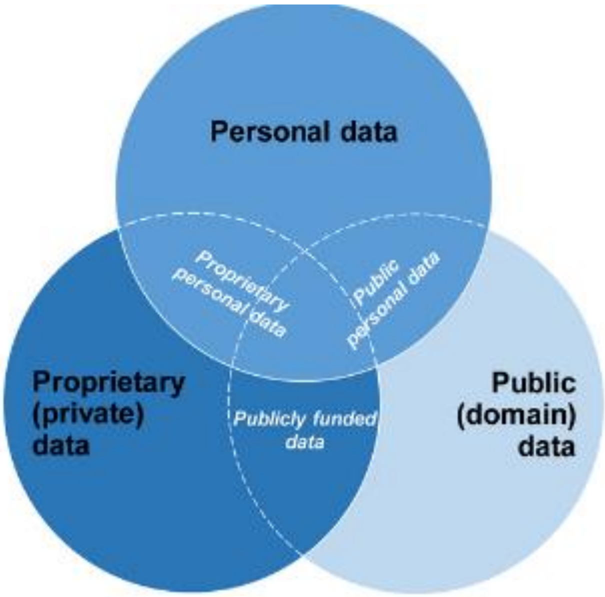
[18] *Teresa Scassa, 'Designing Data Governance for Data Sharing: Lessons from Sidewalk Toronto', 2020.*

[19] *https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024*

[20] *Geospatial, earth observation and environment, meteorological, Statistics, Companies and company ownership, Mobility*

are broad, it still provides sector-based categorisation rather than vague purpose-based definition, giving a direction to prepare eventual rules for HVDs.

**Figure 2: The personal, private, and public domains of data**



*Source: Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies | OECD iLibrary*[21]

On the other hand, an assessment conducted on the perspective of data providers on the HVD datasets proposed by the directive highlighted concerns such as - differences of opinion on the value that can be derived from datasets depending on geographical and sectoral impact; difficulty in assessing the ex-post impact of sharing such data; lack of clarity in prescribing roles and responsibility for specifying and maintaining such datasets; accounting of different local political, cultural and social condition in achieving consistency in the determination of HVDs.[22] Some of these issues are also relevant in the Indian context in prescribing meaning to HVDs and defining its various facets such that a sustainable categorisation could be maintained.

The **stakeholder consultation** with respect to data categorisation highlighted that the sector level sensitivities and the way, in which a particular sector is regulated are critical in determining whether a certain dataset is shareable. For instance, in the case of the 'power sector', which is highly regulated, with varied categories of data being collected at various points, there are already established rules for compliance for data management,

[21] *https://www.oecd-ilibrary.org//sites/276aaca8-en/1/2/2/index.html?itemId=/content/publication/276aaca8-en&_csp_=a1e9fa54d39998ecc1d83f19b8b0fc34&itemIGO=oecd&itemContentType=book#figure-d1e1423*

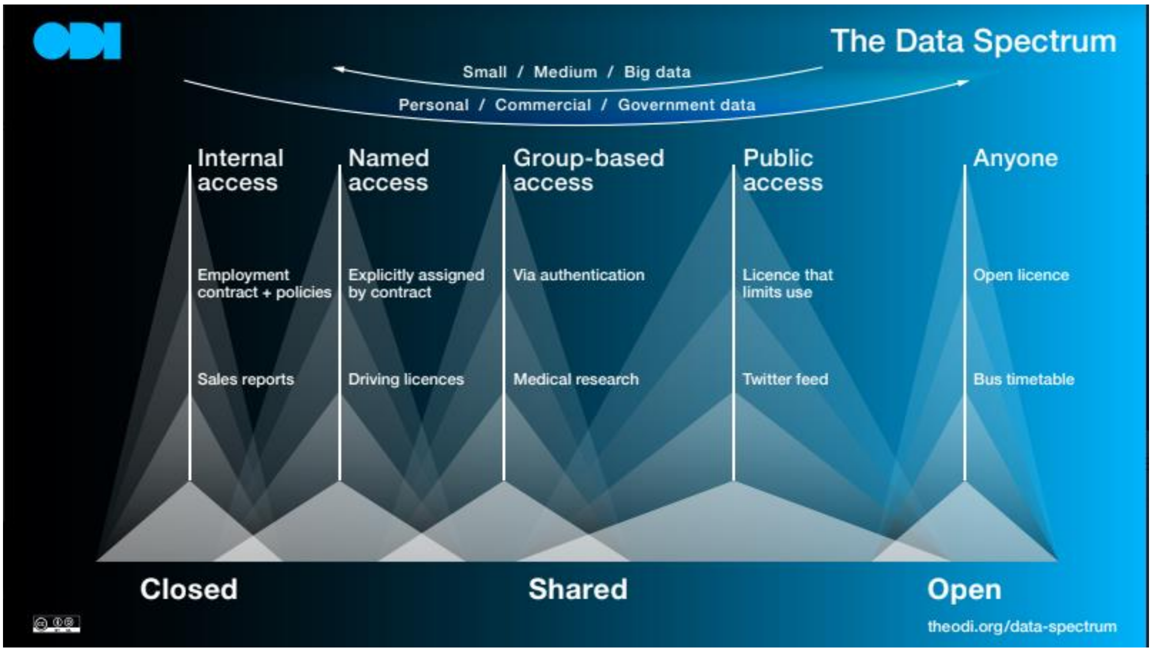
[22] *'High-Value Datasets: Understanding the Perspective of Data Providers.' (LU: European Data Portal., 2020), https://data.europa.eu/doi/10.2830/363773.*

thus sectoral experiences can inform the packaging of data. Notably, along with proposing a holistic data sharing framework for NPD, there are also sector levels sharing proposals such as the Draft National Geospatial Policy[23] and NITI Aayog recently proposed for setting up a central level energy dashboard, which should be given due consideration.[24] These include further categorisation of data based on sectoral definitions. At this point, there seems to be no indication in the Report on reconciliation with these emerging categorisations.

In light of the concerns presented above, the Open Data Institute (ODI) had proposed a data spectrum that could help in mapping the data ecosystem to understand the complication, which may arise in determining the data taxonomy for designing data sharing frameworks. It is important to note that any dataset includes all these categorisations, and thus it is pertinent to assess their level of identifiability, proprietary rights, and the purposes for which they could be useful. This also gives an indication to move away from the 'one size fits all' approach of HVD taken by the Report. Data spectrum could also be helpful in ascertaining sector-specific demarcation in categorising data.

A figurative illustration of the same is given below:

**Figure 3: Data Spectrum (Open Data Institute)[25]**



[23] *https://dst.gov.in/draft-national-geospatial-policy-2021-public-consultation*

[24] *http://www.businessworld.in/article/Niti-Aayog-Launches-India-Energy-Dashboards-Version-2-0/13-04-2021-386378/*

[25] *https://theodi.org/about-the-odi/the-data-spectrum/*

## Conclusion and the Way Forward

It is important to comprehensively understand the scope of data to determine various data typologies across the data value chain due to the lucidity of data; variations that may occur resulting from the way it is collected, the entity which collects it; and the purpose for which it is collected. While the Report has attempted to map various data typologies, however, through the above analysis we have highlighted some fundamental concerns that persist and have also given indication to some alternative approaches to mapping the scope of data.

Moreover, the Report places heavy reliance on anonymisation techniques in categorising personal and NPD, however, due to the complicated application of such techniques, approaching NPD as a separate category becomes difficult. In this context, it could be beneficial to take an ecosystem approach and map the data lifecycle to ascertain stages where anonymisation could be applied with the least risks.

Additionally, while it might seem like a good idea to propose a new category of HVD, a more important first step should be to understand whether such data can be separated from other datasets in actuality. Accordingly, stipulating different treatments to different kinds of data across the spectrum while being sensitive to proprietary and overlapping interests in data, will be beneficial.

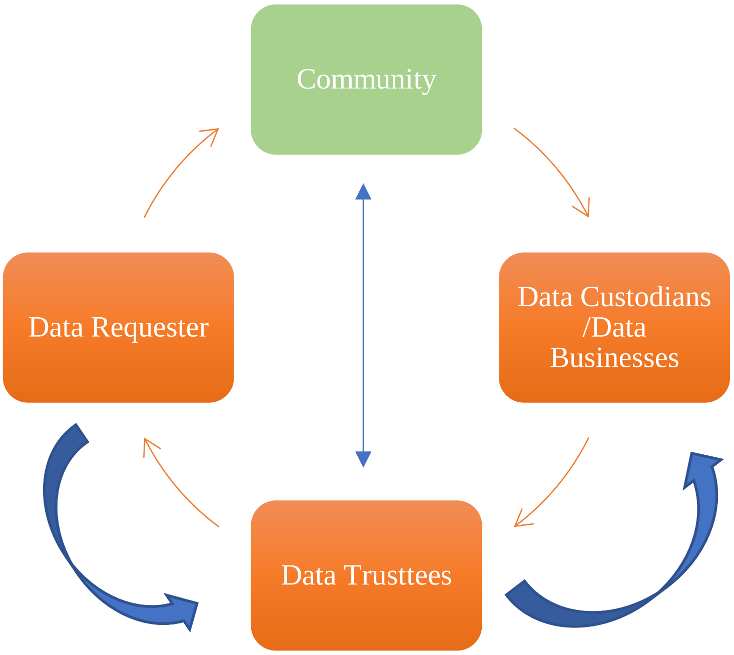
# 2

# Stakeholders' Interactions and Governance Mechanisms

The mapping of stakeholders and defining roles within the data-sharing ecosystem is closely related to the overall rationale of the data-sharing model and the governance mechanisms designed to achieve those rationales. Some of the critical questions in understanding the role of stakeholders in the data-sharing ecosystem are regarding access and controls they have over their data;[26] power dynamics that exist between the stakeholders; and the kind of representation that they have in the decision-making process.[27]

In this context, the Report recognises the following roles (as illustrated in Figure 4) in the data-sharing ecosystem and envisions the value of data to flow in a circular manner so that equitable distribution can be achieved amongst the stakeholders.

**Figure 4: Author's Analysis: The orange arrows in the figure represent the envisioned flow of value of data for different stakeholders and the blue arrows on the side represent the flow of data sharing requests**



## Concerns with Stakeholder Categorisation Proposed by the Report

While this envisioned flow and demarcation of stakeholders seems clear, there are issues with this framework with respect to factors that have been highlighted above and are further elaborated below-[28]

[26] *Marina Micheli et al., 'Emerging Models of Data Governance in the Age of Datafication':, Big Data & Society, 1 September 2020, https://doi.org/10.1177/2053951720948087.*

[27] *Linnet Taylor and Dennis Broeders, 'In the Name of Development: Power, Profit and the Datafication of the Global South', Geoforum 64 (1 August 2015): 229–37, https://doi.org/10.1016/j.geoforum.2015.07.002.*

[28] *https://cuts-ccier.org/pdf/comments-on-revised_npd-governance-framework.pdf*

a) **Community –** The Report has attempted to keep the interest of the community at the core of its vision, however, defining the 'community' as a stakeholder is itself complicated. It has been pointed out that there are limitations in just picking the jurisprudence with respect to natural resources and applying it to define community rights in data, due to the inherent lucidity and fundamental difference in the nature of data.[29] This is specifically true in India, where interests cannot be clearly demarcated in terms of geography, profession and business as communities are not institutionalised on these parameters.[30]

Moreover, the Report has envisaged that the community will achieve its representation in the data sharing chain, through the data trustees, however the mechanism to materialise this is missing, except for loosely stating that data trustees hold 'duty of care' to the community.

Additionally, as the PDP Bill aims to advance data protection rights for individuals, we are still to see the way, in which community rights will interact with these individual rights.

b) **Data Custodians and Data Businesses –** Data custodians include both public and private entities which collect, store and process data, with the exemption of data processors who are not involved in the collection of data.[31]

Data custodians hold stewardship responsibilities and have obligations to share data in the interest of the community upon the request of the data trustee. It is assumed by the Report that no additional incentives, apart from the processing charges are required to be given to the data custodians. In a way, they are imposed with mandatory data-sharing obligations without appropriate incentives. Here, the interest of data custodians seems to be asymmetric in terms of decision making and ensuring accountability regarding the eventual usage of their data. This is because data trustees have the right to decide on data requests without any reciprocating obligation of the data requestor towards them.

On the other hand, data businesses is a horizontal categorisation, which is required to mandatorily share the record of their metadata to the Non-Personal Data Authority. However, the purpose and threshold of information to be shared for the registration is not clear.

[29] *Puneeth Nagaraj, Varsha Rao, and Dedipyaman Shukla, 'Community Rights Over Non-Personal Data: Perspectives from Jurisprudence on Natural Resources', Data Governance Network, 2020, 27.*

[30] *Jyoti Panday, 'Tracking India's Approach to Data Governance: From Localization to Stewardship of Data', Internet Governance Project (blog), 9 February 2021, https://www.internetgovernance.org/2021/02/09/tracking-indias-approach-to-data-governance-from-localization-to-stewardship-of-data/.*

[31] *Page 17 of the Report*

c) **Data Trustees** – The Report has introduced a new kind of intermediary in the form of data trustees, which are either a Government organisation or a non-profit Private organisation (Section 8 company / Society / Trust).[32] They are an important piece in the chain as they hold the responsibility in all three ways – for representing the interest of the community, requesting data custodians for the HVDs, and making it available to the data requestor.

   Here, inadvertently data trustees hold the primary obligation towards the community, however, the Report fails to stipulate mechanisms to materialise these obligations. This is partly because the obligations of the data requestor to the data trustee are not clear. Additionally, the way in which the data trustee can ensure its independence from the government, or the data requester is also not clear.

d) **Data Requestor** – They hold the responsibility to eventually realise the public interest purpose.[33] However, these entities can be both public and private and may also hold business and strategic interests in using the HVD. The way, in which the alternate interest can be balanced with public interest purposes is not clear. Moreover, data trustees and data requestors can have overlapping interests, thus in such cases ensuring the independence of the data trustees may become problematic.[34]

## Comparative Analysis of Stakeholder Interactions and Governance Goals

The above illustrated framework has followed a unique approach in categorising stakeholders and prescribing their roles in data sharing. Through a broad comparative jurisdictional analysis of 19 data sharing frameworks, two broad categories of data flows are determined that is, business-to-business sharing and sharing by the government (public sector) to businesses and other individuals. Depending on the overall objective of the framework various techniques of governance has been identified, which have been illustrated below:

[32] *Page 19, of the Report*

[33] *Page 24, of the Report*

[34] *Chapter 3.1. Navigating the Puzzle of Non-Personal Data Sharing: Three-Pronged Analysis of Rationale and Assumptions. CUTS International. https://cuts-ccier.org/pdf/report-navigating-the-puzzle-of-npd-sharing.pdf*

**Table 1: Business to Business Sharing (also refer to Annexure I)**

| Data Sharing Frameworks | Mechanisms of Interaction | Governance Goals |
|---|---|---|
| EU Framework for Non-Personal Data | Self-regulatory codes of conducts | Transparency, interoperability, and open standards |
| Data Governance Act, EU Data Strategy | Data altruism and trusted data intermediaries to build sectoral data spaces | Building trust, voluntary data sharing, reducing administrative and compliance costs |
| GAIA-X, iSHARE, IDSA | Centralised and decentralised infrastructure, with technical standards | Secure and safe data sharing focusing on European principles of data protection, IPR, and cybersecurity |
| Singapore Trusted Data Sharing Framework | Bilateral, multilateral, and decentralised form of data sharing through agreed-upon standards between parties - related to the value of data, data quality, storage, and access to data | Developing trust among parties in data sharing through agreed-upon principles of trusted data sharing |
| Dutch, Japan, and EU Agricultural data sharing | Stipulating contractual standards with respect to licensing and disputes, technical standards | Remove ambiguities between parties related to rights and responsibilities |
| iSHARE | Certification mechanisms concerning parameters of security, storage, and adequate processing to receive necessary data, with an authority evaluation permits to validate data sharing | Utilising logistics data in a secure environment |
| UK Data Strategy and UK AI Trust Deal | Data Trusts and Data Stewardship | Balancing incentives and equitable distribution of data value across the economy |

| Data Sharing Frameworks | Mechanisms of Interaction | Governance Goals |
|---|---|---|
| FinDATA, EU sectoral framework on sharing of vehicular and power sector data | Establishing a governing authority, certification mechanism | To develop sharing mechanisms for personal data, keeping in mind the sensitivities of the sector |

**Table 2: Sharing from Public Sector to Individuals**

| Data Sharing Frameworks | Mechanisms of Interaction | Governance Goals |
|---|---|---|
| EU Open Data Directive, Data Governance Act | Setting different access points at various levels | Opening public sector data to stakeholders with caveats of protection of intellectual property rights and |
| Australia Data Release and Sharing Reform | Central authority of managing access | Release of data with appropriate safeguards and pre-defined purposes |

It is important to note that a large portion of the frameworks in Table 1 are voluntary in nature. One of the key differentiating factors in other frameworks that can be observed from **Table 1** compared to the Indian framework (also refer to Annexure I), is an attempt to ensure agility and flexibility in setting up interaction amongst the stakeholders while also balancing the objective of realising value from data. Even within the frameworks, which have prescribed using contractual terms, the interactions related to data usage, licensing, purposes, and security of data are given prime importance so that more clarity can be ensured amongst the parties involved in data exchange. Through setting up a clear contractual obligation, both the parties understand their part and incentives with the data sharing process to minimise asymmetries.

In a similar vein, in introducing the data altruism model in the proposed Data Governance Act, the lack of maturity of the data economy-related markets was accounted to prescribe mechanisms that could avoid burden on the stakeholders,[35] unlike the approach taken by the Indian Report. Along with this, it can be observed that sector-level interactions and

[35] *https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-data-governance-data-governance-act*

approaches are more centred around sector-level authorities who may have specialised knowledge.

