Navigating the Puzzle of Non-Personal Data Sharing

A Three-Pronged Analysis of Rationale and Assumptions

CUTS International

#2109

Navigating the Puzzle of Non-Personal Data Sharing

Three-Pronged Analysis of Rationale and Assumptions

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), 'Navigating the Puzzle of Non-Personal Data Sharing: *Three-Pronged Analysis of Rationale and Assumptions*,' 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 one of the two-part CUTS project entitled, 'Examining the Rationale, Assumptions and Approaches to Non-Personal Data Sharing'.

#2109

Table of Contents

List of Abbreviations

AI	Artificial Intelligence
BSA	Baseline Scenario Assessment
B2B	Business-to-Business
CoE	Committee of Experts
CCI	Competition Commission of India
CUTS International	Consumer Unity & Trust Society
DPA	Data Protection Authority
FATEN Principles	Fairness, Accountability and Autonomy, Trust and Transparency, Equity and Efficiency, Non-Maleficence
EU	European Union
FAIR	Findability, accessibility, interoperability and reusability
GDPR	General Data Protection Regulation
HVD	High Valued Dataset
JPC	Joint Parliamentary Committee
NPD	Non-personal data
NPDA	Non-Personal Data Authority
NDSAP	National Data Sharing and Accessibility Policy
PDP Bill	Personal Data Protection Bill 2019
PII	Personal Identifiable Information
RIA	Regulatory Impact Assessment
The Report	Report on Non-Personal Data Governance Framework
NODE	Strategy for National Open Digital Ecosystems
UK	United Kingdom
US	United States of America

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

Executive Summary

Globally, initiatives for sharing non-personal data (NPD) are being explored to formulate frameworks, principles, codes, and mechanisms on different aspects such as: upholding data rights, ensuring trust, fostering fair competition and innovation, protecting consumers, preserving intellectual property, promoting research, enabling legitimate aims of the state, and spurring economic growth.[1]

These initiatives are being led both by the industry and the government. India has also taken a step forward in this direction and thus constituted a Committee of Experts (CoE) on the NPD Governance Framework, which has recently released its Report (the Report).[2]

One of the key rationales of the Report pertains to unlocking the economic value of data for 'public interest purposes' and establishing community rights in data. In this regard, the Report recommends sharing NPD for spurring innovation with appropriate safeguards. Despite having the right intent, there seems to be a broken linkage between the rationale of the Report and pieces of evidence, assumptions, observations pertaining to market failures; regulatory gaps; nature of data; linkage between data access and sharing; intended economic and societal benefits; infrastructure and policy maturity to adopt the proposed data sharing practices.[3]

This raises questions about its perceived objective and risks of unintended consequences and costs. It is therefore vital to dispassionately examine the NPD sharing mechanisms proposed by the Report.

To this end, this study undertaken by Consumer Unity & Trust Society (CUTS International) offers an analysis of the Report in two phases. In this report of the first phase, the report has been dissected from three analytical perspectives as illustrated below:

---

1 For instance, see Analytical report on EU law applicable to sharing of non-personal data, https://eudatasharing.eu/sites/default/files/2020-02/EN_AR%20on%20EU%20law%20applicable%20to%20sharing%20of%20non-personal%20data.pdf

2 https://static.mygov.in/rest/s3fs-public/mygov_160975438978977151.pdf

3 https://cuts-ccier.org/pdf/comments-on-revised_npd-governance-framework.pdf

1) **Baseline Scenario Assessment (BSA):** Through undertaking the BSA, the assumptions behind the rationale, intent, objectives and approaches of the Report have been assessed by conducting an in-depth secondary literature review. The assessment indicates that making assumptions around nature; value and benefits from data sharing; policy maturity and market needs; targeted market and regulated failures should be foregrounded within evidence, market readiness analysis and assessment of infrastructural capacities with clearly defined safeguards. This is paramount as formulating assumptions without clear direction can lead to misalignment of incentives, regulator overreach, exclusion errors, increase in compliance, disincentivising innovation and investments, as indicated from the existing literature.

2) **Comparative Jurisdictional Analysis:** This analysis was undertaken through examining the parameters of - assumption behind rationale and goals; targeted market and regulatory failures; policy maturity; market mapping and needs assessments; and the process followed by nineteen data-sharing frameworks/strategies/policies across various jurisdictions both at umbrella and sector level. This analysis aimed to explore learnings, alternatives and good practices available in other jurisdictions concerning data sharing.

   The comparative analysis highlighted that the intended objectives of data sharing frameworks were closely related to regulatory and market failures. And the assessment of these market failures was derived from conducting impact assessments and extensive stakeholder consultations. Following these assessments of gaps; **the expected value and benefits of data sharing are determined**. In doing so, the countries have followed different approaches that provide unique learnings for India to re-evaluate its approach.

3) **Stakeholder Consultation:** Key informant interviews were conducted with national and international policy experts, legal experts and industry representatives. These consultations aimed to validate our BSA with primary information and develop our understanding of the Indian realities regarding data-sharing. Industry representatives highlighted that while this initiative is in the right direction, details and nuances regarding incentives and balancing costs need to be fleshed out.

   They warned that the industry might not adopt these regulations if certainty is not being provided regarding security, data quality, anonymisation standards, and incentives. Experts also indicated that the first step in making such a policy should be to formulate a clear problem statement and stock existing policies.

Combined reading and assessment of these analytical perspectives presented in this report highlight issues, questions, and nuances that need further deliberation. The CoE needs to go back to the drawing board. In doing so, they must indulge in wider consultation and aim to achieve an appropriate balance between industry and consumer interest. It is also equally important to take stock of existing data-related policies and analyse their lacunae and derive learning instead of proposing yet another set of regulations that carry forward similar issues. The study makes relevant recommendations in this regard, carving out the learnings from these analytical perspectives.

This report, being the output of Phase 1 of our project on Non-Personal Data Sharing in India, focuses on exploring the assumptions behind the rationale, intent, objectives, and approaches taken by the CoE in proposing the regulatory framework for NPD. In the second phase of our project, we have examined the approaches and governance mechanisms of Non-Personal Data Sharing at length, assessing parameters like Scope of Data, Purpose of Sharing, Stakeholder Interactions, Governance Mechanisms, Data Valuation, Accountability, and Consumer Rights from multiple dimensions.