Additionally, in the Indian context, state and central level dichotomies are also important to consider. Many of the states such as Karnataka[36] and Telangana[37] have state-level authorities, which are overlooking their open-data initiatives and are also encouraging data sharing in other sectors in the state. For these states, the Report proposes a parallel mechanism of data sharing, which needs reconciliation with existing mechanisms, especially in sectors that may fall under the state list. Notably, the recently introduced Draft National Geo-Spatial Policy[38] in India has attempted to form a reconciliation between state and central level initiatives through proposing to set up partnering agencies to facilitate the collection of geospatial data from state and Panchayati Raj institutions into a central data repository. This is because various states have existing repositories for geo-spatial data. The efficacy of these partnering agencies to harmonise data collection is still to be seen, however, it indicates the necessity of a policy approach that is sensitive to existing frameworks and sector-specific nuances in developing data sharing approaches.

Looking relatively at the Indian framework proposed by the Report, it seems to follow a unique approach that is centred around community interests in data forgoing sector and state-level sensitivities, which are more institutionalised. Therefore, the access, responsibility, and controls of the stakeholders are also attempted to be defined in that context. In other words, while the business-to-business sharing has been kept out of the scope of the Report, it has targeted data flows amongst private entities through data trustees legitimising data sharing within the broader ambit of public interest.

Furthermore, it is important to note that while the notion of data as a commons or common-pool resource is novel, the interest of different parties, such as the data custodian or data businesses may still stem from a commercial outlook. Thus, incentives, developing licensing practices, and building consensus on appropriate security standards as reflected in other frameworks are equally important to balance the commercial and public interests in data sharing.[39]

Apart from the initiatives highlighted above, other organisations and civic bodies have also come together to identify data sharing and governance strategies by taking a more community-oriented approach. While these initiatives are at a very nascent stage of conceptual theorising, it is useful to take a stock of resembling or similar approaches to

[36] *https://ceg.karnataka.gov.in/ksdc/public/english*

[37] *https://data.telangana.gov.in/policies*

[38] *https://dst.gov.in/draft-national-geospatial-policy-2021-public-consultation*

[39] *Heiko Richter and Peter R. Slowinski, 'The Data Sharing Economy: On the Emergence of New Intermediaries', IIC - International Review of Intellectual Property and Competition Law 50, no. 1 (1 January 2019): 4–29, https://doi.org/10.1007/s40319-018-00777-7.*

understand where the framework proposed by the Report fits relatively and draw learnings from the same.

**Table 3: Alternative Approaches to Data Governance**

| Approach | Initiative | Role of Community | Learnings |
|---|---|---|---|
| **Data Commons –** It is an approach through which data access and control could be democratised through sharing data as a common pool resource. Here, the citizens or community directly participate in making the decisions about their data.[40] | A pilot was conducted at various complementary levels to test the development of data commons in Barcelona at the city level.[41] It led to the creation of Barcelona Now[42], a portal that hosts 21 datasets related to various parameters, donated by the citizen or sources from the municipality data. | The pilot relied on improving the existing community on the platform called decidem. Decidem is an online participatory environment, through which citizens can sign the petition and deliberate on different issues at a city level.[43] However, to integrate the community, different manifestos were opened for voting which elaborated on privacy rights, usage of data, and control that would be provided to the participants of the community. Only after such consultation, the parameters of data | It is important to highlight here that the community was recognised at a city level. It is also a bottom-up approach in which, the first step was to build citizen consensus towards data politics and governance. For this, the questions such as what data they want to share; what they don't want to share; what they want to anonymise, were determined by the members of the community.<br>Notably, the pilot also envisaged the adoption of 'digital sovereignty' in a way that data could be treated as infrastructure and shared as a public good, which is also |

[40] *'What Does It Mean? | Shifting Power Through Data Governance', Mozilla Foundation, accessed 3 May 2021, https://foundation.mozilla.org/en/data-futures-lab/data-for-empowerment/shifting-power-through-data-governance/.,Stuart Mills, 'Who Owns the Future? Data Trusts, Data Commons, and the Future of Data Ownership', SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, 24 September 2019), https://doi.org/10.2139/ssrn.3437936.*

[41] *https://decodeproject.eu/publications/final-report-barcelona-pilots-evaluations-barcelonanow-and-sustainability-plans*

[42] *http://bcnnow.decodeproject.eu/dashboard.html*

[43] *https://www.decidim.barcelona/*

| Approach | Initiative | Role of Community | Learnings |
|---|---|---|---|
| | | control are determined. | similar to what the Indian framework has proposed. However, the procurement of such data is limited to the data accumulated by public authorities at the city level and does not extend to proprietary private data. |
| **Data Collaborative** - Here the focus is on pooling proprietary data through an arrangement between parties for the larger benefit of the society. Here, a group of such data holders appoints an independent authority to manage such data.[44]<br>Pioneering efforts to explore and foster this approach have | Big Data for Social Good is an initiative taken by the GSMA under which Bharti Airtel and Be Healthy (an initiative by WHO and ITU) contributed mobile data to assess high-risk locations of Tuberculosis infection in the Indian states of Uttar Pradesh and Gujarat.[46] | Here, the community is the ultimate beneficiary, however, it is not involved in the arrangement of data sharing. The private sector collaborates for relevant causes. | The highlight here is the efficient matching of demand and supply-side factors such that collaboration involved adequate data and expertise required to use that adequately. Moreover, the data sharing, in this case, was voluntary and on an aggregate level.<br>However, there are also concerns raised in adopting this approach related to privacy and consent of the data principals in the use of the data. Along with this, collaboration may often involve parties with diverse interests thus, trust-building between parties |

[44] *Iryna Susha et al., 'A Research Roadmap to Advance Data Collaboratives Practice as a Novel Research Direction':, International Journal of Electronic Government Research 14, no. 3 (July 2018): 1–11, https://doi.org/10.4018/IJEGR.2018070101,*

[46] *https://aiforimpacttoolkit.gsma.com/resources/Big-Data-for-Social-Good_Airtel-INDIA_TB_Case_Study.pdf*

| Approach | Initiative | Role of Community | Learnings |
|---|---|---|---|
| been taken by GovLab.[45] | | | through contractual rules or other safeguards also becomes crucial.[47] |
| **Data Trusts –** It is based on a legal relationship that is formed between an individual or a group and the trustee to act as a steward of the data. Here, the trustee is responsible to negotiate in good faith the interest of the beneficiaries of that trust.[48] | The framework being proposed by the United Kingdom (UK) Data Strategy and the UK AI Trusts Deal also rely on the data trusts and stewardship models of governance and is relatively closer to the model proposed by the Report. In the UK, the Open Data Institute (ODI) has been actively involved in analysing intricacies in the functioning of data trusts and has initiated data trusts pilots in different sectors to draw learning lessons.[49] | Data trusts in the form of civic data trusts, in which citizens or a group appoints a data trustee for a particular objective involves community participation. For example, in the pilot related to collected data about food wastage, the customers came together to set up a data trust to procure such data from appropriate data holders.<br>However, within other arrangements data trusts may exist in which technically the community may just exist in the beneficiary capacity, however in that case the objective of data trusts should be | The recommendation and learning coming out of the data trust pilot indicate that incentives amongst organisations to set up a data trust to steward their data stem from their incapacity to make data available in the best possible manner. This indicates that data trustees themselves have to be capable. Along with this, while exploring the possibility of mandating data trusts, it was recognised that this should be done on case-to-case basis wherein other authorities such as the competition authorities can adjudicate on such requirements. The maturity of the data ecosystem, the independence of the |

[45] *https://datacollaboratives.org/explorer.html*

[47] *A. J. Klievink, H. G. van der Voort, and W. W. Veeneman, 'Creating Value through Data Collaboratives: Balancing Innovation and Control', Information Polity 23, no. 4 (2018), https://doi.org/10.3233/IP-180070.*

[48] *https://foundation.mozilla.org/en/data-futures-lab/data-for-empowerment/shifting-power-through-data-governance/*

[49] *https://docs.google.com/document/d/118RqyUAWP3WIyyCO4iLUT3oOobnYJGibEhspr2v87jg/edit#heading=h.8c4lvfdze3uy*

| Approach | Initiative | Role of Community | Learnings |
|---|---|---|---|
| | | working towards the benefit of the community. | data trustee, and building trust in communities through certification and disclosures about the functioning of data trustees are also important. |

Holistically looking at the above-mentioned approaches it appears that the community is either a beneficiary or a trigger for data sharing (data commons and civic data trusts). In contrasting these approaches with the one taken by the Report, it seems it has some of the features from all three approaches to form a unique amalgamation, however, also missing some of the central learnings from all these approaches. For instance, it borrows from the data common framework, but the true form of democratic governance and community participation is missing. While the Report has indicated some loosely tied criteria such as getting the expression of interest from the minimum unspecified number of community members and public consultation to map contours of HVD, however, the form and the way in which these are to be implemented to garner meaningful representation is still to be formulated, leaving much scope for deliberation.

Moreover, it has relied on the data trustees' model for governing the exchange of data, however, here also no differentiation is made between situations of mandated data trust and voluntary data trusts. It has also not prescribed the way, in which these trusts could ensure their independence and sustenance to manage the data. A key concept that is being borrowed from the data collaborative model is an underlying assumption that the private sector would see an opportunity in sharing data with other stakeholders, however, here again, the critical mechanism of ensuring trust between parties through contractual standards and laying out the purpose of sharing is again elusive.

The meaning of terms such as 'stewardship responsibilities' and the 'duty of care' is still evolving and remains unclear at this stage. It is also important to note that these concerns have not been addressed in the existing data sharing initiatives in India, including the NDSAP, exhibiting a need for reform in the national approach to data sharing.

Coming back to the concern related to defining community and its representation, in **our stakeholder consultations** it was highlighted that while defining community could be difficult, the starting point should be to identify existing categorisation of demographics and their representation. Along with this, it was also emphasised to closely examine the data lifecycle to understand the linkages between community and benefits intended to be derived. The lifecycle of data presents points where the value of data changes and points where the community or collective value of data becomes more concrete, which could be used to identify the kinds of communities that may have an interest in the dataset.

Alternatively, the concept of Demographic Identification Information (DII)[50], could also be used in identifying collective interests in data. While in the existing scholarship such information has been contextualised to address the data points which may lead to discrimination, however, contextualising data in the form of DII could help in identifying points of commonality, indicating parameters of community identification. At the same time, it is important to note that the first and foremost to any of this is to first unpack the lifecycle of data.

Furthermore, it is important to assess the existing institutional capacity and improve it so that authorities are equipped to understand the data usage in their community and can fulfil the requirement of adequate representation. **Stakeholder recommendations** on this aspect pressed on forming an operational relationship between the political representation and domain experts in data. Through this relationship, a co-design framework could be formulated to identify data trustees, which is also reflected in the learning from the UK pilots.[51]

Some of the learnings in this regard could also be taken from parameters that lead to a sustained data collaborative model, such as building contractual standards, recognising mutual goals, and developing more trust. It was also suggested that more representation could be given to the community through better consultation and open decision-making process and provide a forum for community members to engage in the decision-making of the data trusts.

Currently in India as well, data trusts pilot related to urban mobility data is being undertaken in Delhi. While the implementation of the pilot is underway, the governance mechanism directing the pilot focuses on understanding data stewardship through the data trusts framework in a way that data holders, community and governance intermediaries could come together, such that access to data could be more democratised.[52]

[50] *Lanah Kammourieh et al., 'Group Privacy in the Age of Big Data', in Group Privacy: New Challenges of Data Technologies, ed. Linnet Taylor, Luciano Floridi, and Bart van der Sloot, Philosophical Studies Series (Cham: Springer International Publishing, 2017), 37–66, https://doi.org/10.1007/978-3-319-46608-8_3.*

[51] *https://docs.google.com/document/d/118RqyUAWP3WIyyCO4iLUT3oOobnYJGibEhspr2v87jg/edit#heading=h.8c4lvfdze3uy*

[52] *https://vidhilegalpolicy.in/research/data-stewardship-for-non-personal-data-in-india/*

We are still to see how these principles would be implemented in practice, however, such pilots seem to be a good starting point to understand the data ecosystem itself, which is also critical as highlighted in the UK data trusts pilots.[53]

Another issue is integrating and balancing the interest of the data custodians effectively within the approach proposed by the Report. For this, the **experts suggested** leading by example through conducting pilots and validating some of the assumptions around data trustees. To this end, conducting pilots and establishing stewardship and data trusts model could help data custodians and requesters develop more trust in the process. On the other hand, this will also lead to the evolution of principles and models of governance, through which the current notion of data ownership could be modified. Therefore, setting up bottom-up data trusts is sort of an ideal objective; however, this may start from data marketplaces or data stores.

## Non-Personal Data Authority (NPDA)

The report prescribes for the exclusive jurisdiction of the NPDA on the basis that the objective of the authority is to adjudicate on the rights of the community and provide initial support to the startups and perform both enabling and enforcing functions. It is also stated that NPDA would be created with industry participation and its function will be harmonised with other regulators such as the Data Protection Authority (DPA) under the PDP Bill. While the objective of setting up the authority can itself be questioned as it presents concerns of regulatory overlaps with the Competition Commission of India (CCI) to ensure equitable distribution of data and data protection, which is the responsibility of the DPA[54], however, the way in which the new authority is envisaged is also problematic.[55]

The process, in which a regulator is set up plays a crucial role in determining its functional and financial independence and accountability. While these two factors are important, due to the dynamic and rapid growth of the data economy, the Report recognises that the expertise of the NPDA to predict the changes within the data economy and accommodating industry interests will also be critical. Considering this, the regulatory bodies and policy makers should have proper interactions, so that appropriate expertise

[53] *https://docs.google.com/document/d/118RqyUAWP3WIyyCO4iLUT3oOobnYJGibEhspr2v87jg/edit#heading=h.8c4lvfdze3uy*

[54] *CUTS Comments on Revised Report of Committee of Experts on Non-Personal Data Governance Framework, https://cuts-ccier.org/pdf/comments-on-revised_npd-governance-framework.pdf*

[55] *Page 20 of the Report*

can be introduced, at the same time it should not just work as an extension of the ministry.[56]

However, the mechanisms through which independence could be achieved within the regulatory process are missing. In this context, the Australian Data Release and Legislative Reform also indicated towards unique approach, through which it proposed to set up a National Data Commissioner, which would be supported by the National Data Advisory Council to apprise the National Data Commissioner of "*ethical data use, community expectations, technical best practice and industry, and international developments*."[57]

Notably, within this framework, adequate importance has been given to the independence of the National Data Commissioner. It has also been prescribed for taking a graduated enforcement approach. Through this approach, it is ensured that binding rules are only prescribed to priority areas such as protecting privacy and for other domains, non-binding rules in the form of guidance could be given (also see Annexure I). In this way, regulatory overreach and compliance burden could be accommodated according to the industry readiness.

Furthermore, the constitution of the selection committee, which would be responsible for choosing members of the NPDA is not stipulated by the Report. While it is stated that industry participants will be included, it is also important to include civil society, experts from academia and policy think tanks, and consumer organisations. This is necessary because the objective of achieving public interest is closely tied with adequate representation of the community as well as the industry who would be both the supplier and consumer of the data.[58]

Additionally, the participation of the sectoral regulators or professional bodies is also necessary as the nuances within data management could differ depending upon sector level needs.

Another important factor to consider is the accessibility of the new regulator. The NPDA has an enabling function to maintain a meta-data repository and institutionalise the data-sharing model, however, this would require a sound and easily accessible technological architecture. For this, the Report proposes a technical architecture, however, without

[56] *Vijay Vir Singh and Siddhartha Mitra, 'Regulatory Management and Reform in India' (OECD, 2010), https://www.oecd.org/gov/regulatory-policy/44925979.pdf.*

[57] *https://www.datacommissioner.gov.au/sites/default/files/2019-09/Data%20Sharing%20and%20Release%20Legislative%20Reforms%20Discussion%20Paper%20-%20Accessibility.pdf*

[58] *http://www.cuts-ccier.org/pdf/CUTS_Comments_on_the_draft_Regulatory_Reform_Bill-2013.pd*

assessment of the existing capacity of data management and prescribed standards NPDA would not be able to govern this technical infrastructure.

Along with this, whilst NPDA also has an enforcing function to preserve the privacy of the community, it does not provide any provision for setting up a grievance redressal mechanism. The comparative jurisdictional analysis in this regard indicates taking a principle-based approach that can guide the setting up of technical architecture through prescribing data standards on the lines of findability, accessibility, interoperability and reusability.

Additionally, as indicated in the case of Findata initiative[59] the regulatory bodies should also have mechanisms that could assist the stakeholders in understanding the data-sharing model and help them navigate legal and procedural complications (also see Annexure I).

Furthermore, a collaborative approach is required to deal with regulatory overlaps. Economic regulators have long used this model, the UK being the primary example**.** An authority or a body comprising of all the concerned bodies and regulators (in this case the DPA, CCI, and sector regulators) can be formed to decide and adjudicate on the separation and limitation of each of their jurisdictions. This authority can also be empowered to resolve matters, which cannot fall under any one of the regulations.[60] Memorandum of Understandings (MOUs) is another tool for such regulatory collaboration.

[59] *https://findata.fi/en/*

[60] *Collaboration between Economic Regulators: Options for embedding joint working between economic regulators - government response to the consultation (publishing.service.gov.uk)*

## Conclusion and the Way Forward

The analysis presented above points to two critical aspects in designing data sharing - the role of stakeholders and governance mechanisms that stipulate their interactions. The concern with the framework proposed by the Report is related to achieving representation, balancing the value different stakeholders have in the data value chain, and ensuring trust and impetus amongst the parties to take up data sharing.