We have also explored stakeholder concerns, security concerns, privacy issues, treatment of high-value datasets, and checks & balances to non-personal data sharing in the second phase report. Therefore, the phase two report asks and explores more fundamental questions on data sharing, while phase one dissects the assumptions of the CoE. Combined, the reports present a holistic and all-encompassing analysis of Non-Personal Data Sharing in India.

Overall, it is pertinent to acknowledge that because of the dynamic nature of the data economy, achieving and enforcing targets will not be easy. A lot of experimentation in this regard is going on in various countries. In light of this, policy measures should develop evidence-based evaluation and assessments that can sustain and progressively evolve.

Context Setting

Over the past few years, India has taken strides in developing policies, strategies, infrastructure, and ideas to foster its digital economy. Many of such initiatives aim towards achieving India's trillion-dollar digital economy goal.[4] These strides are not just limited to India and countries around the globe are actively working towards tapping into the potential of 'all things digital'. In doing so, India and other countries have identified data as an important facet of the digital economy.

To unpack this facet and realise its potential, countries have introduced policy frameworks to regulate its usage and sharing. The OECD and the World Bank[5] have also recognised the value of data reuse and sharing for development and economic advancement. At the same time, they have also focused on ancillary concerns and pre-requisites such as having strong cybersecurity and data protection laws; identifying safeguards for the citizens in data sharing; having appropriate frameworks for classifying data categories, robust open data policies and intellectual property regimes and balancing incentives.

The World Bank, in its Report, has specifically focused on integrating data sharing regulation within policies and the economic environment of the countries.

Following the same trend, the CoE in India released the first version of the Report on Non-Personal Data Governance Framework in 2020. The revised Report was released in 2021, which made attempts to incorporate feedback given on the first Report. The Report aims to regulate non-personal data sharing to facilitate its usage for public interest purposes and establish community rights in data.

While CoE's intention in this regard is laudable as it presents much-needed foresight and direction to think about the value of data for the Indian economy, however, it also brings to the forefront the need to further evaluate assumptions around the rationale of the Report to assess ancillary concerns highlighted by the World Bank and OECD reports and placing them

4 https://www.meity.gov.in/writereaddata/files/india_trillion-dollar_digital_opportunity.pdf
5 https://www.worldbank.org/en/publication/wdr2021

within the context of targeted regulatory and market failures, policy maturity, market mapping and needs assessments.

While this framework is novel in its approach and ideation, the political, legal and economic context within which it has originated is equally important to consider. The focus on regulating NPD and fostering data sharing is presented as new and novel. However, the initiatives to unlock the potential of data were initiated in 2012 with the National Data Sharing and Accessibility Policy (NDSAP), which obligated sharing non-personal and non-sensitive data collected by the public sector in the machine human-readable format. The goal of the policy was to achieve transparency and foster innovation. However, it has been observed that the policy failed to gain pace despite setting up an open government data platform (data.gov.in). Many of the data-rich government departments like the NSSO or ISRO did not leverage the platform.

At the same time, many of the departments which had initiated the data sharing have now stopped doing so, for example, the Central Board of Indirect Taxes and Customs (CBITC) and the Bureau of Indian Standards. Business lobbies have also questioned the inadequacy of the quality of data, rendering the platform futile. Data of entities such as the Open Survey of India also have significant errors in their data. A lot of these concerns also emanated from the conflict of data sharing with the copyright regime.[6]

Many of the limitations in sharing geospatial data, such as cumbersome licensing procedures, have only been addressed very recently after years of negligence,[7] while for other sectors, the gaps remain.

In a similar view, the 2019-20 Economic Survey of India had a chapter that focused on the motto of "data of the people, data by the people and data for the people," emphasising the need of harnessing data with the public sector for innovation, however as highlighted above the public sector data-sharing still seems inadequate in India, due to lack of active participation by government departments, inadequate data quality and copyright and licensing issues.[8]

Another impetus in this direction has been the introduction of the Open API Policy,[9] which aims to create an infrastructure through which software-based resources are openly available

6 'Open Data in India: In a Restrictive Copyright Regime, Voluntary Organisations Pitch in to Make Data Accessible', *Economic and Political Weekly*, 5 June 2015, 7–8, https://www.epw.in/engage/article/voluntary-organisations-india-counteract-states-copyright-regime-open-data.

7 https://pib.gov.in/PressReleseDetail.aspx?PRID=1698073

8 https://www.indiabudget.gov.in/budget2019-20/economicsurvey/doc/vol1chapter/echap04_vol1.pdf

9 https://www.meity.gov.in/writereaddata/files/Open_APIs_19May2015.pdf

to start-ups. The Fintech sector has been trying to leverage this initiative to introduce Data Empowerment and Protection Sharing (DEPA), which provides a framework through which consent-based data sharing could be made functional through APIs.[10]

Similarly, RBI had already announced the setting up of 'Account Aggregator' as consent dashboards through which consumers can choose how their financial data will be managed. However, there have generally been concerns regarding usability, trust, and linkages with India's data protection landscape.[11]

Key technological developments in the past few years have also been focused on Artificial Intelligence (AI) to tap into the potential of data to create services. To this end, NITI Aayog released the National AI Strategy in 2018, after the Government infused Rs. 3,660 crores for developing AI, the internet of things (IoT), machine learning and other quantum computing technologies.

The NITI Ayog paper recommended focusing on health, education, agriculture, and smart cities and proposed setting up an AI Research, Analytics and knowledge Assimilation (AIRAWAT) platform, a cloud computing infrastructure. These initiatives have been novel; however, the adoption of AI in India is still low as complimenting data protection frameworks and cybersecurity are not adequately developed.

Along with this, cloud computing infrastructure is just starting to develop and there is no integration of ethics with AI and ML systems. The quality of data is poor and it lacks the professional capacity to develop these systems.[12]

Many of these challenges are also relevant in data-sharing as AI development relies significantly on the availability of data-sharing infrastructures. The most recent development has been the setting up of the National AI Portal,[13] which aims to provide resources and a holistic overview of developments related to AI in India. India also launched the Indian Urban Data Exchange;[14] however, we are still to get results from the pilots of this initiative. Learning from such initiatives and policies is important to be integrated into the umbrella policy frameworks.