Different approaches for achieving these objectives have developed as highlighted from the analysis of different jurisdictions. However, pinpointing any one appropriate approach is very difficult in the Indian context because of the lack of knowledge that exists about the data ecosystem and the divergent interest that exists in its value. At this point, the primary focus should be on identifying principles of governance that are important in data space and then identifying tangible mechanisms in the form of consultation or co-design, through an evidence-based approach.

One of the key parameters should be to explore mechanisms, through which the application of bottom-up data trusts could envision community representation, independence, accountability, and transparency in their functioning could be ensured.[61] Equally necessary is unpacking these parameters in the context of data to understand what representation actually means for communities in India and how do they think they should have control of their data. This kind of unpacking could be seen in the implementation of the DECODE project in Barcelona. Thus, deriving a new form of institutions may be an eventual process, but this should go beyond just the vaguely identifying 'duty of care' or 'stewardship' or 'public interest'.

[61] *Sylvie Delacroix and Neil D Lawrence, 'Bottom-up Data Trusts: Disturbing the "One Size Fits All" Approach to Data Governance', International Data Privacy Law 9, no. 4 (1 November 2019): 236–52, https://doi.org/10.1093/idpl/ipz014.*

# 3

# Purpose of Sharing Data

Ascertaining the purpose or rather the expected value creation from data sharing is critical to determine the acceptance, expectation, and motivation of stakeholders to indulge in data sharing. In this regard, in order to ensure that data is shared for "socially progressive objectives" and there is equitable distribution of benefits,[62] questions such as - what stakeholders would be benefited from such sharing, whether it will benefit public or private interests, what sectors of the economy will it effect, will it solve societal issues become critical.

The Report stipulates that the data could be shared for public interest[63] and sovereign purposes.[64] In defining public interest or public good purposes the Report gives an open-ended list of examples, which makes its scope broad and vague. While this in itself is a pertinent issue, for the current analysis, our focus would be on investigating approaches and mechanism, through which data sharing model intend to achieve these purposes. In this context, the Report makes the data trustees responsible to determine whether there exists a legitimate purpose of sharing, however, there seem to be missing links between the purpose of sharing and mechanisms to achieve that purpose as highlighted below –

1) **Streamlining process of data-sharing** – The Report states that the data trustees must ensure that the dataset (HVD) meant to be shared should be used for public interest purposes. For this, the Report indicates that the NPDA will form guidelines regarding the intended objective and impact of HVDs. This leaves much room for uncertainty and deliberation. A key starting point to determine the objective and impact should be to ascertain ways to identify a problem statement, then to identify necessary data to address that problem and ensure the technical feasibility and develop mechanisms trust can be ensured between parties in data sharing. This process is important, to match the supply and demand-side factors.

   For example, there can be a situation where a combination of multiple datasets is required to fulfil a public interest objective, however, without actually determining a clear problem statement indicating the need for multiple datasets and related feasibility of sharing every single dataset, the objective would not be achieved. Equally important is the management of this process, where a single entity with appropriate expertise should be able to handle the smooth flow of data.[65]

   Thus, within the Indian context, the ways to identify particular problem statements and associated datasets; resolving conflicts amongst rights to a particular dataset

[62] *Marina Micheli et al., 'Emerging Models of Data Governance in the Age of Datafication':, Big Data & Society, 1 September 2020, https://doi.org/10.1177/2053951720948087.*

[63] *Page 23 of the Report*

[64] *Page 24 of the Report*

[65] *'Towards a European Strategy Onbusiness-to-Governmentdata Sharing for the Public Interest' Final Report Prepared by the High-Level Expert Group on Business-to-Government Data Sharing' (European Commission, 2020), https://www.euractiv.com/wp-content/uploads/sites/2/2020/02/B2GDataSharingExpertGroupReport-1.pdf.*

between the data trustees, or dealing with the regulatory and technical complications for smooth processing of data sharing requests are vital.

2) **Data Stewardship** – Another mechanism that the Report relies on to achieve the public interest purpose is through data stewardship responsibility, which it has stipulated for both data trustees and data custodians. Data stewardship is a model through which intermediaries take on the responsibilities on the behalf of the users or the communities to govern and manage data flow such that the data is available for the public good. This responsibility could be recognised at different levels and in varied forms.[66]

   The effective implementation of data stewardship to ensure achieving the purpose of data sharing amongst entities relies on trust and assurance of the parties that the data will be used for the intended objective without any intentional over-spill.[67] In building this trust it is also important to have proper accountability mechanisms for each of the parties to ensure appropriate data usage.

   Some of these concerns have been rightly pointed out in analysing stewardship responsibilities in data collaborative arrangements as different organisations involved in sharing data may belong to different sectors and have different ways, in which they treat and use data. In such situations, it would be important for them to come to a consensus on basic parameters of data usage, security, data sharing beyond data requestor, storage mechanisms, and most importantly to build a shared understanding of the purpose of sharing.[68] Thus, the way in which stewards collaborate, act and protect to achieve the purpose of sharing is crucial.[69]

   However, while the Indian report stipulates such responsibilities, but the parameters to sustain it are missing, for example, the Report stipulates for some obligations pertaining to data trustees and data custodians, no responsibility concerning appropriate usage has been stipulated for the data requestor. Moreover, the mechanisms to enable a sense of trust regarding the purpose of data sharing is also lacking.

3) **Data equity** – One of the larger aims of the Report is equitable distribution of data, through ensuring the protection of community interest in data sharing. For this, the power dynamics between the state, data providers, requestors and community is

[66] *Sidharth Manohar, Astha Kapoor, and Aditi Ramesh, 'Understanding Data Stewardship: Taxonomy and Use Cases' (Aapti Institute, 2020), https://thedataeconomylab.com/2020/06/24/data-stewardship-a-taxonomy/*

[67] *'Data Sharing and the Public Interest in a Digital Pandemic*', Verfassungsblog (blog), accessed 11 May 2021, https://verfassungsblog.de/data-sharing-and-the-public-interest-in-a-digital-pandemic/.*

[68] *Iryna Susha and J. Ramon Gil-Garcia, 'A Collaborative Governance Approach to Partnerships Addressing Public Problems with Private Data', 2019, https://doi.org/10.24251/HICSS.2019.350.*

[69] *'(Re-)Defining the Roles and Responsibilities of Data Stewards For An Age Of Data Collaboration' (GOVLAB, 2020), https://www.thegovlab.org/static/files/publications/wanted-data-stewards.pdf.*

important to ensure that the benefits culling out of the data-sharing achieve a larger public benefit with minimum exclusions.

Within the governance model proposed by the Report exclusions may occur in the following forms - "*(a) exclusion from (personal or other) data and information entering a digital common; (b) exclusion of people from using data and information held in the digital commons; (c) exclusion of people from benefitting from the digital commons (both data and infrastructure).*"[70]

Along with this, even the small and medium enterprises, which are envisioned to be able to tap into the potential of data will be influenced by the way in which the HVD is packaged and made available to them. It has to be ensured that the data is interoperable, and they are motivated to use it toward the public interest objective. In this context, it is important for the Report to recognise that equity is multi-faceted and would require appropriately defining the scope of data, governance mechanism, and having redressal mechanisms in case of exclusions. However, this also calls for ex-post mechanisms to measure the way. in which objectives are achieved and discrepancies thereof. This would be crucial in creating a feedback loop for the entire data-sharing chain.

Considering these factors and observations, it would be beneficial to look at data sharing ecosystems that stood out in our **comparative jurisdictional analysis** (also refer to Annexure I) as they provided a unique perspective on the mechanism to realise their respective purpose of data sharing. Some of these cases are mentioned below:

1) **Findata**[71] – Finland passed the Act on the Secondary Use of Health and Social Data in 2019, to ensure more accessibility in health data in a secured ecosystem. For this, the act prescribes for establishing a permit granting authority that is, Findata. This permit authority is responsible for managing data requests; processing and granting data permit requests; aggregation, anonymisation, and pseudonymisation of data. Thus, it provides one-stop governance of health-data sharing. These functions are facilitated by a portal that gives information to the potential data requestors on the permit process and data description in the form of metadata. Notably, each data request has to specify the purpose of use, the contents of the data required, the time span for which data will be used, and for doing so the authority has provision for assisting the requestors in the procedure.[72]

[70] *Barbara Prainsack, 'Logged out: Ownership, Exclusion and Public Value in the Digital Data and Information Commons', Big Data & Society 6, no. 1 (1 January 2019): 2053951719829773, https://doi.org/10.1177/2053951719829773.*

[71] *https://findata.fi/en/*

[72] *'A Finnish Model For The Secure And Effective Use Of Data', Sitra (blog), accessed 16 May 2021, https://www.sitra.fi/en/publications/a-finnish-model-for-the-secure-and-effective-use-of-data/.*

This ecosystem checks out on relevant factors necessary to achieve its intended purpose - as it provides for a mechanism, through which the purpose of sharing is to be made explicit and cross-checked with a centralised governing authority, which has adequate expertise in the sector. Along with this, the model aims to ensure accessibility through establishing an online portal and offering assistance and legal counselling in case of security and other related complications.

Moreover, in assessing the performance of this model, acquiring public trust and consumer centricity were considered important facets to justify data-sharing for societal benefit. It also highlighted some important factors such as - specifying who can use the data and for what purpose and being sensitive to different interpretations that can be accorded to purposes depending on different sectors or stakeholder interests should be adequately considered to develop a successful data sharing ecosystem.[73]

2) **Lessons from the UK Data Trust Pilots** – As part of the UK AI Trust Deal and the UK Data strategy, pilots of the data trust model were undertaken in different sectors. The learnings from the pilots specifically indicate that in a case where intermediaries such as data trustees have involved the mission statement or the purpose of the trust should be clearly specified. This should also include the geographical limitations, kind of the data that the trust could hold, and its targeted beneficiaries.

   Additionally, it was recognised that the purposes of data usage can evolve, however, data trusts should have a flexible mechanism in place, through which a shared understanding of the purpose of sharing could be developed through a collaborative effort. We must be cautioned from the Sidewalk Lab data trust experiment, which lacked this collaborative and shared understanding of the usage of data, leading to community distrust in the project resulting in its failure.[74]

   Thus, the lessons from the pilots recommend that the purpose of establishing data trust must adequately reflect the purpose of sharing, which should be formalised within the governance mechanisms of the data trusts.[75] In the Indian context, with varied community interests and risks of overlaps, a clear understanding of the purpose of data trust and the community that they aim to serve is pertinent.

3) **Australian Data Sharing and Release Reform** – This framework proposes for sharing of public sector data for specified public interest purposes in the form of

[73] *A Finnish Model For The Secure And Effective Use Of Data', Sitra (blog), accessed 16 May 2021, https://www.sitra.fi/en/publications/a-finnish-model-for-the-secure-and-effective-use-of-data/*

[74] *Teresa Scassa, 'Designing Data Governance for Data Sharing: Lessons from Sidewalk Toronto', 2020.*

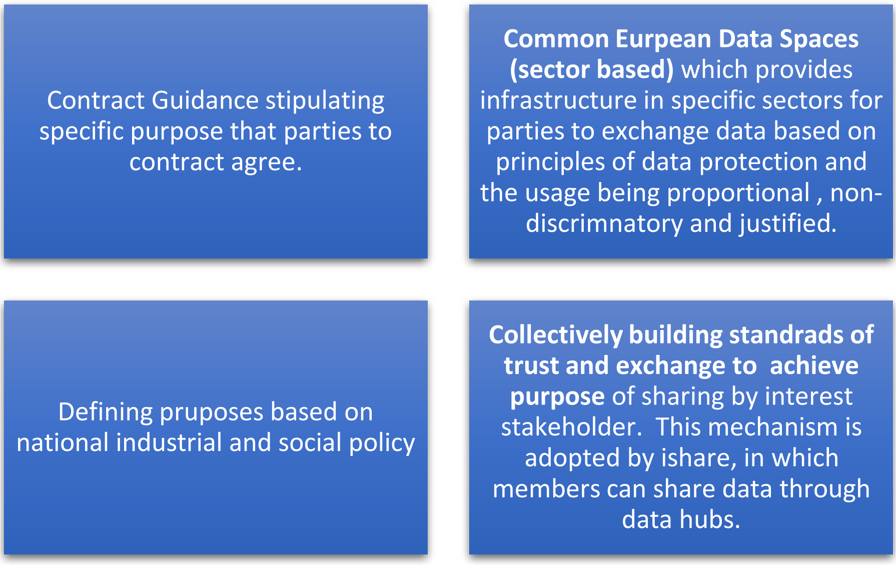
[75] *https://docs.google.com/document/d/118RqyUAWP3WIyyCO4iLUT3oOobnYJGibEhspr2v87jg/edit#heading=h.sd38vishuemj*

government policy programmes, research development and government service delivery. Each of these purposes has been elaborated in a manner that their meaning and intended objective becomes clear. For example, for the first two purposes, the benefits should likely extend beyond the individual to the community, while the third purpose may be targeted individually. It also provides clarifications on the purposes, which are not included within these categories.

Moreover, it stipulates for the satisfaction of purpose test, through which it can be established that such data is reasonable and necessary and in consonance with the above-mentioned purposes. Some important learning points from this framework are regarding the way, in which clarity can be ensured in giving meaning and limiting the scope of public interest purposes and providing purpose tests, which can facilitate data equity in the sharing process.

While these are some cases that stood out, there are other approaches to achieve clarity in the purpose of sharing. These approaches depend upon the context of sharing that is, whether it's business-to-business (B2B) or business-to-government sharing or *vice-a-versa*. The broader strategies also stipulate purpose in consonance with overall industrial and social policy priorities. In such cases rather than focusing on defining the purpose objectively, emphasis is placed on clearly stipulating standards and conditions of re-use on a case-to-case basis. Some of these approaches are illustrated in the figure below.

**Figure 5: Mechanisms to approach the purpose of data-sharing**



Contract Guidance stipulating specific purpose that parties to contract agree.

**Common Eurpean Data Spaces (sector based)** which provides infrastructure in specific sectors for parties to exchange data based on principles of data protection and the usage being proportional , non-discrimnatory and justified.

Defining pruposes based on national industrial and social policy

**Collectively building standrads of trust and exchange to achieve purpose** of sharing by interest stakeholder. This mechanism is adopted by ishare, in which members can share data through data hubs.

Seeing there are already quite a few data access initiatives, we attempted to explore Indian models of data access that are entrenched to achieve public interest purposes, our **stakeholder consultations** shed light on one such initiative on the Electricity Supply Monitoring Initiative (ESMI) undertaken by the Prayas Energy Group.[76] While this is an open data initiative, it was insightful to trace its life cycle to understand the way in which public interest purpose could be anticipated and appropriate technical and ethical considerations need to be evaluated to realise its benefits (Box 2).

**Box 2: Lessons from the Electricity Supply Monitoring Initiative (ESMI) by Prayas**

The ESMI initiative was started in 2007, was undertaken due to the data gaps in supply quality of electricity and frequent power outages, which was identified by Prayas, a civil society group that has been working towards consumer protection in the sector. From the start, the project **had a clear objective** to understand the performance of utility through comparing the prescribed load sharing protocol with on-ground realities. However, to sustain the project, ESMI deployed the GSM monitoring to build a more robust infrastructure to take the data back to the consumers and other regulators. In the past few years, the open data from this initiative have been used by the regulators as well as the ministries and consumers at instances to hold the utilities liable.

[76] *Prayas (Energy Group), 'Electricity Supply Monitoring Initiative (ESMI) - Prayas(Energy Group)', accessed 16 May 2021, https://www.prayaspune.org/peg/resources/electricity-supply-monitoring-initiative-esmi.html.*

The initiative also faced challenges at the initial stages related to data collection techniques and getting accustomed to the cultural and geographical sensitivities of different areas. In dealing with such challenges, the consent related to data collection and being cognizant of political risks, tracing of data locations are critical. While increasingly the data is being available, however, the public interest in such data can only be ensured through the **proper quality of data being available**. Thus, to hold regulators responsible and utilise data, **the authenticity and valuation of data are vital**. Through this initiative, many consumers were able to hold authorities and utilities accountable.

The ESMI life-cycle highlights that to realise the public interest in data sharing a well-defined objective, proper technical capacities, authenticity, and quality of data are vital. At the same time, it is also important that the data reaches the right stakeholders in a way that it can be utilised for social progressive objectives. While the ESMI was an open data initiative, some of the factors identified here could be crucial for the data trustees to validate data requests such as the expertise of the data requestor, their vision, and intended beneficiaries. It also gives a good example of data equity, as here the envisioned beneficiaries were in fact able to use this data for public interest data.

Furthermore, it was emphasised by the **stakeholders** that a prerequisite to effectively implement a public interest objective is to define a community in a limiting manner. For doing so the community may be restricted to a region, locality, or other tangible criteria, which can enable in identifying the uniqueness of their interest effectively and then move on to understand mechanisms, through which their interest can be made more representational through mandatory consultation.

Data Trustees should undertake the responsibility of conducting these consultations in a transparent manner. Even within the data trustees' model, there should be sub-committees that are more approachable. Moreover, the challenge also lies in holding the entities accountable to the initially identified purpose. It is here that ensuring data equity through transparency and accountability, which can be given through giving people various avenues, through which they can approach relevant authorities, is vital.