10 https://pn.ispirt.in/a-great-leap-forward-to-transform-fintech-data-empowerment/
11 http://financelawpolicy.umich.edu/files/raghavan-singh-regulation-of-information-flows-as-central-bank-functions-implications-from-treatment-account-aggregators-india.pdf
12 https://indiaai.gov.in/article/five-challenges-for-ai-adoption-in-india-and-what-are-we-doing-about-them
13 https://indiaai.gov.in/
14 http://www.rbccps.org/wp-content/uploads/2018/12/The-Indian-Urban-Data-Exchange.pdf

Along with this, development towards data sharing is also underway in the healthcare sector. Parallelly, states like Telangana[15] and Karnataka[16] have developed fairly advanced open data initiatives and push for developing AI and data sharing infrastructures. In contrast, other states have still not signed off on the open data initiatives to make their data available. The integration and impact assessments of these initiatives could be precious for the NPD sharing framework.

The above context also illustrates that the developments related to data-sharing have been happening in different contexts and sectors. The Personal Data Protection Bill 2019 (PDP Bill) is still in the pipeline, creating parallel and conflicting narratives with the NPD Governance Framework.

The aim of setting this context and overview is to highlight that India is on the path towards building a fostering data-sharing environment. However, this cannot be achieved without addressing the issues from the open data initiatives to the PDP Bill. The Report's rationale needs to be questioned, examined and analysed in light of such issues before developing an umbrella framework for non-personal data sharing. There is a need to unpack further the assumptions which could continue to cause unintended consequences, pushing India back from its goal to build a sustainable digital economy.

At the same time, it is important to think about the functionality of the NPD Governance Framework in the context of challenges of infrastructure, capacity, adequate safeguards, trust, inadequacies of legal regimes already faced by the existing data, AI and other related policies.

[15] https://data.telangana.gov.in/,https://www.livemint.com/Companies/iZKdgU1KkQh4azdSKCZ31O/Telanganas-open-data-policy-to-help-startups-address-pub.html

[16] https://egov.eletsonline.com/2009/12/karnataka-state-data-centre-sharing-common-infrastructure/

Key Learnings

1) **Clear 'problem-statement and objectives' backed by evidence:** It is laudable that the Report intends to foster innovation in the public interest by enabling greater value realisation from NPD. However, the envisaged goalposts signalling fulfillment of such intent are broad and unclear.

   The evidence to link these objectives to rationale, targeted gaps, market need, and policy landscape in India is missing. It should learn from the experiences of superimposing a novel governance structure over weak institutional capacities negating optimal interaction, transparency, and accountability frameworks among citizens, industry, market, and the state.

   This necessitates undertaking a Regulatory Impact Assessment (RIA) of non-personal data governance framework for India, which would require identifying the problem, defining the problem statement unambiguously, establishing linkages with intended objectives, analysis of costs and benefits of different regulatory options (including no-regulation, self-regulation, and co-regulation) on different stakeholders and sectors, and selecting the most appropriate option, costs of which are likely to be significantly outweighed by benefits.

2) **Agile and Flexible Data Sharing Framework:** Data economy is very dynamic, and it evolves and grows fast. Considering that India's digital economy will grow rapidly, the data sharing models should be agile rather than stringent to avoid compliance burden. It should have the flexibility for businesses at the initial stages so that, going forward, technological development and new business practices could be imbibed to scale up its operations.

   The government's objective should be to foster the businesses by providing flexible guidelines rather than imposing stringent regulations for data management. To build such a framework, wider consultation is required with industry stakeholders and small and medium enterprises to develop the required agility.

3) **Identifying Adequate Pre-Requisites:** Data is very different from other resources that existing regulations and infrastructure have dealt with. Therefore, it is imperative to take stock of current open data policies and other sectoral-level data-sharing initiatives, providing important lessons regarding the availability, demand, and usability of data.

   Additionally, it is important to assess the gaps in the technical capacities of start-ups pertaining to data management before suggesting a framework that can stimulate businesses to upgrade. In this regard, we could also be informed by global developments, including the approach taken by other jurisdictions that have adopted the principles of findability, accessibility, interoperability and reusability (FAIR) to ensure quality data.

   These principles should be the starting point of data standardization and the government should help businesses achieve these standards. Along with this, having appropriate data protection, cybersecurity and intellectual property regimes is also equally important.

4) **Adopting Appropriate Policy Sequencing:** Considering the pre-requisites of the data-sharing framework, the starting point should be establishing appropriate mechanisms for open data-sharing policies. It is important first to unlock public sector data to be used for public interest purposes. This is also informed from the experiences of other jurisdictions, which started from developing their open data policies, which also informed them about data usage practices.

   To avoid conflicting narratives around Personal Data Protection Bill and the non-personal data governance framework, one framework should follow the other. As seen in other jurisdictions where a personal data regulation is the first step in regulating any kind of data, it might be beneficial if the committee formulates its approach after the implementation of the proposed Personal Data Protection Bill 2019 (PDP Bill). This will help establish a 'rights-based approach' from which a common understanding of values and principles can be internalised within the data economy.

   Moreover, the next steps in data sharing should be mindful that anonymisation techniques are not a full-proof solution, which convolutes the basis of the categorisation of NPD.

   The CoE must also take a liberal and settled view in the global context on how to treat data and what kind of rights and responsibilities are to be associated with it. Given that several other jurisdictions are defining public data, it's likely that international trade agreements will come with a caveat of liberal use of data in the future. Therefore, the CoE

must ensure that the framework is aligned with other policies that the government is pushing for, such as foreign trade policy, to ensure greater convergence.

5) **Assessing Market Needs and Incentives:** It is important to undertake an assessment of market discrepancies and assess the risk of incentive misalignments which could stifle innovation and dis-incentivise investments, before even considering mandatory data sharing alternative.

   These unintended consequences emerge due to neglect of 'proportionality' between private costs (current and foreseeable) and public benefits, which is vital in understanding the loss of incentive for smaller start-ups. Equally vital is the consideration of incentives for all actors in the data-sharing process, even for the data trustees to maintain sustainable infrastructure and avoid leaning towards private interest due to lack of incentives.

6) **Appropriately understanding of the purpose of NPD Sharing:-** The nature of data is multi-faceted. Its value is lucid. As such, attribution to its nature and value cannot be done in a vacuum. There are multiple factors to consider: proprietary rights in data, power dynamics in data usage, quality of data, the cost involved in data management, and interoperability challenges. To address this, the CoE should try to avoid the "one size fits all" approach.

   It should strive to achieve 'proportionality' between private costs and public benefits, examine the 'necessity' of achieving public benefits through mandatory data sharing while establishing the inability to facilitate sharing otherwise, and ensure 'reasonableness' of achieving such objectives in practice to give a more purposive construction.