## Conclusion and the Way Forward

While defining the purposes and the objective for data sharing is vital, it is equally important to have mechanisms in place, which can extract specific problem statements and the kind of data required to achieve a particular objective. This helps in building trust between stakeholders involved in data sharing. In this regard, data stewardship responsibilities of the data trustees play a critical role as they are responsible to initiate discussions around brokering appropriate data usage between parties as well as adequately represent the interest of the community in achieving such an objective. In this regard, the comparative jurisdictional analysis indicates that having a proper procedure to entertain data usage requests, appropriately defining the purpose of data trustees in data sharing, and stipulating conditions of data re-use are vital.

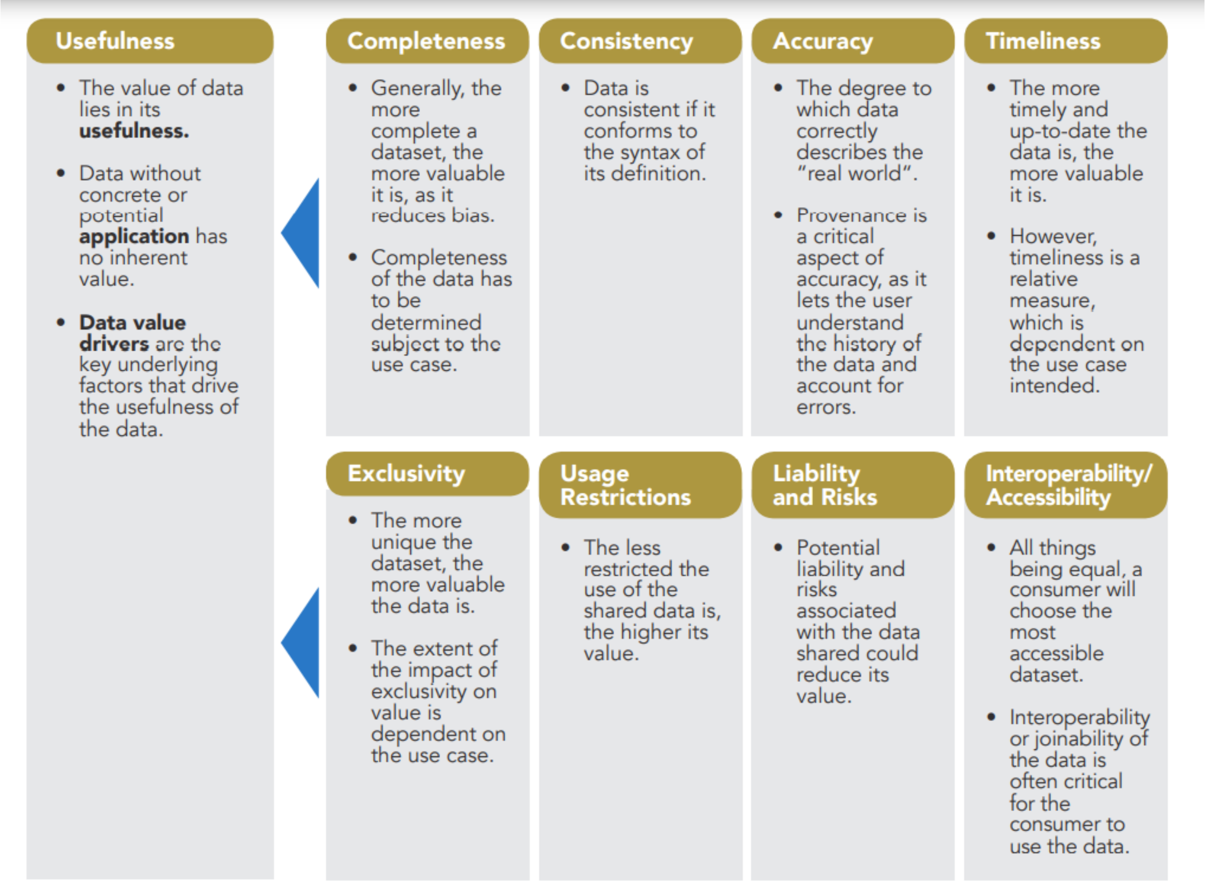
Different jurisdictions have used different methods to attain these vitalities, however, having a clear problem statement, targeted beneficiary, and mechanisms to hold stakeholders accountable for over-spill. As indicated by the analysis above the Indian Report misses on clear stipulating these parameters, without which a functional mechanism to approach 'public interest' purpose.

# 4

# Data Valuation and Incentive Mechanisms

To derive the value out of data, as well as to design a policy that can do so, it is important to ascertain the value of data in more tangible terms, through which a better understanding of the trade-offs in data sharing could be developed. These trade-offs could justify the development of much-needed incentive-sharing mechanisms, not just to encourage data sharing, but also to assess the value of data that is being exchanged. In a traditional sense, the value of data relies on several parameters and is based on the nature and economic characteristics that are associated with it. Data has been treated both as an asset as well as a public good. Therefore, the categorisation of data essentially defines the way the data could be valued.[77]

**Figure 6: Data Value Drivers according to Singapore's Data Valuation Guide for Sharing[78]**



**Usefulness**

- The value of data lies in its **usefulness.**
- Data without concrete or potential **application** has no inherent value.
- **Data value drivers** are the key underlying factors that drive the usefulness of the data.

**Completeness**

- Generally, the more complete a dataset, the more valuable it is, as it reduces bias.
- Completeness of the data has to be determined subject to the use case.

**Consistency**

- Data is consistent if it conforms to the syntax of its definition.

**Accuracy**

- The degree to which data correctly describes the "real world".
- Provenance is a critical aspect of accuracy, as it lets the user understand the history of the data and account for errors.

**Timeliness**

- The more timely and up-to-date the data is, the more valuable it is.
- However, timeliness is a relative measure, which is dependent on the use case intended.

**Exclusivity**

- The more unique the dataset, the more valuable the data is.
- The extent of the impact of exclusivity on value is dependent on the use case.

**Usage Restrictions**

- The less restricted the use of the shared data is, the higher its value.

**Liability and Risks**

- Potential liability and risks associated with the data shared could reduce its value.

**Interoperability/ Accessibility**

- All things being equal, a consumer will choose the most accessible dataset.
- Interoperability or joinability of the data is often critical for the consumer to use the data.

[77] *Bennett Institute for Public Policy, University of Cambridge. The Value of Data: Policy Implications. February 2020. https://www.bennettinstitute.cam.ac.uk/media/uploads/files/Value_of_data_Policy_Implications_Report_26_Feb_ok4noWn.pdf*

[78] *https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Guide-to-Data-Valuation-for-Data-Sharing.pdf?la=en*

## Variables at Play

While the CoE has talked about the nature of data, the discussion surrounding the valuation (monetary or incentive) from its established nature has been missing. Wide-ranging variables directly impact and define the value of data. Economic characteristics such as the non-rival nature, excludability, externalities in tandem with its informational characteristics like the subject matter, quality, sensitivity, and interoperability are used to assign the value of data. Further, factors such as exclusivity, accuracy, timeliness, restrictions, liabilities, and risk are used to pin down the value of data (as indicated in Figure 6).

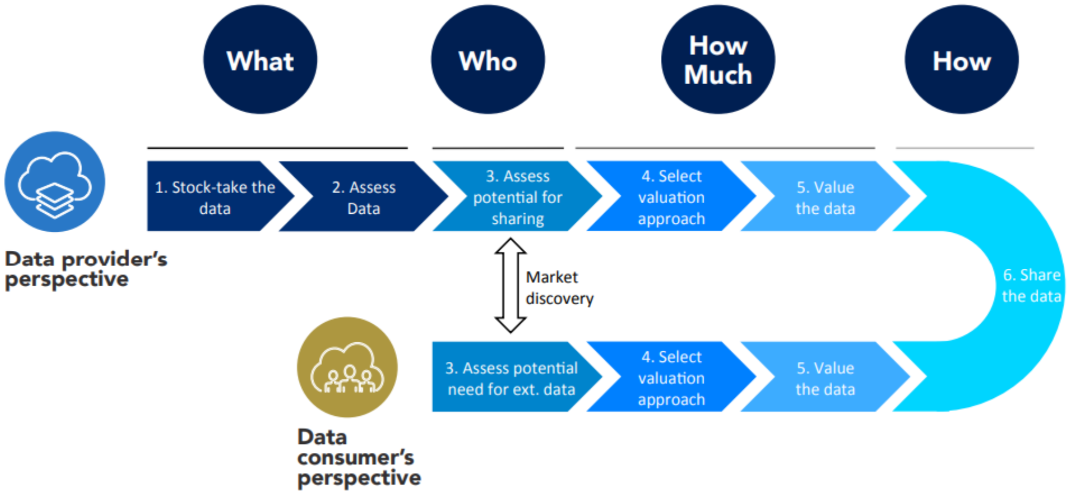
Without a nuanced discussion and understanding of these variables, the implementation of the proposed framework cannot be foreseen. There is a need for informed discourse about the way data will be valued. According to **our comparative jurisdictional analysis**, nation-states have been experimenting with several different ways to establish ways to value data, with pilot studies, wide-ranging discussions, and detailed road maps to achieve the same. Without such a blueprint, a mere statement of data value is unlikely to roll out as an efficient policy but is instead likely to create uncertainty among investors and our adolescent market itself (also see Annexure I).

Several data governance models, like the Data Trustee model, must have a different evaluation strategy than the marketplace buying and selling of data. During **our stakeholder consultations**, a variety of stakeholders pointed out that transparent data valuation is a necessary characteristic of a healthy data market. This not only facilitates data sharing between parties but also can incentivise the sharing. Further, it was noted that a blanket method to value data through mandatory sharing cannot work as different kinds of data hold or create different values.

Therefore, a deeper insight is needed in order to determine multiple ways of deriving the value of data. Data also has to be valued differently when it comes from an end-user or consumer, in comparison to a data fiduciary. Therefore, the stages of data sharing are different for different stakeholders. The common denominator in the stages among all data-sharing stakeholders is the recognition of the data market.

Once the assessment of the potential of sharing data is determined, the appropriate data valuation approach is chosen. This closely ties in with the notion that incentive or impetus to share data cannot be created within the industry without a clear mechanism, through which the value of their data can be realised and, in some way, can flow back to them.

## Figure 7: Data Sharing Process from different stakeholder perspectives in a data market according to Singapore's Data Valuation Guide for Sharing[79]


What
Who
How Much
How
Data provider's perspective
1. Stock-take the data
2. Assess Data
3. Assess potential for sharing
4. Select valuation approach
5. Value the data
6. Share the data
Market discovery
Data consumer's perspective
3. Assess potential need for ext. data
4. Select valuation approach
5. Value the data

*(In this figure Data Providers can be considered creators or owners of data, for example, government agencies and businesses. Whereas Data consumers collect or buy external data to generate additional insights and supplement internal functions)*

## Approaches

Depending on the data sharing use case in mind, governments, institutions, businesses, and data consumers can value data for their individual purposes. It is to be noted that the value of data is different for each actor in the process as one kind of data may be much more valuable to one actor than another actor.

Depending upon the purposes and actors involved, among other variables mentioned above, different approaches are used to ascertain the value of data:

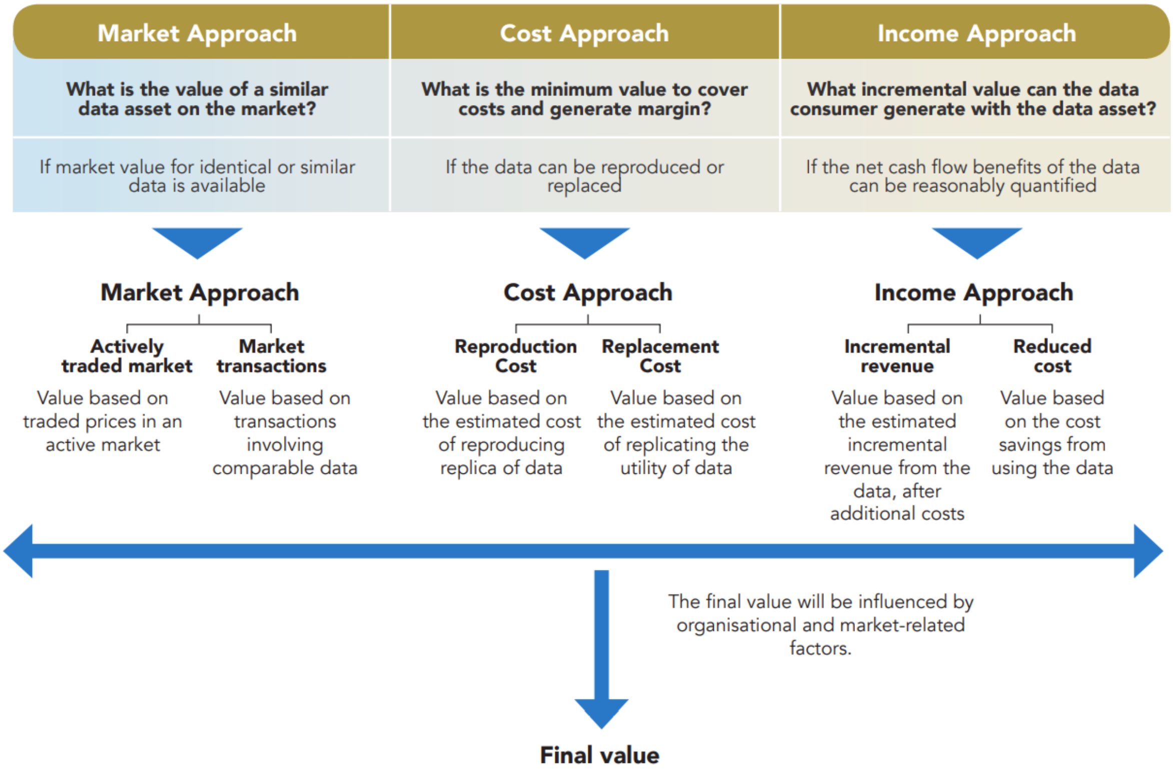
a. **Market Approach**- In this approach, the value of data is determined by using the market value of identical data or a data asset similar in nature.

b. **Cost Approach**- In the Cost Approach, the costs incurred to create the data are used to ascertain the value of data. This also involves the "data reproduction costs" and "data replacement costs" methods. This approach provides a base value of data, which may be coupled with the potential of the said data in economic returns.

**Income Approach**- In this approach, the value of data directly correlates with its ability to generate economic value in the future. There are several technical variables that come into play in this approach, leading to a reliable estimate of data value.

[79] https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Guide-to-Data-Valuation-for-Data-Sharing.pdf?la=en

## Figure 8: Overview of different kinds of data valuation approaches according to Singapore's Data Valuation Guide for Sharing[80]



| Market Approach | Cost Approach | Income Approach |
|---|---|---|
| **What is the value of a similar data asset on the market?** | **What is the minimum value to cover costs and generate margin?** | **What incremental value can the data consumer generate with the data asset?** |
| If market value for identical or similar data is available | If the data can be reproduced or replaced | If the net cash flow benefits of the data can be reasonably quantified |

**Market Approach**

- **Actively traded market**: Value based on traded prices in an active market
- **Market transactions**: Value based on transactions involving comparable data

**Cost Approach**

- **Reproduction Cost**: Value based on the estimated cost of reproducing replica of data
- **Replacement Cost**: Value based on the estimated cost of replicating the utility of data

**Income Approach**

- **Incremental revenue**: Value based on the estimated incremental revenue from the data, after additional costs
- **Reduced cost**: Value based on the cost savings from using the data

The final value will be influenced by organisational and market-related factors.

**Final value**

While most data sharing policies and initiatives do not specify data valuation mechanisms or approaches since they are in either discussion stages or the data sharing is for public purposes only. However, in the European Strategy for Data, the market approach where the valuation of data is based on the contracts is proposed.[81]

Similarly, the Japanese Contract Guidance on Utilization of AI and Data by the Ministry of Economy Trade and Industry refers to the market approach, specifying contractual terms to specify licensing and profit-sharing.[82] The Japanese Act on Special Measures for Productivity Improvement, 2018[83] further proposes tax breaks to businesses who are certified with an innovative plan for data use. And Singapore's Trusted Data Sharing Framework discusses the different approaches and ways to value data at length, which has been covered in the figures above (also see Annexure I).

[80] https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Guide-to-Data-Valuation-for-Data-Sharing.pdf?la=en
[81] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN
[82] https://www.meti.go.jp/english/press/2018/0615_002.html
[83] https://www.meti.go.jp/english/press/2018_06/0606_001_00.html

## Allied Factors

In conjunction with the data valuation, another important factor that has not been deliberated upon by the Report at length is the Intellectual Property Rights (IPR) considerations. The Report bypasses discussing and placing value in IPR, citing public interest from data sharing. IPR considerations need to be taken into account when it comes to data sharing, otherwise, it will stifle innovation and deter investments in the data economy since we do not have trade secret protection laws. There is a need for the recognition of ownership of not only data as well as copyrights over data. As seen in other jurisdictions, data itself cannot be copyrighted, compilations of data that display sufficient creativity in the arrangement, annotation, or selection can and must be protected.[84]

Japan's Contract Guidance on Utilization of AI and Data by the Ministry of Economy Trade and Industry 2018[85] has been analysing the intellectual property and ownership rights on data, while also studying group steps for exploring intellectual property rights in the fourth industrial revolution.[86]

Similarly, the Trusted Data Sharing Framework of Singapore even specifies the compilations, which have been afforded copyright protection. Additionally, it stipulates that acquisition of the ownership of the data allows for broadly unfettered usage of the data, while licensing may place limitations on the use of the data, depending on the scope and terms of the licence. Thereby establishing the need for the organisations to understand licensing terms before engaging in data sharing.