   This test can aid in rationalising the scope of public interest and laying down appropriate principles in this regard, which can protect the interests of data principals, communities, and start-ups.

7) **Consumer Welfare and Appropriate Safeguards:** The CoE must consider specifying how "duty of care" should be approached by data custodians and data trustees, which can be private as well as government entities. In this regard, there is a need to have pilot and feasibility studies before establishing clear rules with regard to the fiduciary responsibility of data trustees with appropriate safeguards, certification mechanism, clear mission statements, unbiased representation, and taking into account interoperability and sustainability requirements.

The focus here should be on building trust between the intermediaries, community and consumers, through contractual or other legal safeguards, with clear liability and accountability frameworks. Some of these could be informed from existing data-sharing and data protection practices within the private sector for the public interest. Thus, mapping existing initiatives in the Indian context is vital to understand the gaps and risks involved in data-sharing.

Furthermore, it is vital to keep in mind that there can be risks of exclusion errors by giving a 'public good' treatment to data. Thus, there can be liabilities and risks of exclusion errors. Hence, it is essential to pay due attention to the definition of community, misalignment of incentives, risk of biases with data trustees. In this regard, it is important first to consider the current institutional capacities and the need to learn from our experiences.

8) **Re-assessing Existing Legal Regimes:** It will be beneficial if the CoE re-evaluates its jurisprudential and legal basis of establishing community rights through understanding the differences between traditional notions of 'identifiability' of community with a resource and that related to data. It is also equally important to internalise that the community identification with data is a gradual process. As people will understand privacy rights and the importance of data, communities will emerge on their own and may not necessarily require regulatory stimulation.

To nudge this process, open data practices and voluntary data sharing mechanisms, developed through robust consultative mechanisms could be explored in publicly funded institutions and projects. When determining the data sharing for sovereign purposes, the committee must recognise the importance and applicability of the three-pronged test of proportionality, legality, and necessity that the Supreme Court of India has interpreted in Justice *KS Puttaswamy v. Union of India.*

Additionally, it is important to consider and analyse, the regulatory objective of the Non-Personal Data Authority, as it may overlap with the jurisdiction of the Competition Commission of India (CCI) if it is intended to address the problem of inequitable distribution and of the Data Protection Authority (DPA) in the case where there are no clear boundaries between personal and non-personal data.

At this stage, it might be important to take stock of the regulatory landscape and tools available, specifically sector-level regulations, to examine their applicability within the data economy.

Baseline Scenario Assessment

Introduction

Globally, initiatives for sharing NPD are being explored through frameworks, principles, codes, and mechanisms on different aspects such as: upholding data rights, ensuring trust, fostering fair competition and innovation, protecting consumers, preserving intellectual property, promoting research, enabling legitimate aims of the state, and spurring economic growth.[17] These initiatives are being led by the industry as well as the government. India has also taken a step forward in this direction and thus constituted a CoE on the NPD Governance Framework, which has recently released its Report (the Report).[18]

One of the key rationales of the Report pertains to unlocking the economic value of data for 'public interest purposes' through establishing community rights in data. In this regard, the Report recommends sharing NPD for spurring innovation with appropriate safeguards. Despite having progressive intentions, there seems to be broken linkage between the objectives and rationale of the Report and pieces of evidence, assumptions, observation and the approach taken by the Report regarding market failures; regulatory gaps; nature of data; linkage between data access and sharing and the intended economic and societal benefits; maturity levels in terms of infrastructure and business practices; and policy maturity.[19]

This raises questions about the objectives that are intended to be achieved and the unintended consequences and costs emerging from such a framework. It is, therefore, important to dispassionately examine the proposed NPD framework in India.

To this end, an in-depth assessment is undertaken through secondary research in the form of baseline scenario assessment (BSA) to examine the rationale of the Report i.e., **to come up with a set of recommendations for realising the value of data for the benefit of citizens and communities in India and to leverage its public and economic value.**

This assessment is undertaken in the light of the arguments in the context of –

1. Lack of clarity on the market, regulatory, and government failures that the Report address.
2. Lack of adequate evidence in formulating assumptions around maturity and facets of India's data market and digital economy.
3. Insufficient assessment of data protection, competition, intellectual property regimes, etc. in India to determine the policy maturity, inter-linkages and conflicts.

---

17 For instance, see Analytical report on EU law applicable to sharing of non-personal data, https://eudatasharing.eu/sites/default/files/2020-02/EN_AR%20on%20EU%20law%20applicable%20to%20sharing%20of%20non-personal%20data.pdf

18 https://static.mygov.in/rest/s3fs-public/mygov_160975438978977151.pdf

19 https://cuts-ccier.org/pdf/comments-on-revised_npd-governance-framework.pdf

The BSA (as given below) presents concern in the form of questions to identify gaps, ambiguities, and missing steps in establishing rationale; accounting for evidence, and assessing Indian realities as stated by the Report in making a case for non-personal data sharing and regulation.

1. Assumptions behind the Rationale

1.1 Nature of Data

Details from the arguments as presented in the Report

- Throughout, the Report makes strong claims about data as a resource to be used 'as' public good and 'for' public having social, public and economic value. **The Report picks up from the previous version and looks at data through two lenses, i.e., an economic lens through which it can create value and an informational lens that creates a potential for privacy harm.**[20]

  Thus, the Report asserts that to realise the economic value and protect against potential privacy harms, a regulatory framework is required to enforce community rights in NPD and enable its sharing.[21]

  However, the Report does not adequately support the link of this assertion with the need for regulation. It does not adequately identify a comprehensive problem statement that it wants to address through enabling data sharing for public interest purposes in the same way as other material resources.
- The Report perceives high valued dataset (HVD) (a particular subset of a dataset) to be mandatorily shared as a **public good** for larger public and community interest.[22] A legal justification of this is also emerging from the committee's observation that data should be considered a 'material resource'. Its equitable distribution should be ensured under Article 39 of the Indian Constitution. Here the Report assumes that data can be compared to other material resources such as forests or water to establish community rights in data.

  It seems that Report confuses 'data as public good' and 'data for the public good'. Moreover, the Report identifies data as 'public good' without considering the externalities that might emerge from such interpretation. Furthermore, it assumes that regulatory and technical capacities could be developed to make data available as a public good.
- Alternatively, the Report also has an underlying tone of assumption for data to be considered as a 'public infrastructure' such as roads on which other services

[20] Page 42, Appendix 2 of the Report
[21] Page 6 of the Report
[22] Page 19 of the Report

could be built. However, here again, it assumes that data is comparable to other tangible infrastructures.

- The report also states that data is non-rivalrous, and several organisations can consume its value without degrading its value to the relevant community.[23] Moreover, it states that benefits accruing from sharing the community data must also flow back to the community; thus, in economic, social, and public interest purposes, community data should be shared.[24]

  Here, while the Report considers the non-rivalry as characteristic of NPD, it negates the factoring of non-excludability of such data. Additionally, it assumes that considering data as a public good or other material resources and enforcement of community rights in NPD will benefit the community without any appropriate evidence or analysis of this assertion.

Concerns

- Studies and experts have indicated that data is different from other natural resources such as 'oil' because of how the value is derived from data. Such a balance sheet approach to the economy cannot be applied to data. This is because data itself does not have inherent value and is not finite. It is only one piece in the value generation chain; it is only when organisations process data for insights that value is created.

  Thus, any data regulation that focuses on access/ sharing will not automatically benefit society. On the contrary, it may deter businesses from innovating and finding unique solutions to customer needs, minimal and medium enterprises and start-ups, who may be forced to share the data they collect and process, with great difficulties.[25]

  This necessitates a further examination from the economic lens as proposed by the Report. There is a need to question - is data similar to other economic resources? And, in what form is the value derived from data, specifically in the Indian context?

- Conceptually public goods are non-rivalrous and non-excludable. The form in which NPD currently exists with private companies makes it non-rival and excludable only due to the option of transferability through voluntarily foregoing the existing intellectual property rights and commercial value in data. This makes data a 'club good' or 'impure public good'.[26]

---

23 Page 16 of the Report

24 Page 6 of the Report

25 Dr Michael Mandel, *The Economic Impact of Data: Why Data Is Not Like Oil.* Progressive Policy Institute. 2017.

26 "Are Data More like Oil or Sunlight?," *The Economist*, February 20, 2020, https://www.economist.com/special-report/2020/02/20/are-data-more-like-oil-or-sunlight.

While ensuring the fact that the data that is protected by the Intellectual Property is not treated as a public good, **there is a need for further assessment to evaluate the possibility, requirements, and costs of transitioning data from a 'club good' to a 'public good', which would lead to disincentivising of businesses?**

- Research has also pointed that considering data as a public good should be closely assessed in the light of power dynamics, as giving access to one stakeholder may conflict with the interest of the other, which may create externalities such as increasing risks of misuse of data and affecting stakeholders who in the first place collected such data.

  Therefore, before considering data to fulfill national or public interest purposes, the moral necessity of sharing such data needs to be identified.[27] This also warrants a deeper examination of - **whether and what data should be considered as a 'public good' and in what circumstances?**

  And, if so, what kind of data can be classified as a public good, and how do we separate proprietary and non-public data from public sector 'open' data? And can one size fits all approach be the right way forward when regulating data for 'public good'?

- Similarly, another study examining data as the infrastructure through assessment of three legal regimes i.e., Public Sector Information Directive, Vehicular data sharing and the Electricity directive in Europe, observed that if data sharing is to be mandated, the granularity of 'purpose' of sharing should be determined, instead of taking a one size fits all approach in all cases. It further stated that data is not a commodity and since its value changes with each transfer and reuse, the purposive infrastructure would be beneficial.[28]

  Thus, before assuming data as public infrastructure, it needs to be questioned – what is the purpose of assigning data as a public infrastructure?

- A study while examining the rights over data indicated that factors such as incentives for data collection, market dynamics, and existing intellectual property rights play an important role in establishing the legal right over data.[29]

  These points towards the need for determining - how the legal rights over data exist in the current scenario? Are there any existing proprietary rights in data affecting the excludability factor? If so, will there be conflicts with existing

---

27 Taylor L. 2016 The ethics of big data as a public good: which public? Whose good? Phil. Trans. R. Soc. A 374: 20160126. http://dx.doi.org/10.1098/rsta.2016.0126

28 Charlotte Ducuing, "Data as Infrastructure? A Study of Data Sharing Legal Regimes," *Competition and Regulation in Network Industries* 21, no. 2 (June 1, 2020): 124–42, doi:10.1177/1783591719895390.

29 Stepanov, Ivan. "Introducing a Property Right over Data in the EU: The Data Producer's Right – an Evaluation." *International Review of Law, Computers & Technology* 34, no. 1 (January 2, 2020): 65–86. https://doi.org/10.1080/13600869.2019.1631621.

proprietary claims in data and the envisaged idea of data as an economic good having beneficial ownership for realising its economic, public, and sovereign value?

- Research also indicates that creating commons rights in the resource for making it available for public interest may not transform into equitable access to a resource.[30] In such cases, access to a resource may be limited by access to knowledge, financial and technical resources with a certain section of the community, which may also create exclusion.[31]

  Considering this, it is important to assess how community rights in data translate into deriving public benefits in data? What are the risks involved with such a prescription?

1.2 Value of Data

Details from the arguments as presented in the Report

- The report states that data creates economic value and wealth, apart from social and public value.[32] On this basis, the Report states that regulation is necessary to enforce community rights and create a framework to unlock its economic benefits.[33]

  Here, the Report assumes that regulation is a silver bullet to unlock and internalise the value of data without identifying problems with the existing scenario, establishing market failure, considering the capacities of the community, data trustees and data requestors to leverage the data access such that its value can be realised.

- The report states that data holds value to the community as such, the access to data as a 'public good' for public interest purposes would be useful. However, the Report makes this assumption by equating data with 'material resources' distributed for the common good.[34]

  The Report itself identifies that there are multiple ways to understand and treat data. Here, the Report assumes that a community's understanding of data is similar to other resources, and through data trustees, communities will be able to leverage the value of data.

---

30 Anupam Chander and Madhavi Sunder, "The Romance of the Public Domain," SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, July 7, 2004), https://papers.ssrn.com/abstract=562301.

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

32 Page 5 & Page 39, Appendix 2 of the Report

33 Page 6 of the Report

34 Page 34 of the Report

- It also carries forward the assumption from the previous version that public agencies collect a lot of data; however, much of the required data is accumulated by the private companies, which leads the value of data to be untapped.[35]

  Here, the Report assumes that there is evidence of the untapped value of data in the private sector without considering the value of public sector data in this regard.
- The Report also presents some examples of beneficial uses of data to indicate the value expectation from data. However, most of these cases are hypothetical and may vary considerably depending on the sector.
- The Report also emphasises giving open access to meta-data for spurring innovation**; however, it does not assign an incentive mechanism for sharing meta-data and portrays it just as a hook for attracting interested parties to pay for the data, not considering the value of meta-data itself.**

Concerns

- Various studies and research[36] have pointed out that data does not have inherent value in itself and the quantity of data cannot be equated with the value that could be created from it. The value realisation from data often faces challenges such as a lack of interoperability and quality standards for data.