Considering the lack of incentives in the framework, the report should keep in mind other proposed policies in the domain of data economy or that may be affected by the NPD framework. For instance, in the Draft National Geospatial Policy 2021, a sector-specific initiative must not be made part of the mandatory data sharing process of high-value datasets as proposed in the NPD framework as the Geo-Spatial Data is proprietary and the proposed policy already deals with data sharing in the domain.[87]

Further, the concerns regarding reasonable charges echoed across stakeholders in our consultations. It was pointed out that those reasonable charges are not enough to cover appropriate costs of data sharing and compliance. In the current form, it is likely to become a norm that the data businesses have to start bearing these costs, adding

[84] https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Trusted-Data-Sharing-Framework.pdf

[85] https://www.meti.go.jp/english/press/2018/0615_002.html, https://www.meti.go.jp/press/2019/04/20190404001/20190404001-1.pdf.

[86] https://www.meti.go.jp/english/press/2017/0419_001.html

[87] https://dst.gov.in/sites/default/files/Draft%20NGP%2C%202021.pdf

additional burden over them. While no other jurisdictions have specified "reasonable charges".

The Public Sector Information Directive 2019[88] (PSD2) of the European Union, it is stipulated that the recovery of the marginal costs incurred for the reproduction, provision, and dissemination of data as well as for anonymisation of personal data and measures taken to protect commercially confidential information could be allowed, however, this is not applicable to the private sector. Along with this, member states may exempt bodies making high-value datasets available free of charge that are required to generate revenue to cover a substantial part of their costs.

## Conclusion and the Way Forward

The Report aims to create a strong data economy. To do that and ensure such an economy is sustainable, there is a need to foster discussions and define possible approaches in data valuation in the context of Indian economies. This has to be done in conjunction with discussing the nature and scope of data itself, where the Report itself is unclear in treating data only as an economic resource.

Data Valuation approaches need to be defined and adapted to Indian complexities and regulations while ensuring appropriate provisions for consumers and data principals. There is also a need to address the binaries between Personal Data and NPD when evaluating the data. The CoE must also rely on pilots and studies that are trying to study the valuation of data in a data-sharing economy.

In harmony with other frameworks dealing with data, the CoE must also make a tangible effort to keep the proprietary data out of the NPD Sharing Framework, thereby ensuring that the valuation of data is not on factors or kinds of data that have been regulated elsewhere. The CoE must also reconsider the "public benefit" argument while justifying the violations of IPR. The committee must ensure that the IPR is protected when it comes to datasets as well as copyrights over data itself, to protect investments in Indian markets and to attract further investments. The CoE must also attempt to evaluate the way costs will be borne in the data market and simplify the "reasonable charges" principle, making it more appropriate for a multi-model data sharing market, to ensure that unclear models do not set precedents for harmful norms for consumers and the markets.

[88] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L1024&from=EN

# 5

# Accountability and Consumer Rights

While the Report has made significant considerations for key issues like privacy and security, it does not have a roadmap or a design in itself on how these proposed considerations will actually work. The Report does not rely on use cases or empirical evidence to recommend the solutions to privacy, security, or collective privacy, and falls short on providing the technical details associated with these recommendations.

## Privacy

Our comparative jurisdictional analysis also indicated that in the Indian context, these considerations are likely to fall short as they do not account for the level of security and privacy NPD may need. Evidence suggests that identity can never be excluded from data and therefore misuse of NPD can be potentially just as risk-prone to individuals and communities as Personal Data,[89] which is generally provided with a higher degree of security and privacy protocols.[90]

The proposal of collective privacy in the Report also needs to be scrutinised as empirical research indicates that collective privacy as proposed may not protect consumer rights and there may be a need to further distinguish and classify collective and group privacy concepts.[91] Hybrid solutions like using the Commons to protect Data Subjects can be deliberated on, allowing for a balance between generating economic value of data while also protecting consumer privacy.[92]

Regulations across the globe are trying to strike a balance where they can derive economic and public value out of data, without risking the privacy and security of the individuals and communities. There are innovative ways, in which different regulations are approaching this, however, the Report does not take into consideration the entire spectrum of pitfalls and risks that are associated with NPD, and therefore lags in recommending solutions for the same.

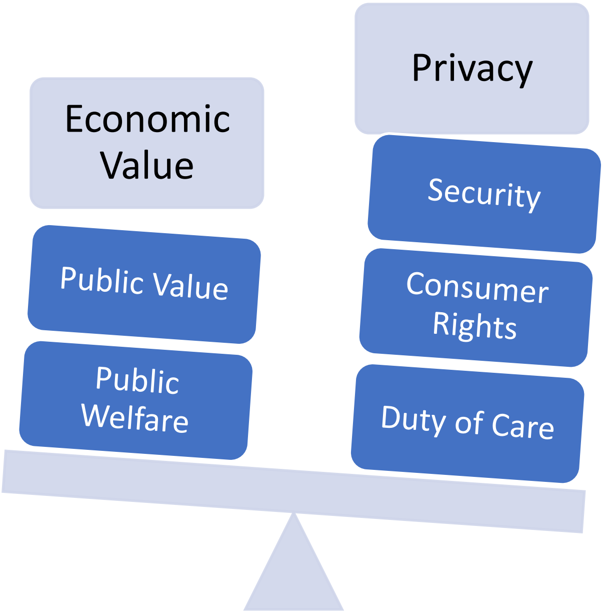
[89] *Kolata, G. (2019, July 24). Your Data Were 'Anonymized'? These Scientists Can Still Identify You. The New York Times. https://www.nytimes.com/2019/07/23/health/data-privacy-protection.html*

[90] *de Montjoye, YA., Hidalgo, C., Verleysen, M. et al. Unique in the Crowd: The privacy bounds of human mobility. Sci Rep 3, 1376 (2013). https://doi.org/10.1038/srep01376*

[91] *Taylor, L., Floridi, L., van der Sloot, B. eds. (2017) Group Privacy: new challenges of data technologies. Dordrecht: Springer. https://www.stiftung-nv.de/sites/default/files/group-privacy-2017-authors-draft-manuscript.pdf*

[92] *Wong, J & Henderson, T 2020, ' Co-creating autonomy : group data protection and individual self-determination within a data commons ' , International Journal of Digital Curation , vol. 15 , no. 1 . https://doi.org/10.2218/ijdc.v15i1.714*

**Figure 9: The delicate balance between extracting value out of data while also protecting consumer rights**



## Checks and Balances

There is a noticeable lapse in the checks and balances as well as the grievance mechanisms when it comes to non-personal data sharing. The framework does not talk about the three-pronged proportionality doctrine that has been established in the *Puttaswamy* case by the Supreme Court to determine the validity of rights restricting measures.[93] The doctrine postulates that the nature and extent of the State's interference with the exercise of a right must be proportionate to the goal it seeks to achieve.[94]

During the stakeholder consultations on the need for necessary checks and balances, it was presented that establishing a duty of care by the state, data trustee and other involved parties is essential. However, defining this duty of care is closely related to the culture and historical context of individual states. For example, in the United States (US), there is a clear exemption from duty of care because this allows for economic benefits to flow to the US and its inherent capitalist model, leaving the rest of the world to follow their rules since a majority of technology companies are American.[95] In contrast, the EU

[93] *Justice K.S.Puttaswamy(Retd) vs Union of India. (2017) 10 SCC 1*

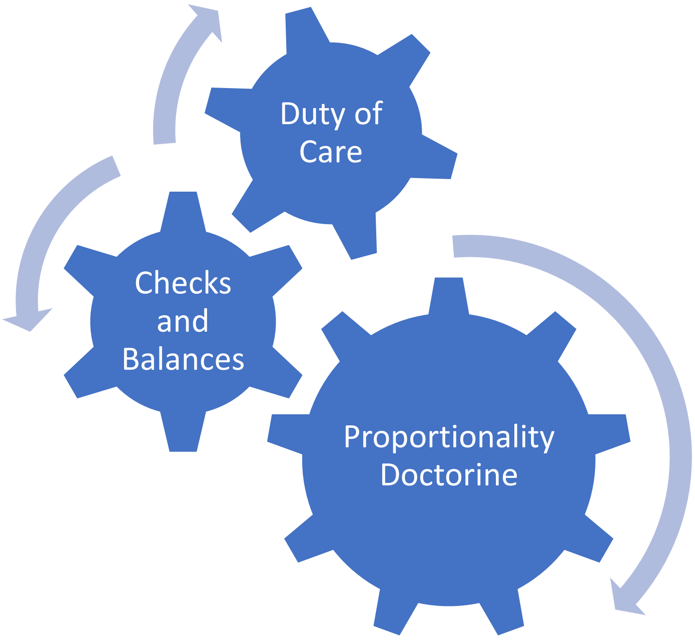
[94] *Bhandari, Vrinda; Kak, Amba; Parsheera, Smriti; Rahman, Faiza. "An Analysis of Puttaswamy: The Supreme Court's Privacy Verdict". IndraStra Global. 003: 004. ISSN 2381-3652*

[95] *Wolf Sauter, A duty of care to prevent online exploitation of consumers? Digital dominance and special responsibility in EU competition law, Journal of Antitrust Enforcement, Volume 8, Issue 2, July 2020, Pages 406–427, https://doi.org/10.1093/jaenfo/jnz023*

has a rights-based framework,[96] and in Singapore, there is a duty to protect commercial interests.[97]

However, in India, this duty of care is not clearly defined and is generally linked to cases of negligence and tortious liability.[98] While there are legal checks and balances prescribed by the Supreme Court under the set principles of "legality, necessity, and proportionality". In the Indian context, it has been observed that the question of duty of care is closely tied to the associated target population.

**Figure 10: The interconnected prerequisites ensuring consumer protection in data sharing**



In terms of state capacity to implement this duty of care and ensure inclusive public interests, regulation may itself play an important role. For example, in the case of Kenya, after having introduced a digital identity programme in place, the government excluded refugees to join the programme and therefore, left them without benefits. In response and protest, several refugee groups along with the civil society groups came forward and were able to register themselves, thus establishing a duty of care towards the refugees on the Government's part.[99] However, this might not be the case with every country, for

[96] *Chris Jay Hoofnagle, Bart van der Sloot & Frederik Zuiderveen Borgesius (2019) The European Union general data protection regulation: what it is and what it means, Information & Communications Technology Law, 28:1, 65-98, DOI: 10.1080/13600834.2019.1573501*

[97] *OneTrust DataGuidance (2020) Comparing privacy laws: GDPR v. Singapore's PDPA, https://www.dataguidance.com/sites/default/files/gdpr_v_singapore_final.pdf*

[98] *Choudhry, S., Khosla, M., & Mehta, P. B. (Eds.). (2016). The Oxford handbook of the Indian constitution. Oxford University Press.*

[99] *Refugees and Identity: Considerations for mobile-enabled registration and aid delivery (2017) GSMA Intelligence & DFID. https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2017/06/Refugees-and-Identity.pdf*

instance, in the Netherlands, the use of technology for economic surveillance led to the fall of the government.[100]

Here, it is important to note that the duty to care does not merely come out of data protection or technology regulations and discussions but also from constitutions, existing social protection laws, the internationally recognised principles of human rights, and principles of natural justice.[101]

In several of these cases, strong civil society and academia that have not necessarily been working on technology issues have made interventions to raise these concerns, as seen in South Africa, where the state contracted with a commercial intermediary to distribute welfare payments and the intermediary committed large-scale fraud and civil society stepped in the fight to protect the communities.[102]

This again ties back to the exclusion issues and the power dynamics. It circles back to the fact that those who actually control the benefits and have a "duty of care" usually use these towards targeted and specific sections of the community.

## Grievance Redressal

While the Report states that grievance redress mechanisms would be set up to address concerns by the data trustees, however, communities or consumers would not be able to make use of the redress mechanisms without a clear prescription and understanding of harms, and approachable avenues for redressal. This was also highlighted in a CUTS survey, which observed that most consumers are not aware of avenues for grievance, and only half of those who have earlier experienced a privacy breach went on to complain about it.[103]

Such issues will dilute the community benefit objective and place consumers at the margins of the data sharing value chain, without any necessary recourse. Overall, this points to insufficient focus on the onus of the government, regulators, and intermediaries to create an environment where consumers feel empowered to contest the decision at various levels.

In the following table and explained in detail in Annexure I, we compare various data-sharing initiatives and policies from around the world in contrast to the NPD governance framework proposed in India. The **comparison** indicates that along with addressing the

[100] *A benefits scandal sinks the Dutch government. Jan 23rd, 2021. The Economist. https://www.economist.com/europe/2021/01/23/a-benefits-scandal-sinks-the-dutch-government*

[101] *Nolan, D. (2013). Deconstructing the Duty of Care. The Law Quarterly Review, 129, 559-588.*

[102] *Gabriella Razzano. Sassa Grants: The small information win hiding in the grant crisis. 24 April 2017. Daily Maverick. https://www.dailymaverick.co.za/opinionista/2017-04-24-sassa-grants-the-small-information-win-hiding-in-the-grant-crisis/*

[103] *https://cuts-ccier.org/pdf/survey_analysis-dataprivacy.pdf*

aforementioned concerns, the policies have given due consideration to appropriate accountability and grievance redressal despite differences in the governance mechanisms of the data-sharing initiatives.

**Table 4: Accountability Mechanisms (also refer to Annexure I)**

| Policy/Initiative | Governance Methodology | Accountability |
|---|---|---|
| Framework for the free flow of non-personal data in the European Union 2019[104] | Open Standards, Self-regulatory code based on the principles of transparency, interoperability, and accountability. | Anonymised data that has the possibility of de-anonymisation will be considered as personal data. |
| Proposal for a Regulation on European data governance (Data Governance Act) 2020[105] | Three modes of governance of data sharing are based on the use and purpose. | Right of privacy under the GDPR and e-privacy directive. |
| Trusted Data Sharing Framework[106] | Multiple governance models. Supervisory authority not directly involved. Open to new future models. | Based on the trust Principles of Transparency, Accessibility, Standardisation, Fairness and Ethics, Accountability and Security and Data Integrity. |
| Data Sharing and Release Legislative Reforms, 2019[107] | Independent oversight to promote sharing and safe practices while acting as a watchdog. Independent data sharing falls under this umbrella regulation. | Privacy, Transparency, and a detailed Grievance Redressal System |

[104] https://ec.europa.eu/digital-single-market/en/free-flow-non-personal-data

[105] https://ec.europa.eu/digital-single-market/en/news/proposal-regulation-european-data-governance-data-governance-act

[106] https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Trusted-Data-Sharing-Framework.pdf

[107] https://www.datacommissioner.gov.au/sites/default/files/2019-09/Data%20Sharing%20and%20Release%20Legislative%20Reforms%20Discussion%20Paper%20-%20Accessibility.pdf

| Policy/Initiative | Governance Methodology | Accountability |
|---|---|---|
| Contract Guidance on Utilization of AI and Data by Ministry of Economy Trade and Industry 2018[108] | Governed by contractual terms with certain prerequisites. | Personal information protection, liability on operators and data providers. Grievance redressal. |

In contrast with the policies from other jurisdictions, it is evident that the Report leaves very little possibility for future improvements. This is an even bigger concern given the fact that this report is acting as a conversation starter, described as such by the members of the expert committee, rather than the final version of the legislation itself.

## Conclusion and the Way Forward

Based on our learnings from the stakeholder consultations and the comparative analysis of policies from various jurisdictions, it can be said that the Report severely lacks accountability, as well as the protections needed for consumers when it comes to non-personal data sharing.

There is a need for in-depth consideration to protect any personally identifiable data or data that can be used to identify a group or a community. Group privacy needs to be appropriately defined to ensure such a definition can work with future regulations. To come to such a definition, deeper insights on collective privacy, its opportunities, and more importantly the possible risks and harms need to be studied, in contrast to individual privacy, particularly in the Indian context.

Further, to avoid consent fatigue and user rights protection, the consent must be redesigned as an "opt-in", given the limitations of the digital literacy in the country, ensuring that consumer does not have to go out of the way to revoke their consent, thereby promoting a consumer empowering privacy architecture.

There is also a need for independent judicial or quasi-judicial oversight over the executive authority to ensure that the executive power is kept in check and not abused. A thorough and transparent grievance redressal system is also needed to ensure that consumers can resolve their complaints in this technically complicated process of data sharing.

A harm-based approach towards ensuring consumer welfare, where the approach consists of the best parts of all approaches, modified for the Indian context may be

[108] *https://www.meti.go.jp/english/press/2018/0615_002.html, https://www.meti.go.jp/press/2019/04/20190404001/20190404001-1.pdf.*

considered, where a duty to protect consumer interest is enshrined in the regulation and harm minimisation is the dictating principle for non-personal data sharing.

# Key Takeaways

This study has undertaken the analysis of the approaches to data sharing and governance model proposed by the Report in India. Through the analysis, it can be deduced that governing data economy is a challenging task due to the multifaceted nature of data and stakeholders involved in its management. The primary objective should be first to gather evidence and conduct an impact assessment to adduce the current status and expected value creation within the data economy.

Considering India is at a nascent stage of regulatory developments in the context of data, these pieces of evidence are the key to provide grounds for developing principles of governing data. The analysis conducted in this study also indicates that the approaches prescribed by the Report have missing linkages and unclear framing, which leaves open room for building uncertainty for all stakeholders. In light of this some key takeaways from the analysis are presented below:

1) **Scope of Data –** It is important to comprehensively understand the scope of data to determine various data typologies across the data value chain due to - lucidity of data; variations that may occur resulting from the way it is collected; the entity which collects it; and the purpose for which it is collected. The Report places heavy reliance on anonymisation techniques in categorising personal data and NPD, however, due to the complicated application of such techniques, approaching NPD as a separate category becomes difficult. In this context, it could be beneficial to take an ecosystem approach to understand data typologies and map data lifecycle within the ecosystem to ascertain stages where anonymisation could be applied with the least risks.