  Considering this, and as stated in the previous section defining the nature of data into one particular bracket, i.e., public good or public infrastructure, is difficult. Thus the questions of reusability, findability, and access to data become important in determining its value.[37]

  Thus, before finalising the expectation of value creation from data (either economic or public), it will be important to assess the pre-requisites for such value? And, who are the actors involved in such value creation?
- To assess the value expectation from data, it is important first to identify the benefactors of that data. The value of data is very lucid and changes with every reuse and depends on the context of its use. The value of data can also evolve overtimes and its users may also change with changing priorities. In such cases, the assurance of public value derived from data becomes uncertain and

35 Page 41, Appendix 2 of the Report

36 Yuri Demchenko, Wouter Los, and Cees de Laat, "Data As Economic Goods: Definitions, Properties, Challenges, Enabling Technologies For Future Data Markets," no. 2 (2018): 10; "The Economic Value of Data: Discussion Paper," Discussion Paper (London: UK Government, 2018).

37 "Fair Value? Fixing the Data Economy," MIT Technology Review, 2020, https://www.technologyreview.com/2020/12/03/1012797/fair-value-fixing-the-data-economy/.

ambiguous as communities or the data trustees may not envisage prospective data usage.[38]

Additionally, most of this value addition will occur with the private sector. **Thus transferring value from businesses to the community without incentives can be problematic and may create disincentives.**[39]

In such cases, it will also be important to assess – will the data trustees and communities be capacitated to identify the value of data for their communities?

- While the report takes stock of evolving data sharing frameworks,[40] it should also assess the value propositions these frameworks have proposed along with data sharing. Other jurisdictions, such as Singapore, have specifically issued guidelines for determining data value parallel to introducing data-sharing guidelines. The guidelines included identifying the kinds of data available; the costs and benefits of sharing; the restrictions involved in sharing the data; and the stakeholders that may be affected.[41]

  Additionally, the EU, Australia and UK assessed the value of data in terms of the purpose it is utilised for through assessing use cases of data sharing and conducting impact assessment regarding the evolution of data economy in their respective countries in determining their expectations. [42]

  Considering these practices, it would be essential to examine-- what considerations should be taken for assigning value to the data in the Indian context? Are there any use cases in which data has been valued in quantifiable terms? And, who are the stakeholders involved in such value creation? [43]
- Before considering the proposition of data sharing through regulatory means, the government should examine the linkage between access to data and the benefits**. For this, we need the question - is there substantial evidence available to prove that regulation and access to data would lead to enhanced value realisation?**

---

38 James E Short and Steve Todd, "Many Businesses Don't yet Know the Answer to That Question. But Going Forward, Companies Will Need to Develop Greater Expertise at Valuing Their Data Assets," *MIT Sloan Review*, 2017, 5.

39 Puneeth Nagaraj, Varsha Rao, and Dedipyaman Shukla, "Community Rights Over Non-Personal Data: Perspectives from Jurisprudence on Natural Resources," 2021, 27.

40 Page 5 of the Report

41 "Guide to Data-Valuation for Data Sharing" (Singapore: Infocom Media Development Authority, 2019), https://www.imda.gov.sg/-/media/Imda/Files/Programme/Data-Collaborative-Programme/Guide-to-Data-Valuation-for-Data-Sharing.pdf?la=en.

42 https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN ,https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy

43 Frontier Technology Quarterly, Data Economy: *Radical transformation or dystopia*? 2019 https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2016.0126

- Research in the context of meta-data has pointed out that collecting, organising, and presenting meta-data requires investment in technical and human resources, which transcends into financial costs. It also states that not investing in meta-data will have adverse consequences, thus concluding that meta-data may have a similar asset value as data is processed.[44]

  However, the Report undermines these arguments mandating open access to meta-data without any remuneration. **Thus, there is a need to assess the related costs in organising, collecting, and sorting meta-data? Should these costs be added to the value of meta-data while giving open access?**
- We also need to assess further - if there is any, the surety that meta-data access, will in fact, to greater demand in data sharing and may cover the cost of sharing meta-data.
- It has been argued that mandatorily sharing of data for public interest may adversely affect investments in data collection by the business,[45] **thus, we need to closely examine how the expectation and realisation of value from data affects businesses. And, will the presumed value creation incentivise a business to keep collecting and sharing data without giving due consideration to the cost of data collection?**
- An OECD assessment has pointed out that, "many datasets are not of the requisite quality, are not adequately documented or organised, or are of insufficient (or no) interest for use by others". The lack of a common understanding of what quality means in data is a major source of uncertainty among organisations.[46]

  Thus, before proposing value realisation from data, should we not ensure quality standards within the available data to prescribe standards for the same?
- Research has indicated that before sharing data for the public good, the moral necessity of sharing data should be established as it may have adverse effects on the power dynamics in the current environment.[47]

  Hence, we need to assess - on what basis can we ascertain the gaps in data value, which has necessitated the need for regulation for data sharing?

---

44 Jane Greenberg, "Metadata Capital: Raising Awareness, Exploring a New Concept," *Bulletin of the Association for Information Science and Technology* 40, no. 4 (2014): 30–33, doi:https://doi.org/10.1002/bult.2014.1720400412.

45 Mohit Chawdhry, "The Non-Personal Data Report: Do We Really Need Another Data Regulator?," *THE BASTION*, September 21, 2020, https://thebastion.co.in/data-protection/the-non-personal-data-report-do-we-really-need-another-data-regulator/.

46 "Business Models For Sustainable Research Data Repositories" (OECD, 2017). http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/STP/GSF(2017)1/FINAL&docLanguage=En

47 Taylor L. 2016 ,The ethics of big data as a public good: which public? Whose good? Phil. Trans. R. Soc. A 374: 20160126. http://dx.doi.org/10.1098/rsta.2016.0126, the argument that data should be treated as public good should be evaluated and the ethics of sharing must follow-up this argument.

1.3 Benefits of Sharing

Details from the arguments as presented in the Report

- The Report prescribes data sharing for 'public interest purposes'. However, how **'public interest purpose' is defined gives an ostensibly wide mandate to cover all kinds of purposes.** There are no purpose limitations prescribed. Here it seems that under the garb of public interest, the government is mandating sharing amongst private entities.