   Additionally, while proposing the new category of HVD may seem like a step in the right direction, but the first step should be to understand whether such data can be separated from other datasets in actuality. According to those stipulating different treatments to different kinds of data across the spectrum while being sensitive to proprietary and overlapping interests in data, will be beneficial.

2) **Stakeholder Interactions and Governance Mechanisms –** The analysis points to two critical aspects in designing data sharing - the role of stakeholders and governance mechanisms that stipulate their interactions. The concern overall with the framework proposed by the Report is related to achieving representation, balancing the value different stakeholders have in the data value chain, and ensuring trust and impetus amongst the parties to take up data sharing.

Different approaches for achieving these objectives have developed as highlighted from the analysis of different jurisdictions. However, pinpointing any single appropriate approach is very difficult in the Indian context because of the lack of knowledge that exists about the data ecosystem and the divergent interest that exists in its value. At this point, the primary focus should be on identifying principles of governance that are important in data space and then identifying tangible mechanisms in the form of consultation or co-design, through an evidence-based approach.

Equally necessary is unpacking these parameters in the context of data to understand what representation actually means for communities in India and how do they think they should have control of their data.

3) **Purpose of Sharing and Expected Value Creation –** While defining the purposes and the objective for data sharing is vital, it is equally important to have mechanisms in place, which can extract specific problem statements and the kind of data required to achieve a particular objective. This helps in building trust between stakeholders involved in data sharing.

   In this regard, data stewardship responsibilities of the data trustees play a critical role as they are responsible to initiate discussions around brokering appropriate data usage between parties as well as adequately represent the interest of the community in achieving such an objective. In this regard, the comparative jurisdictional analysis indicates that having a proper procedure to entertain data usage requests, appropriately defining the purpose of data trustees in data sharing, and stipulating conditions of data re-use are vital.

4) **Data Valuation and Incentive Mechanisms –** The Report aims to create a strong data economy. To do that and ensure such an economy is sustainable, there is a need to foster discussions and define possible approaches of data valuation in the context of Indian economies. This has to be done in conjunction with discussing the nature and scope of data itself, where the Report itself is unclear in treating data only as an economic resource.

   The CoE must ensure that the IPR is protected when it comes to datasets as well as copyrights over data itself, to protect investments in Indian markets and to attract further investments. The CoE must also attempt to evaluate the way costs will be borne in the data market and simplify the "reasonable charges" principle, making it more appropriate for a multi-model data sharing market, to ensure that unclear models do not set precedents for harmful norms for consumers and the markets.

5) **Accountability** – There is a need for in-depth consideration to protect any personally identifiable data or data that can be used to identify a group or a community. Group privacy needs to be appropriately defined to ensure such a definition can work with future regulations. To come to such a definition, deeper insights on collective privacy, its opportunities, and more importantly the possible risks and harms need to be studied, in contrast to individual privacy, particularly in the Indian context. Further, to avoid consent fatigue and user rights protection, the consent must be redesigned as an "opt-in", given the limitations of the digital literacy in the country, ensuring that consumer does not have to go out of the way to revoke their consent, thereby promoting a consumer empowering privacy architecture.

   There is also a need for independent judicial or quasi-judicial oversight over the executive authority to ensure that the executive power is kept in check and not abused. A thorough and transparent grievance redressal system is also needed to ensure that consumers can resolve their complaints in this technically complicated process of data sharing.

   A harm-based approach towards ensuring consumer welfare, where the approach consists of the best parts of all approaches, modified for the Indian context may be considered, where a duty to protect consumer interest is enshrined in the regulation and harm minimisation is the dictating principle for non-personal data sharing.

To read further on the Rationale, and Assumptions on Data Sharing, visit our project page here. You can access the first report of this study here.

# Annexure - I

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| **Cross-Sectoral and Umbrella frameworks/initiatives/strategies/ guidelines for data sharing** | | | | | | |
| **European Union (EU)** | | | | | | |
| **Framework for the free flow of non-personal data in the European Union 2019**[109] | The objective of the framework is to achieve efficiency in data processing and creating the 'EU Digital Singles market through increasing data' **mobility across countries which have been inhibited due to data localisation practices of member states such as imposing technological requirements for storing of data in the geography of specific member states and other vendor lock-ins (cloud service providers) practices.**<br>**Before introducing the report impact assessment studies were conducted.**[110] | **Applies to non-personal data**. In the case of mixed data sets, it only applies to the non-personal part of datasets and in cases where **personal and non-personal data are intrinsically linked, the General Data Protection Regulation (GDPR) prevails**.<br>While the framework states that data that is anonymized is included within non-personal data, such assessment would have to be made on a case-to-case basis, depending on the technology of anonymization.[111] | It aims to facilitate the flow of data to competent authorities for official and legal duties as well as amongst private sector organizations and companies for commercial and economic purposes. These purposes are not further elaborated and will be based on self-regulatory codes developed by industry bodies. | The member states are required to update the European Commission about any new data localization framework introduced by them. The commission is responsible for updating the details of the same and making them available publicly through a website.<br>The framework encourages the **development of self-regulatory codes by the industry to facilitate porting of data based on the principles of transparency, interoperability and taking due account of** | **Based on a self-regulatory code of conduct.**<br>**No valuation mechanisms for data are given.** | The framework prescribes for following conditions for the flow of data-porting data in a structural and readable manner, sufficient information to be given to users before porting certification mechanism to compare quality management, information security and generate awareness about code of conduct<br>The framework specifically states that **any anonymised data that has the possibility of de-anonymization will be considered as personal data**.<br>The commission has been directed to submit a report |

[109] https://ec.europa.eu/digital-single-market/en/free-flow-non-personal-data

[110] file:///C:/Users/Shubhangi/AppData/Local/Temp/ImpactAssessmentSummary.pdf, https://ec.europa.eu/digital-single-market/en/news/facilitating-cross-border-data-flow-digital-single-market-study-data-location-restrictions

[111] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019DC0250&from=EN, the assessment of whether data is properly anonymised depends on specific and unique circumstances of each individual case17. Several examples of re-identification of datasets that were supposedly anonymised have showed that such an evaluation may be demanding18. To establish whether an individual is identifiable, one has to look on all means reasonably likely to be used by a controller or by another person to identify an individual directly or indirectly

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | | | | **open standards.** | | evaluating the implementation of this framework by 2022. |
| **GAIA- X[112] (expected launch in 2021)** | Project GAIA-X is a cloud initiative to create a data-sharing space (open digital ecosystem) in Europe, the lead of this initiative is taken by Germany and France. It connects centralised and decentralised infrastructures in order to turn them into a homogeneous, user-friendly system. The resulting federated form of data infrastructure strengthens the ability to both access and share data securely and confidently.<br>This initiative has also come as fostering the goals for EU Strategy for Data. | Participants can choose which data they wish to share with other companies or contribute to open data infrastructure. | Initially, the project has identified 40 uses cases over domains including –<br>**Industry 4.0/SME**<br>**Smart Living**<br>**Finance**<br>**Health**<br>**Public Sector**<br>**Mobility**<br>**Agriculture**<br>**Energy** | **In order to implement the federated data infrastructure, it is proposed to establish a central organisation at the European level. This organisation would lay the economic, organisational, and technical foundations of a federated data infrastructure. Its task will be to develop reference architecture, define standards, and determine criteria for certifications and product quality seals. It should be a neutral mediator and the hub of the European eco-system** | The incentives are to be decided between the parties. However, the infrastructure provides opportunities to parties involved to be engaged in a platform that provides secured usage sharing of data. | Depending on individual and sector-specific requirements, GAIA-X provides the platform for users to choose from services meeting their demands relating to e.g. rigorous information-security requirements, legal certainty within the framework of the European General Data Protection Regulation (GDPR), data storage within certain countries or regions or other specific attributes that users can leverage in making their choice. **The initiative is to be set up on European values of data sovereignty, user-friendliness, transparency, privacy, security, openness.** |
| **European Strategy for Data 2020** [113] | The measures laid out in this paper contribute to a comprehensive approach to the data economy that aims to | **Both personal and non-personal data for the government to business, business to** | **The strategy recognizes data sharing for the public good and gives examples such as climate change,** | Under the strategy **general principle is to facilitate voluntary data sharing**. | The strategy proposes evaluating existing IPR frameworks to further enhance data access and | The strategy focuses on increasing the competence of data principals by empowering them to be in |

[112] https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html

[113] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | increase the use of, and demand for, data and data-enabled products and services throughout the Digital Single Market in Europe.<br><br>The strategy at the outset establishes **that the EU has everything which can lead to the development of this initiative - technology know-how, implementation of regulation and policies like GDPR, FFD, Open Data Directive, Cybersecurity Act.**<br><br>There also has been sector-specific legislation and frameworks already in place for data sharing. Additionally while introducing this strategy there was parallel guidance issued on private-sector data sharing, which specifically notes the outcome of the public consultation which indicated that at this stage the horizontal legislation for private sector data sharing is not necessary and this could be proposed at a later stage.[114] | **business, business to government, and sharing amongst public authorities are prescribed for.**<br><br>**For mixed datasets, the strategy notes that businesses and governments should follow practical guidance prescribed for the businesses for mixed datasets by the earlier directive.**<br><br>Rights for co-generated data (such as IoT data in industrial settings), typically laid down in private contracts. | **predicting, and coping with natural disasters.**<br><br>**However, it does not prescribe mandatory sharing for such purposes. It also encourages data sharing for economic and commercial purposes.**<br><br>It proposed to set up sector-based European Data Spaces, which can work in an interoperable manner. | It stipulated that only where specific circumstances so dictate, access to data should be made compulsory, where appropriate, under fair, transparent, reasonable, proportionate, and/or non-discriminatory conditions.<br><br>**Additionally, mandatory sharing is only prescribed when there is a market failure in a particular sector.**<br><br>**The strategy proposes to explore the need for a legislative framework in the form of the Data Act of 2021- which would focus on sectoral needs, voluntary data sharing, and formulating data pools.** | use (including a possible revision of the Database Directive and possible clarification of the application of the Trade Secrets Protection Directive as an enabling framework) Concerning the valuation of data, private contracts are proposed. Additionally, it states that organisations would voluntarily contribute to data pools in return for data from other organisations, license fees, and data analysis tools. | control of their data through tools and means to decide at a granular level about what is done with their data ('personal data spaces'). For this, it also **proposes to enhance the portability righ**t for individuals under Article 20 of the GDPR.<br><br>It also proposes to increase data literacy and digital competence amongst the users. |

[114] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0232&from=EN

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| **Public Sector Information Directive 2019**[115] | The objective of this directive is to make public sector data available for commercial and non-commercial purposes.<br>The framework establishes an open data sharing mechanism for sharing public sector data with all entities and individuals. | **It covers existing documents and research data held by public sector authorities.**<br>The directive does not apply to –<br>Documents on which third parties hold IPR.<br>Documents that have sensitive data pertaining to national security<br>Under the directive, re-use of documents shall be open to all potential actors in the market, even if one or more market actors already exploit added-value products based on those documents.<br>The Directive also introduces the concept of "high-value datasets", defined as documents the re-use of which is associated with important benefits for the society and economy. The directive indicates forming a separate set of rules ensuring their availability free of charge, in machine-readable formats, provided via APIs, and where relevant be available as a bulk download. | **Both commercial and non-commercial purposes.** | Request for re-use of the data will be made to public authorities which will take such a decision within 20 working days.<br>The public authority will also assess if a license is needed for the requested re-use of the data. | Data is made available free of charge.<br>However, the recovery of the marginal costs incurred for the reproduction, provision, and dissemination of documents as well as for anonymisation of personal data and measures taken to protect commercially confidential information could be allowed.<br>Member states may exempt bodies for two years, where making high-value datasets available free of charge by public sector bodies that are required to generate revenue to cover a substantial part of their costs. | The directives prescribe that the re-use of documents shall not be subject to conditions unless such conditions are objective, proportionate, non-discriminatory, and justified on grounds of a public interest objective.<br>When re-use is subject to conditions, those conditions shall not unnecessarily restrict possibilities for re-use and shall not be used to restrict competition. |

[115] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L1024&from=EN

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| **Proposal for a Regulation on European data governance (Data Governance Act) 2020**[116] | The objective is to introduce governance, guidance, and standards that could facilitate data reuse and availability. | The **proposed act covers both personal (in an anonymized form deleting commercially confidential information) and non-personal data.** It gives a broader definition of data which covers digital representation of acts, facts, or information and any compilation of data in the forms of sound, visual or audiovisual recording. In defining, non-personal data it states that it means all other data that is not covered within the definition of personal data in the GDPR. Additionally, it also covers the definition of 'meta-data' which includes the date, time, and geo-location data, duration activity, connection to other natural persons. This act is likely to affect public sector undertaking, private sector data intermediaries, and consumers. | **The act does not lay down a specific purpose for data re-use and availability,** however specifically lays down the condition and standards for re-use. The larger aim of the act is to make diverse data available through various stakeholders in a trusted environment**.** | **The act largely introduces three modes of governance of data sharing and re-uses i.e. - conditions of re-use of public data which is not covered in the PSI directive on the grounds of commercial, statistical confidentiality, protection of IPR and covered by protection under personal data; data sharing through trusted data intermediaries ; and data altruism.** For re-use of public sector data, it stated that public sector bodies may impose conditions which are non-discriminatory, proportionate, and objectively justified, anonymisation conditions in case of personal data; re-use must be compliant with IPR, however with exception to the certain provision to the database directive. For this, the Commission proposes for member stated | **The act provides provisions for charging fees for the re-use of public sector data**.<br>However, in the case of other data sharing, the act does not prescribe any particular valuation mechanisms. | The act states that in any case the data cannot be used for purposes other than those specified.<br>Additionally, the regulation gives due consideration to the rights of data holders in the intellectual property regime, the **fundamental right of privacy under the GDPR and e-privacy directive, and freedom to conduct business.** |

[116] https://ec.europa.eu/digital-single-market/en/news/proposal-regulation-european-data-governance-data-governance-act

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | | | | to designate a competent body and setting up single information points to support public sector bodies that grant access to data. The commission also introduces notification requirements for intermediaries who will be involved in data exchange services.<br>The act also introduces the concept of data altruism which could be exercised through organisations that are to be registered with competent authorities. | | |
| **Singapore** | | | | | | |
| **Trusted Data Sharing Framework** [117] | The Framework is aimed to address concerns over trust and security hindering the mass sharing of data, despite the benefits that can be gained from leveraging large volumes and a variety of data for analytics, including machine learning artificial intelligence.<br>**This Framework is just a guide for industry and not for compliance** | For this Framework, "data" refers to both personal and business data (derived in the process of business, including non-personal data).<br>It states that in the case of personal data, additional safeguards should be followed by the parties.<br>This framework is intended for use | **The framework highlights that data sharing would help in developing Artificial Intelligence in Singapore. In this regard, the framework highlights some use cases of data sharing.** | The framework recommends that an institution or organisation empowered to operate a supervisory function related to the ecosystem may be set up. Such supervisory authority -<br>• May refer to the regulator (or other governing bodies), or industry | The framework recommends for where there is a need to assess the value of data on its own (e.g. when approached by business partners for data), organisations may consider the following three key actions:<br>**Take Stock of Own Data** - what are the kinds of data that exist like identifiable data sets, observed data, authored | This Framework introduces six trust Principles: **Transparency, Accessibility, Standardisation, Fairness and Ethics, Accountability and Security and Data Integrity** as foundations to forming a trusted data-sharing partnership<br>The framework also introduces risk assessment |

[117] https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Trusted-Data-Sharing-Framework.pdf

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | | in the commercial and non-governmental sectors but excludes data sharing in or with the public sector. | | bodies with oversight mandates or other practical influence (e.g. industry associations, standards institutes)<br>• Usually not directly involved in data sharing, but can influence the data sharing activities through legislative reviews, issuance of the guidelines, standards, or accreditation schemes.<br>**The framework also proposes the kinds of data sharing models that may be developed**.<br>**Bilateral** – two parties agree to share data, where sharing can be one-way or two. Trust principles can be decided between the parties.<br>**Multilateral** – three or more parties agree to share data, each acting as a Data Provider, a Data Consumer, or | data, derived data. The aim should be to form a data taxonomy.<br>**Assess Potential for Sharing** - When assessing potential use cases and data partners for the data, an organisation should consider all potential stakeholders in the whole value chain or ecosystem that the organisation operates in<br>**Consider Data Valuation Approaches-** market approach, cost approach, the income approach | parameters- lack of control over the use of data, lack of control of change in exchange or platform modification, insolvency, and reputational risks. |

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | | | | both. Trust can be established directly by the parties or institutionally.<br>**Decentralised** – includes peer-to-peer (“**P2P**”) and other distributed systems. These are designed to grant control over data access and sharing to a community of participants. Participants in this community may share data on a bilateral or multilateral basis, using advanced platforms governed by a system of incentives and crowd consensus. | | |
| **Australia** | | | | | | |
| **Data Sharing and Release Legislative Reforms, 2019**[118] | **The report forms the basis of the new regulation to be introduced for purposes of sharing such data. It introduces the standards for legislation that will** empower government agencies to safely share public sector data with trusted users for specific purposes. It aims to streamline | The new legislation will empower government agencies to safely share public sector data with trusted users.<br>**Public sector data is data held by the Australian government as it fulfils its various functions.** This may include data on topics as diverse as weather patterns, | Under the proposed Data Sharing and Release Legislative Reform, data sharing may occur for public benefit. The framework prescribes a purpose test to this end. This test is satisfied if sharing is reasonably necessary - to inform government policy, program and service | The report recommends setting up the **National Data Commissioner as an independent authority with oversight of the new data-sharing system**. The Commissioner will play an important dual role: championing greater data sharing while | Any cost and resource-related matters will be part of the data-sharing agreements**.**<br>**If the costs are to be incurred by the users, they will be informed about the same.**[119] | **The framework has proposed data sharing principles which are based on –**<br>Data sharing is for an appropriate project or program of work<br>Data is only available to authorised users<br>The environment in which the data is |

[118] https://www.datacommissioner.gov.au/sites/default/files/2019-09/Data%20Sharing%20and%20Release%20Legislative%20Reforms%20Discussion%20Paper%20-%20Accessibility.pdf

[119] https://www.pmc.gov.au/sites/default/files/publications/data-sharing-principles-best-practice-guide-15-mar-2019.pdf

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | and modernise data sharing, overcoming complex legislative barriers and outdated secrecy provisions. | who is coming and going from Australia, and administrative data about access to government services by both businesses and individuals. Such data may exist at different levels of detail, including aggregated to the category or population or at the more detailed unit record. | delivery or for research and development Commercial uses of public sector data by the private sector could be limited to non-sensitive data that is openly released.<br>The first two (government policy and programs and research and development) may involve the sharing of personal information but should result in outcomes for the entire community. In contrast, the final purpose (government service delivery) will involve the sharing of personal information and support better outcomes targeted at individuals no matter what community they belong to. | promoting safe data sharing practices. That framework recommends that the Commissioner should be empowered to apply strong penalties to intentional or negligent misuse and should cooperate with other regulators, including the Australian Information and Privacy Commissioner.<br>A National Data Advisory Council will be formed, advising the National Data Commissioner on the ethical database, community engagement, technical best practices, as well as industry and international developments.<br>**Data sharing agreements will be a requirement for all data sharing under the Data Sharing and Release legislation** | | shared minimises the risk of unauthorised use or disclosure.<br>Appropriate protections are applied to the data<br>Outputs are appropriate for further sharing or release<br>Along with safeguards of the Privacy Act of 1988. The report proposes of privacy by design approach in data-sharing agreements and will follow the principles laid out in the Privacy Act. However, it does not give a concrete view on consent and leave of National Data Commissioner.<br>To increase transparency, the registers of Accredited Data Service Providers and Accredited Users will show who has been accredited to offer data services, to access and work with data.<br>Include a complaints mechanism for Data Custodians, Accredited Users, and |

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | | | | | | Accredited Data Services Providers to raise system-specific complaints with the National Data Commissioner. |
| **Data Exchange Framework IT Strategy Action Plan 2017-18**[120] | This data exchange framework creates a standardised whole of Victorian government (WOVG) data exchange approach regardless of the datatype, classification, exchange method, platform, or intended use<br>The framework came about as support Victorian Centre for Data Insight's (VCDI). Data Reform Strategy, API (application programming interface) gateway. | This framework covers structured data i.e. data in the form of a database with appropriate contextual information.<br>It creates an exchange framework primarily for the government departments, however, the target audience for such data can be data custodians, data owners, etc. Hence, the framework focuses more on the government to government and non - government sharing. | There are specific purposes that are stipulated, however, such purpose should broadly be interest in the interest of the government, department, or public in Victoria. | In this framework data requestor, will have to submit a data request which underlines the kind of data requested, the purpose of use, whether such data is openly available. The request will be made to the provider after approval from the relevant government department.<br>Such requests will then be assessed under the Privacy Act 1988 (Cth), Victorian Data Sharing Act 2017, Public Records Act 1973, and Freedom of Information Act 1982. If there is no legal mandate to share the data contract agreement will be formulated.<br>Every data request will be assessed based on risk-based assessment and | No incentive strictures are defines, in case of any legal obligations concerning data ownership contractual agreements will support creative license requirements and terms. | **This data exchange framework is built-on – transparent and collaborative accountability,** data privacy, confidentiality, security, and intellectual property is respected and protected during and after the exchange of data**,** data is exchanged with the assurance provided for the appropriate use of data after the exchange |

[120] https://www.vic.gov.au/sites/default/files/2019-07/Data-Exchange-Framework_0.pdf, https://www.vic.gov.au/sites/default/files/2019-09/Data%20Exchange%20Guideline.PDF

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | | | | most data should be made unidentifiable. | | |
| **Japan** | | | | | | |
| **Contract Guidance on Utilization of AI and Data by Ministry of Economy Trade and Industry 2018**[121] | IoT and AI, data use is expected to create new value-added and solve societal issues through data collaboration that transcends business boundaries. It is often difficult, however, for businesses to conclude contracts related to the utilisation of data or AI technology **due to a lack of sufficient experience in contract practices and the gaps in understanding between the parties involved.**<br>**The guidelines highlight the questions and details that should be formulated while contracting for data sharing.** | The guidelines divide the contracts into different categories based on the purpose of sharing and include different kinds of data based on that –<br>**From one data provider to another** - whether to use derivate data or not, **notice to be given when data includes personal information.**<br>**Where data is newly created due to the involvement of multiple parties** – only the parties involved in data creation can use it, there might be a restriction on sublicensing to third parties.<br>**Sharing data through the platform** – type of data to be specified | Different contracts based on the purpose of data sharing – From one data provider to another – The purpose for which data is not allowed to be used should be mentioned.<br>Where data is newly created due to the involvement of multiple parties – terms of usage between the parties is to be specified<br>Sharing data through the platform - describing usage range of data or scope of usage in the agreement. | Data sharing would be governed by contractual terms for models of sharing which would include data sharing from one data provider to another, creation and sharing of data by multiple parties, or creating a data-sharing platform.<br>Contracts for any of these models would include clauses such as –<br>**Responsibility for disputes with third parties due to provided data**<br>**Scope of license to use provided data.· Guarantee / non-guarantee of data**.<br>**Liabilities of platform operators.· Liabilities of data providers and users.· at withdrawal/termination.** | **Contractual terms would specify licensing terms and profit-sharing** in case the data is created by multiple parties.<br>Additionally, analysis for exploring the intellectual property and ownership rights on data have already been undergoing since 2019, with a study group step for exploring intellectual property rights in the fourth industrial revolution.[122]<br>**There is no specific costing mechanism** prescribed for the data. | The guidance recommends for clauses to be included in the contract with regards to –<br>**Notices when data includes personal information,** Management method, security Liabilities of platform operators.<br>Liabilities of data providers and user. |
| **Act on Special** | This act had been enacted in the | This act includes both public and | On energy, industrial | The Act establishes a | **There is no specific incentive** | In case the data contains |

[121] https://www.meti.go.jp/english/press/2018/0615_002.html, https://www.meti.go.jp/press/2019/04/20190404001/20190404001-1.pdf.

[122] https://www.meti.go.jp/english/press/2017/0419_001.html

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| **Measures for Productivity Improvement, 2018[123]** | backdrop of Japan's economic policy of 2017, which aimed at attracting investment and facing international competition, and increasing productivity in the IoT, big data, and artificial intelligence.<br>Notably, the provision under this act are subject to the Basic Act on the Advancement of Public and Private Sector Data Utilization[124] and Act on the Protection of Personal Information | private sector information (excluding information that is likely to damage national security, hinder the maintenance of public order, or be an obstacle to the protection of public safety) | machine, and logistics and to solve social problems like accident prevention, energy management | certification system for business plans that aim at data sharing or collaboration, allowing certified business operators to take advantage of tax breaks and other measures for investing in facilities, equipment, and so on used for efforts stipulated under the Act. In addition, the Act is to establish new procedures through which data sharing business operators who receive confirmation in terms of predetermined levels of cybersecurity, are eligible to request that the government, independent administrative agencies and other public entities provide them with necessary data. | **structure specified in the Act.**<br>**However, the Act proposes to give tax breaks to business operators who are certified and make a plan for innovative data use.** | personal information as under the Act on the Protection of Personal Information, the minister and authority concerned will examine the application appropriately and liaison with the Personal Information Protection Commission. It will also examine the necessity of prompting such use of information |
| **Netherlands** | | | | | | |
| **Dutch Digitalisation Strategy: Dutch Vision on Data Sharing Between** | **The strategy recognises that data is a resource for the 21st century and its re-use and sharing will benefit businesses.** | The strategy **covers personal, non-personal, and data generated out of pieces of equipment** and recognizes that such data can be | The strategy covers data **sharing for innovation and increasing competition**.<br>It also recognizes that **compulsory** | **The strategy first and foremost encourages voluntary data sharing based on the principles of FAIR (data** | This will be determined through contractual agreements between the businesses agreeing to share the data. | **The strategy specifies that while sharing data the rights and obligations must be clearly specified-**<br>- **Sharing of personal** |

[123] https://www.meti.go.jp/english/press/2018_06/0606_001_00.html

[124] http://www.japaneselawtranslation.go.jp/law/detail/?id=2975&vm=02&re=

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| **Businesses 2019**[125] | **However, it recognises that the government can play a role in this if the markets themselves have failed to do so and to reduce the risk of privacy breaches and ensuring cybersecurity in data sharing. Additionally, the strategy is inspired by the analysis of the used cases of data sharing in the Netherlands following different arrangements and principles.** | shared amongst businesses with proper compliance and agreements. | **data sharing may be introduced for sharing of data for public interest** such as competition, freedom of choice, innovation, good health or free-flowing traffic, and a green economy. | **must be findable, accessible, interoperable, and reusable)** through sets of agreements between parties and common technical principles. The government may facilitate such sharing through proper infrastructure.<br>The strategy recognises the need for **mandatory data sharing only for public interest purposes** when data cannot be easily produced or gathered; it is not possible to make appropriate sharing agreements; and such an obligation would not reduce the incentive for innovation, consequences for intellectual property and necessity to obtain the consent of the data subject.[126] | The strategy recommends that such sharing agreement must **specify the intellectual property clauses, trade secrets, ownership of data within such agreement.** In such cases, the government will only play a facilitator's role.<br>Even for the cases where mandatory sharing may be proposed the strategy suggests that due attention needs to be given to its effect on intellectual property. | **data should comply with the GDPR**<br>- **Frameworks related to consumer law where relevant must also apply** |
| | | | **United Kingdom** | | | |
| **National Data** | This strategy looks at how to leverage existing | The strategy refers to data as information about | They have identified five concrete and | The strategy does not recognise any | **While no definite valuation mechanisms are** | **The strategy to build on the Data Ethics** |

[125] https://www.government.nl/documents/reports/2019/02/01/dutch-vision-on-data-sharing-between-businesses

[126] The strategy prescribes for a decision tree under which the government will first as ask : Does data sharing offer opportunities in regard to (for example) productivity and innovation, competition and choice, or societal challenges?-Will data sharing take place in markets and communities even if the government does not take a role?- Could private data sharing come about with targeted financial and/or organisational assistance? And then decide on its role.

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| **Strategy 2020[127] ( Under Consultation)** | UK strengths to boost better use of data across businesses, government, civil society, and individuals. The strategy focuses on using data to deliver new and innovative services, promote stronger competition, and better prices and choices for consumers and small businesses.<br>This strategy comes at the backdrop of **used cases of data sharing by private companies and amongst various sectors, which has also inspired the parameters and focus of this strategy.** Moreover, the strategy also notes that the government has considerably invested in research and partnered with organisations with expertise in the field to develop and test models of data sharing. | people, things, and systems, which means it includes both personal and non-personal data. | significant opportunities for data to positively transform the UK in the following domains:<br>1. Boosting productivity and trade<br>2. Supporting new businesses and jobs<br>3. Increasing the speed, efficiency, and scope of scientific research<br>4. Driving better delivery of policy and public services<br>5. Creating a fairer society for all. | concrete mechanism for governance and proposes for exploration of government as a collaborator, steward, customer, provider, funder, regulator, and legislator. The strategy of open for consultation and proposed to get views of the stakeholder on the kind of government intervention that might be apt.<br>It specifically notes that mechanisms to make the data available should ensure that an appropriate balance is struck between maintaining incentives to collect and curate data, and ensuring that data access is broad enough to maximise its value across the economy. | **proposed, the strategy observes that the** aim should be to maintain and bolster a data regime that is not too burdensome for the average company – one that helps innovators and entrepreneurs to use data legitimately to build and expand their businesses, without undue regulatory uncertainty or risk in the UK and globally. | **Framework published by the government and ensure to maintain transparency in the AI use of data.**<br>It also aims to ensure that any governance model would ensure the privacy of consumers and the intellectual property of businesses. |
| **UK AI Sector Deal (Data Sharing Infrastructure)[128][129]** | This Sector Deal sets out actions to promote the adoption and use of AI in the UK, and delivers on | It includes both personal and non-personal data. Although, in the case of personal data consent need | No specific purpose for setting up data trusts has been identified. | The AI Sector Deal proposed a data trust model for un-tapping the data sets from both | In the pilots conducted, broadly the incentive to contribute to the data trust rested | The proposed data trusts have to comply with rules and regulations concerning |

[127] https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy#data-1-3

[128] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/702810/180425_BEIS_AI_Sector_Deal__4_.pdf

[129] https://docs.google.com/document/d/118RqyUAWP3WIyyCO4iLUT3oOobnYJGibEhspr2v87jg/edit#

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | the recommendations of the independent AI review, 'Growing the AI industry in the UK'. The strategy proposed for setting up of data trust to tap on datasets help by the public and private sector.<br>It was also pointed by the report published by the UK Digital Competition Expert Panel - Unlocking digital competition report', which identified that increasing access to data – potentially through data trusts – can be a regulatory tool to improve competition | to be taken along with appropriately informing the use of how his/her data will be used. Alternatively, such data could be anonymised or aggregated. | However, three pilots have been initiated with –<br>• the Greater London Authority and the Royal Borough of Greenwich to explore the creation of a data trust in an **urban space**, focusing on data about electric vehicle parking spaces and data collected by heating sensors in residential housing.<br>• WILDLABS Tech Hub to explore the creation of a data trust to tackle the **international illegal wildlife trade**, focusing on image and acoustic data, and data acquired by officials at borders.<br>• food and drink manufacturers and retailers to explore the creation of a data trust to tackle **global food waste**, focusing on food waste and sales da | public and private sectors. Data trust is defined as a legal structure that provides independent stewardship of data. Under this kind of data trust, there are independent collaborations or organisations, which become stewards of data. A data trust can decide who can access the data and for what purpose.<br>This was piloted in three sectors in Europe to consider the viability of the system | in – delegate data steward responsibilities i.e. costs related to sharing of data goes to the data trusts, data trusts then also become responsible for mediating between prospective data users, data trusts would also engage with citizens and consumers, sharing data might create more efficiency in products, services, and supply chains, reputational benefits for companies for giving some data and enhance consumer trust, financial returns as data trust can be designed in a way to create remuneration and responsibility on trust for compliance of regulation.<br>In its design, the data trust proposes for model through which data holders can make arrangements with data trusts on incentive structures. Additionally, intellectual property rights in the data will be licensed or transferred based on an agreement between data holders and data trusts. | privacy, however in the case of no legal rule 'consent of the governed' would be the norm to be followed by the data trust authority. |
| **Sectoral Data Sharing Frameworks** | | | | | | |

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | | | **European Union (EU)** | | | |
| **Payment Services Directive 2015** [130] | The directive stipulates rules for sharing customer's payment data across service providers.<br>This Directive aims to ensure continuity in the market, enabling existing and new service providers, regardless of the business model applied by them, to offer their services with a clear and harmonised regulatory framework. | Consumers and companies using payment services will have **to grant access to their payment data to third parties providing payments-related services (TPPs). These are, for example, payment initiation service providers (PISPs) and account information service providers (AISPs).**<br>PSD2 regulates the provision of new payment services which require access to the payment service user´s data. For instance, this could mean initiating a payment from the customer's account or aggregating the information on one or multiple payment accounts held with one or more payment service providers for personal finance management. | It requires banks to maintain an infrastructure through which customers can transfer their payments data between different service providers other than banks | Banks will be required to build **application programming interfaces (APIs) — sets of code that give third parties secure access to their back-end data.** | **No incentive model has been specified.** | The directive establishes rules to provide more flexibility and freedom to customers regarding their payment data. **They can make their data available to third-party service providers – who must also, meet supervisory and security requirements - while maintaining the confidentiality of these data**. The directive prescribes conditions for – explicit consent, users have personalised security credentials, purpose limitations. |
| **Commission Delegated Regulation (EU) No 886/2013 for data and procedures for the** | This directive aims for the traffic -data to be made easily available for exchange and reuse for the provision of information | **It covers data on – Slippery roads, animals on the road, accident area, road works, reduced visibility,** | The main purpose of such data sharing would include giving real-time access to the public regarding road safety. | **The Member States shall manage a national access point to the data, which regroups the access points established by** | All the data under this directive is to be provided free of charge to the end-users.<br>However, for sharing real-time traffic data with | It provides for the protection of personal data through compliance with existing regulations on personal data. In |

[130] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015L2366&from=EN

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| **provision, where possible, of road safety-related minimum universal traffic information free of charge to users**[131] | services, public and/or private road operators, and service providers. | **blockage of the road, etc.**<br>**It will include data from both public and private road operators.** | | **public and/or private road operators and/or service providers operating on their territory**. These data shall be accessible for exchange and reuse by any user of road safety-related minimum universal traffic information:(a) on a non-discriminatory basis;(b) within the Union irrespective of the Member State of establishment; `(c) in accordance with access rights and procedures defined in Directive 2003/98/EC;(d ) within a timeframe that ensures the timely provision of the information service;(e) through the national access point.<br>The provision under the real-time traffic data sharing with public authorities for increasing efficiency of their systems could be determined on contractual terms without prejudice to this | public authorities or other service providers, private operators may enter into a contract that can define remuneration and terms of use. | this case, it will be the GDPR. |