  The Report here assumes that a fixed meaning can be ascribed to "public interest", such that the benefits will go back to the public without any intended consequences. Moreover, it assumes that an interpretation ascribed to 'public interest' by the data trustee will represent the community.
- **The Report also assumes that data sharing with the government and other stakeholders leading to 'public interest' purposes as envisaged by the Report will motivate data custodians to sustain their data collection and processing mechanisms.**
- The shared Non-Personal Data may be useful for Indian entrepreneurs to develop new and innovative services and products from which citizens may benefit (economic purpose).

  The Report here assumes **that businesses will be able to leverage data sharing opportunities and deliver on the promise of public interest. Thus, the assumption that smaller companies would tap into the value of shared data rather than more prominent players in the market at this stage would be incorrect.**
- Moreover, framing the case for public interest use of data, the Report makes an underlying assumption that currently, the benefits in data are not flowing back to the public and there exists a market failure, which had necessitated the need for public interest regulation. While the report presents examples of used cases of data, these are mere hypotheticals that do not indicate the building blocks, challenges and risks associated with materialising such benefits in data.
- Collective data is needed for the social and public good. For this, the Report suggests data trusts and data trustee frameworks protect the community interest.[48]

  Here, the Report assumes that trust relationships can be established within the data economy in India without considering existing realities and dynamics between different stakeholders.

[48] Page 18-19 of the Report

Concerns

- The 'public interest theory of regulation' has pointed out that the need for regulation for the public interest is based on two underlying assumptions. First, unhindered markets have failed to lead to problems of abuse by monopolies and externalities. Second, governments are benign and capable of correcting these market failures.[49]

  Thus, there is a need for further examination if such assumptions are correct in the Indian context and any evidence to support such assertions.

- Moreover, how 'public interest' has been presented in the Report makes it vague and ambiguous. No fixed meaning has been ascribed to 'public interest' by Indian jurisprudence in other areas such as natural resources, right to information and intellectual property rights, etc. In fact, in the current form, it appears that, under the garb of public interest, mandated sharing of HVD amongst private entities through data trustees would be undertaken for both 'for profit' and 'non-profit purposes'. Thus, 'public interest' for one may not be the public interest for the other.

- While there is no widely accepted definition of "Public Interest", the definition has been derived from judicial precedents. The Supreme Court in the case of *Bihar Public Service Commission v. Saiyed Hussain Abbas Rizwi* referred to the definition of public interest stipulated by Black's Law dictionary i.e. "the expression "public interest", like "public purpose", is not capable of any precise definition. It does not have a rigid meaning, is elastic and takes its colour from the statute in which it occurs, the concept varying with time and state of society and its needs. It also means the public's general welfare that warrants recognition and protection, something in which the public as a whole has a stake.[50]

  Without a concrete and widely recognised definition of the term, ambiguous interpretations carry a high risk of expropriation in the name of public interest, as is exhibited in another case of land expropriation.

  In this case, the court stated that *"[We] must examine these questions very carefully when little Indians lose their small property in the name of mindless acquisition at the instance of the State. Suppose public purpose can be satisfied by not rendering the common person homeless and exploring other avenues of acquisition. In that case, before sanctioning an acquisition, the courts must exercise their power of judicial review and focus their attention on the concept*

---

49 Peter Drahos, "The Regulation of Public Goods," Journal of International Economic Law 7, no. 2 (2004): 321–39, https://www.anu.edu.au/fellows/pdrahos/articles/pdfs/2004regulationpublicgoods.pdf.

50 (2012) 13 SCC 61

of social and economic justice. While examining these questions of public importance, the Courts, especially the Higher Courts, cannot afford to act as mere umpires."[51]

- **These rulings indicate that while no fixed definition can be ascribed to the public interest, the risk of its misuse through expropriating resources can lead to adverse consequences for the very beneficiaries it is intended to help; in the present situation, those beneficiaries are community, smaller start-ups and businesses, and eventually the citizens to whom the data relates to**.

  The court, in the aforementioned matter, also stressed the need for exploring alternative mechanisms to achieve` stated objectives. Thus, it becomes important to question the 'one size fits all' approach to prescribe mandatory data sharing for all kinds of 'public interest purposes'. Here, it is important to question **the costs and benefits and pre-requisites essential to achieve the envisaged benefits of data sharing?**

- Studies, while examining data as commons to be used for the public interest, have highlighted that such interpretation cannot be done in a vacuum, it is important to explore the interaction of collective interest with individual interest in data. This is relevant because they may be in conflict with each other since the unit of data collection is an individual.[52]

  Thus, it is also important to assess – what can be the tensions between data sharing for public interest and individual interests in data? What are the different dimensions of data ownership?

- **We also need to assess - there are any use cases that can establish the benefits of data sharing in India? If so, what are the lessons that can be learned from such use cases like in Australia? Should the public authorities' open data sharing strategy be the first step before NPD sharing framework to release the siloed data**?

- An assessment conducted on the concept of using data for 'public interest purposes' challenged the assumptions that incentives to collect data are a 'given' and that firms will continue to collect data not-withstanding governmental access to such data. The assessment pointed out that an inverse relationship exists between incentives for collecting data and sharing it for governance in some instances. [53]

---

51 Radhy Shyam(D) Thr. Lrs & Others vs. State of U.P. & Others, (2011) 5 SCC 553

52 Patrik Hummel, Matthias Braun, and Peter Dabrock, "Own Data? Ethical Reflections on Data Ownership," *Philosophy & Technology*, June 15, 2020, doi:10.1007/s13347-020-00404-9.

53 Niva Elkin-Koren and Michal S Gal, "The Chilling Effect of Governance-by-Data on Data Markets," *The University of Chicago Law Review*, 2019, 29.

In light of this argument, there is a need to examine whether data for public interest creates a self-applied limitation in collection and processing, leading to a chilling effect in the existing data economy in the Indian context?

- Research has pointed that to use data for good public purposes, it should be shared voluntarily. To make voluntary data sharing work, factors such as resource dependency and incentives that could be created for the stakeholders in the process need to be assessed to create a 'win-win' situation. [54] For this to happen, businesses may undertake initiatives to develop 'collaborations' for sharing to realise what context and situation best suit them, without government interventions.[55]

In light of these findings, it might be essential to the first question if- the private sector will is motivated to share data for the purposes or benefits as prescribed by the Report? If so, should we let the private sector develop collaborations and cross-sector linkages that may help achieve the benefits of sharing?