[131] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013R0886&from=EN

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | | | | directive. Similarly, they may also enter into a commercial contract with other service providers . | | |
| **EU Code of conduct on agricultural data sharing by contractual agreement 2018**[132] | This framework **provides non-binding guidelines for contractual agreements for agricultural data sharing in the EU**. The framework recognises that while data sharing can bring greater efficiency in the agricultural sector, the issues surrounding data protection, ownership, and intellectual property need to be addressed appropriately. To this end, the framework gives guidelines on what components are to be considered while formulating data-sharing contracts. | It includes a code of conduct for **personal data, anonymised data, publically available data, raw data, metadata, primary data, and aggregated data**.<br>The right to determine who can access and use the data is attributed to the **Data operator, who is the person or entity that can claim an exclusive right to license the data. It is this person which has collected/created this data** either by technical means or has commissioned data providers for this purpose.<br>**This does not include data that is aggregated, but provisions for such data should also be included within the agreement. For instance, the rights regarding data produced on the farm or during farming** | There is no specified purpose that is prescribed for sharing, however, **the code indicates that the purpose of using the data must be specified in the contract of data sharing**. | **The guidelines specify important terms of contract which should include- important terms, the purpose of collecting, sharing, and processing of data rights and obligations of the parties related to data sharing**, security, storage software, or applications used in storage and use of data **v**erification mechanism for the data originator, transparent mechanisms for adding new or future users.<br>Data originators should also have the right to transmit data to another user. | The framework states licensing conditions under the contract **should adequately protect the IPR of the parties in the data value chains by specifying licensing conditions.** | The code of conduct specifies appropriate requirements for anonymization and pseudonymisation for personal data, and it recommends that if the data is used to decide the data originator the GDPR will apply. |

[132] https://www.ecpa.eu/sites/default/files/documents/AgriDataSharingCoC_2018.pdf

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | | **operations are granted to ("owned by") the farmer and may be used extensively by them**.<br>The parties (originator, provider, user, the third party should establish a contract clearly specifying conditions for data collection and sharing. | | | | |
| **Finland** | | | | | | |
| **Act on the Secondary Use of Health and Social Data, Finland 2019[133] (the Act)** | The objective of this act is to facilitate effective and safe processing and access to personal social and health data for steering, supervision, research, statistics, and development in the health and social sector. A second objective is to guarantee an individual's legitimate expectations as well as their rights and freedoms when processing personal data. | The Act stipulates for the following kinds of data to be shared -<br>• data from several different controllers are combined<br>• the register data originates from private social welfare and health care service providers<br>• the data is stored in Kanta services (database of medical records and other related information).[134]<br>**All the data is to be anonymised or pseudonymised** | The **data permit requests are required to stipulate the purposes of data sharing, data utilisation plan** and after the assessment of such purposes with the authority grants data permits.<br>Along with this they also have to specify what controller of data they want to target. | **The Act stipulates the creation of the Health and Social Data Permit Authority (FinData).**<br>The Authority gives access to data after permit requests are made and processed by it. If the permit is processed they gather data from a controller or request from a private service provider and then combine, pseudonymise, and anonymise the data or produce statistical data converting and combining the permit holder's data. | Pricing of the processing permit request includes the costs of -<br>1. Fee for Findata for data request or data permit<br>2. Costs incurred by data controllers for the extraction and delivery of data, based on each controller's regulations<br>3. Working hours used by Findata for combining, pre-processing, pseudonymising, and anonymisation the data<br>4. Remote access environment charge for data permit holders. | The Act requires compliance with GDPR |

[133] https://stm.fi/documents/1271139/1365571/The+Act+on+the+Secondary+Use+of+Health+and+Social+Data/a2bca08c-d067-3e54-45d1-18096de0ed76/The+Act+on+the+Secondary+Use+of+Health+and+Social+Data.pdf

[134] https://www.kanta.fi/en/what-are-kanta-services

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| **Common rules for the internal market for electricity and amending directive 2019**[135] | EU has implemented the policies and directives for the internal electricity market since 1999. It had 'A Framework Strategy for a Resilient Energy Union with a Forward-Looking Climate Change Policy' followed by 'Launching the public consultation process on a new energy market design'.<br>Through such policies and initiatives over the years this directive intends to create an internal electricity market, Member States should foster the integration of their national markets and cooperation among system operators at the Union and regional level, and incorporate isolated systems that form electricity islands that persist in the Union. | For the Directive, **data means 'data of the final customer' and 'include[s] metering and consumption data as well as data required for customer switching, demand response and other services** | The data access would be given to eligible parties which will be decided by the competent authority.<br>It further states that list of eligible parties would have to be specified by the Member States and would at least include 'customers, suppliers, transmission and distribution system operators, aggregators, energy service companies, and other parties which provide energy or other services to customers<br>However, **the eligibility requirements are purposes for accessing the data are not laid out**. | The Member States or, where a Member State has so provided, the designated competent authorities, **shall authorise and certify or, where applicable, supervise the parties responsible for the data management, to ensure that they comply with the requirements of this Directives.**<br>The Member State shall 'organise the management of data to ensure efficient data access and exchange'. Access to data shall be granted in a 'non-discriminatory' manner among the eligible parties. It shall be 'easy and the relevant procedures for obtaining access to data shall be made publicly available. | The **price for accessing data shall be regulated by the Member States, but shall, in any case, be 'reasonable and duly justified**. This is only applicable to eligible parties.<br>Data to the customers is to be provided free of charge. | The directive stipulates for the Commission to adopt, through implementing acts, interoperability requirements and non-discriminatory and transparent procedures for access to data. |
| **Sectoral Level Framework/ Initiatives/Strategies for Data Sharing** | | | | | | |
| **International/ Global Initiatives** | | | | | | |
| **Dawex**[136] | Dawex Data Exchange and global marketplace allow users to deploy free or | This global data marketplace **hosts all kinds of data aggregated data missed datasets etc**. | The users of the market places are free to set the purpose of usage conditions on the data. The | **It's an open marketplace, where data can be monetised, shared** | The marketplace can be joined for free, however, the valuation of the data will have to be determined by | To secure your data exchanges beyond national borders, Dawex has chosen to obtain |

[135] https://eur-lex.europa.eu/eli/dir/2019/944/oj

[136] https://www.dawex.com/en/

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | monetized business models and multiple use cases including internal data exchange, data sourcing, free data sharing, open data, data monetization, and data marketplace orchestration between customers, suppliers, partners, subsidiaries, and many other organizations.<br><br>They note the necessity of such marketplace on account of –<br>Many organisations and companies are already launching specialised marketplaces in different regions<br>Governments are supporting such initiatives<br>Governments are also adopting regulations such as GDPR and other data flow regulations<br>Associations are already building new forms of trust data sharing models. | However, all the data is encrypted and is hosted at servers closest to the location of the organisation - North America, South America, Europe, or Asia with technical infrastructure meeting the strictest worldwide standards. | marketplace also provides pre-set contracts for this.<br><br>While the marketplace caters to all industries there are specific focus industries stipulated – Agriculture, Automative, Bank Insurance and Financial Services, Energy, Retail and Consumer Goods, Health, Environment, Media and Entertainment, Public Sector, Shipping and Logistics, Tourism and Sports. | **according to specific business models of organisations/ companies**. | the users themselves.<br><br>There are different kinds of packages available on the platforms for increasing the valuation and making data visible to more people. –<br>Community- Free joining of the marketplace<br>Business – fee per month<br>Enterprise - customised pricing<br>Regarding data usage rights between parties licensing contracts could be set-up. | certification from independent data protection authorities.<br><br>They follow the Privacy by Design concept in their marketplace.<br><br>They ensure compliance with GDPR and help their customers comply as well. |
| **International Data Spaces Association** [137] | International Data Spaces is run by International Data Spaces Association via a European non-profit, which takes an active part in designing a trustworthy | It includes **all kinds of data including both personal and non-personal data, however, IDS adheres to European principles of privacy and data security**. | IDSA is suitable for almost every industry. The orientation of its members is wide-ranging, from medium-sized businesses to multi-corporate enterprises: from urban data space | The data provider – i.e. the company – determines who may use the data and how to use them. As a result, partners in a value chain can individually or jointly access | Each business is free to propose its valuation and pricing models. | **Data security and data sovereignty are the essential features of Industrial Data Spaces**.<br><br>Data owners can always keep control over |

[137] https://www.internationaldataspaces.org/our-approach/#about-us

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | architecture for the data economy.<br>More than 101 companies and institutions of various industries and sizes from 20, global acting medium-sized companies, software, and system houses are members of the association.<br>The IDSA aims to guarantee data sovereignty by an open, vendor-independent architecture for a peer-to-peer network that provides user control of data from all domains | | to material data space, medical data space, mobility data space, etc<br>For the exchange of data IDSA architecture creates different roles for different parties which include - Data Provider, Data User, Data Broker. | certain data by mutual agreement to start something new, develop new business models, design their processes more efficiently, or otherwise initiate additional value creation processes.<br>**Each participant and each component in this network is certified and can be identified as a conclusive identity. Certification prescribes and verifies the implementation of generally accepted safety standards and mechanisms. The participants in the data space are obliged to observe both the general rules for dealing with each other and the data usage guidelines specified by the data providers**. IDS provides technologies to implement and control this at a technical level (usage enforcement) | | their data and can also fulfil their standards of data security. The data are exchanged safely on demand if they are requested by certified, trustworthy partners.<br>The main feature of the International Data Spaces is that data providers – i.e. companies that want to make their data available for digital services – can always keep control over their data and enforce their standards of data security (keyword: “Privacy Enforcement”).<br>The data remain with their provider and are exchanged securely on demand. They are only exchanged if they are requested by certified, trustworthy partners. If necessary, the data themselves are not exchanged, but analysis procedures are applied to the data. |
| **Netherlands** | | | | | | |

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| **iShare**[138] | The iSHARE project is an initiative of the Neutral Logistics Information Platform (NLIP), which is the leading platform promoting data exchange in the transport and logistics sector and part of the Netherlands' Logistics Top Sector programme.<br>The iSHARE uniform set of agreements for identification, authentication, and authorisation enables everyone to share data with everyone else in the logistics sector in a simple and controlled way – including with new and hitherto unknown partners. Through iSHARE, NLIP is keen to eliminate data-sharing barriers, stimulate supply chain collaboration and scale-up, accelerate and successfully connect existing digital data-exchange initiatives. This initiative has been supported by relevant Dutch Ministries. | iSHARE is developed in conjunction with organisations that represent a cross-section of the sector: all modalities, organisations of all shapes and sizes, public-sector and private-sector organisations, data providers/data recipients, and their software suppliers.<br>Before becoming part of the iSHARE platform, the organisation requires the companies to sign standardised **agreements for data sharing in which type of data to be shared, with whom it is to be shared and licensing terms are specified.**<br>**Once the organisation/company is issued an iSHARE identity they can share and access data through data hubs organised by iSHARE** | The participants in the scheme – which includes more than 20 public and private organisations – focus on how to share information as effectively as possible. **By building agreements and standards together, they have created an atmosphere of trust.**<br>The conditions for data use are recorded in the agreements system. The data owner's authorization specifies the purpose and the conditions under which his or her data can be used.<br>Some of the beneficiary categories which have been identified include –<br>Freight Forwarders<br>Platforms<br>Shippers<br>Software Suppliers<br>Transport Companies. | Once an organisation has an iSHARE identity they can use it to authorise the data hub to release data to third parties. In the iSHARE authorization, you specify which party is permitted to access which data. If the situation changes, you can withdraw or modify your authorization.<br>Through the data hub, all parties and organisations then have digital access to the data of the owner and also to that of many other contracting parties.<br>A precondition is that they also have an iSHARE identity. A machine-to-machine link, for example in the form of an API, is also required to receive the right data rapidly, securely, and entirely automatically. | These conditions may be stipulated in the contracts, however, no explicit incentive or valuation of data has been prescribed. | **The iSHARE agreements ensure compliance with the GDPR and other applicable legal obligations.**<br>It also gives complete control of the data to the owner and they can withdraw from sharing at any time. |
| **South Africa** | | | | | | |
| **Biodiversity** | The South African National | SANBI was mandated to | Balancing the interests of open | South Africa was one of the | The demand for data is mainly for | The policy took shape by |

[138] https://www.ishareworks.org/en/ishare

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| **Information Policy Framework** [139]**/ SANBI Data Sharing Agreement** [140] | Biodiversity Institute was established under the National Environmental Management Act, 2004. The model data-sharing agreement between SANBI and its partners was introduced in 2018. | collect, generate processes, coordinate and disseminate information about biodiversity and sustainable use of indigenous biological resources and maintain databases. To help achieve that mandate and meet the demands of international partners like UNEP, the agreement was put forward to share data strategically with its partners. | access to data to increase the quality and efficiency of research and innovation with the need for restriction of access in some instances to protect social, scientific, and economic interests is the purpose of this framework. This framework will lead to enhanced biodiversity research in both the public and private sectors. | first countries to join the open access to data initiative as far back as 2000 and introduced the Promotion of Access to Information Act. The Act ensured that all publicly funded institutions are legally bound to make their data accessible. Over time, in 2010, the SANBI Biodiversity Information Policy Framework was developed, which strives to ensure easy access to information whilst simultaneously providing protection to sensitive data and maintaining intellectual property rights. | research and policy purposes. The research is conducted by industry players, governments, and civil society. The framework and agreement, therefore, serve a multifunctional role of bringing transparency along with data sharing. | building upon the open government policy adopted by the South African government in the early 2000s. This followed by several other policies at both national and regional levels based on the demand were brought forward. Eventually, the policies evolved and contributed to forming a national framework to share data on biodiversity based on set standards. |
| **Ethiopia** | | | | | | |
| Agronomy and Soil Data Sharing Policy, 2020[141] | Agriculture remains the least digitised sector across developing countries. And while the open data policy has been proposed everywhere, including by FAO and UN, many key partners don't share their data. Based on this, the Ethiopian Ministry of | Under this data sharing strategy, government, industrial farms, small farmers, fertiliser suppliers and producers, seed suppliers, local agricultural traders, agro-exporters as well as agro researchers will be directly affected. | Being a predominantly agrarian economy, the government has decided to introduce policies that improve the agricultural outcomes of the county. As a part of the larger Agriculture Extension Strategy introduced in 2017, the | A civil society-led "coalition of the willing" (CoW) created by soil and agronomy experts eager to share their data, or support data access. The mechanisms of governance of data sharing policy will also be done by this coalition in | Exports are almost entirely agricultural commodities, and coffee is the largest foreign exchange earner for Ethiopia. To that extent, the government seeks to increase and expand its diverse agricultural market. The value creation is therefore | Inspired by the moves from the civil society, the Ministry of Agriculture established a national task force to develop a soil and agronomy data-sharing policy for Ethiopia. The task force developed data-sharing guidelines and a |

[139] biodiversityadvisor.sanbi.org/wp-content/uploads/2012/09/Biodiversity-Information-Policy-Framework-Principles-Guidelines.pdf

[140] http://biodiversityadvisor.sanbi.org/wp-content/uploads/2018/01/2.DataSharingV2.pdf

[141] New Ethiopian Ministry of Agriculture data sharing policy supported by WLE/CIAT and GIZ to improve food production while building landscape health | Water, Land and Ecosystems (cgiar.org)

| Parameters for Synthesis | Description | Scope of Data Covered and Stakeholders Affected | Purposes of sharing and expectation of value creation | Mechanisms of Governance | Incentives and valuation of data | Checks and Balances |
|---|---|---|---|---|---|---|
| | Agriculture established a national task force to develop a soil and agronomy data-sharing policy for Ethiopia. | | government had been exploring policies to improve the agricultural outputs of the country.[142] | partnership with the government and with international aid agencies. | expected in form of increased net exports. | way forward for the CoW based on the evidence presented by the civil society and the CoW.[143] A draft was presented at several CoW meetings with a finalized policy launched in June 2019. |

[142] 51050623-b954-46cf-bea3-aaefece29408 (moa.gov.et)

[143] studySummary.do (cgiar.org)

## About the Project

Globally, initiatives are being launched to explore frameworks, principles, codes, and mechanisms for sharing non-personal data (NPD). India has also taken a step in this direction, and a committee of experts on the NPD governance framework have recently released its report. One of its key recommendations pertains to sharing NPD for spurring innovation and enabling digital economy growth, with appropriate safeguards in place.

Despite being well-intentioned, the hypothesis, rationale, assumptions approach and recommendations regarding sharing of NPD appears to be ambiguous and inconsistent with the broader objective of the committee.

Since discussions around NPD sharing framework in India are in their formative stage, there is a need to examine the issues dispassionately, question the assumptions underlying the recommendations, and consider appropriate evidence by taking a comparative and multi-stakeholder perspective.

To this end, Consumer Unity & Trust Society (CUTS) is undertaking a study to assess and question the rationale and assumption of the report of Kris Gopalakrishnan Committee on the NPD sharing framework and analyse its recommendation approach and presenting a multi-stakeholder perspective in the Indian context.

For more, please visit:
*https://cuts-ccier.org/npd/*

## CUTS International

Established in 1983, CUTS International (Consumer Unity & Trust Society) is a non-governmental organisation, engaged in consumer sovereignty in the framework of social justice and economic equality and environmental balance, within and across borders. More information about the organisation and its centres can be accessed here: http://www.cuts-international.org.

CUTS®
International

D-217, Bhaskar Marg, Bani Park, Jaipur 302 016, India
Ph: 91.141.228 2821, Fax: 91.141.228 2485
Email: cuts@cuts.org, Website: www.cuts-international.org

Also at Delhi, Kolkata and Chittorgarh (India); Lusaka (Zambia); Nairobi (Kenya); Accra (Ghana); Hanoi (Vietnam); Geneva (Switzerland) and Washington DC (USA).