- The benefits of data sharing are dependent on various factors and also an assessment of challenges and risks. For example, data segregation on some particular subjects like transportation would involve segregation costs for removing a subset of data borne by the businesses (data custodians) for specific purposes like transportation. Similar provisions are also recommended for crucial healthcare data as well**.**

The Report also provisions for data sharing by all major cell tower companies for specific purposes. Data on the frequency of data transfer by cell towers can be requested for specific purposes. Here, it is also important to questions what are the underlying costs involve for benefits creations? What are the privacy risks in sharing such data? And, how can accountability be ensured for the same?

- The Report **proposes** to create a trust framework for data-sharing without thoroughly assessing the Indian experience with intermediaries and trust framework. **Thus, we need to understand - if we can utilise the existing legal regime of 'trusts' in the Indian data economy?** [56]
- Research has pointed out that, in practice, organizations need to continuously realign work practices, organizational models, and stakeholder interests to reap the benefits from big data. In this context, it was pointed out that two socio-

---

54 Iryna Susha et al., "Identifying Mechanisms for Achieving Voluntary Data Sharing in Cross-Sector Partnerships for Public Good*," 2019, 227–36, doi:10.1145/3325112.3325265

55 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, doi:10.4018/IJEGR.2018070101.

56 India has trend of mistrust in intermediaries which hampers trust relationships, thus it cannot be assumes that data trusts or trustees will lead in beneficial access to data.

technical features of big data that influence value realisation are portability and interconnectivity. [57]

In light of these findings, we need to assess - whether we have adequate and secure data infrastructures addressing interoperability, usability, and readiness of data to realise the envisaged benefits of sharing?

- A study assessing the impact of data sharing for governance purposes highlighted that the data principals know that their data may be shared and may affect change their data sharing behaviour and withhold data.[58]

However, currently, we lack an understanding of consumer behaviour regarding data sharing in the Indian context. Thus, the government may need to assess - how do data principals currently share data? Will they still be willing to share data if they know that data may be shared with government entities in some form? And, even if data sharing remedies such as portability exist, can they leverage its full benefits?

2. Assessment of Targeted Market and Regulatory Failures

Details from the arguments as presented in the Report

- The Report also builds from the underlying assumption that the economic benefits of data are not flowing to the community because there is currently no regulation that ensures that the benefits of data go beyond organisations that collect and process data. It also justifies this claim stating that jurisdictions worldwide are undertaking this exercise to unlock the potential value of data.[59]

  Furthermore, the report states that the goal of the regulation will be to create the community's right in the data as the benefits if the data should accrue not only to the organisation but also to communities that are the source of the data.[60]
- The Report does not explicitly provide any evidence of regulatory gaps and market failure; however, it picks up from the previous Report in this regard, building on the assumption that Big Tech companies have large data pools giving them an unbeatable techno-legal advantage and there is the untapped value of data. And, this provides them with a first-mover advantage and network effect, and data collection which hampers community welfare.[61]

---

57 Wendy Arianne Günther et al., "Debating Big Data: A Literature Review on Realizing Value from Big Data," *The Journal of Strategic Information Systems* 26, no. 3 (September 1, 2017): 191–209, doi:10.1016/j.jsis.2017.07.003,

58 Niva Elkin-Koren and Michal S Gal, "The Chilling Effect of Governance-by-Data on Data Markets," *The University of Chicago Law Review*, 2019, 29

59 Page 5 of the Report

60 Page 6 of the Report

61 Page 41, Appendix 2 of the Report

However, it does not indicate how the market forces have failed or will fail in the foreseeable in proving consumer welfare in data-driven services. While there exist challenges such as misinformation, censorship, surveillance, and data leaks, however, these challenges may be resolved through data protection rather than data-sharing.

- **Here, while the Report assumes that without regulation, the benefits may not flow to the community without any evidence. Additionally, it does not consider the different stages of development of data economy, and stages of development, timelines and policy consultations that other jurisdictions might have undertaken before prescribing any data sharing regulation**.

  The report also states that it wants to create regulatory certainty for different actors in the NPD[62] ecosystem. However, here again, the Report assumes that the new regulation will create more certainty. There currently exist gaps in this regard. **Thus, it assumes here that such uncertainty exists in the market and start-ups in India and entrepreneurs have been left behind in the digital economy and this can be addressed through data sharing and providing access for public interest would stimulate innovation**

- Furthermore, the report also states that this regulation will help address harms related to privacy and re-identification of anonymised personal data and prevent misuse of and harms.[63] The Report here assumes that it will enforce this framework parallel to the data protection framework, which may be enacted soon.

Concerns

Overall, the Report misses on stipulating what it is the gap that it is trying to achieve through regulating data sharing for public interest i.e. '**what is the benefit of data sharing'; 'whose benefit the report is trying to achieve'; and 'how is the proposed framework optimally suited for India's data economy'.** Thus, concerns pertaining to the assumption on market failures and there is a need to assesses the following issues -

- As stated before, the need for regulation in the public interest must emanate from existing market failure and establish the government capacities for a need to regulate. Still, any evidence of the assumption of market failure has not been presented. Here, it is important to consider that competitive advantage such as data accumulation does not itself lead to abuse, and the practices that lead to abuse of dominance must be established.

  Thus, before an *ex-ante* approach of public interest regulation as proposed by the Report, there is a need to investigate if there is any evidence of market failure and

62 Page 6 of the Report
63 Page 6 of the Report

imbalance in bargaining power. **The markets are working against community welfare and producing externalities? And, if they do exists, whether such imbalances are directly related to the accumulation of data?**

- The Report prescribed data sharing to enforce the rights of the community in NPD and unlock economic benefits of innovation, however, it needs to be assessed regulation is an optimal tool to correct these imbalances**?**
- The Report recognises a need to improve public sector data sharing, however, it emphasises that most of the untapped value in data is with the private sector. **However, we need first to assess public sector data sharing to examine if there is any evidence if the government can tap into data value. As such, the focus should be now on private-sector data sharing?**
- **The Report considers data-sharing as a silver bullet ensuring public interest through data sharing and providing regulatory certainty. Therefore, we need to examine if there is any evidence that the process of sharing data will in fact, be competitive and not create unintended consequences and uncertainty? What are the mechanisms to achieve this?**
- In research assessing the governance of data by public authorities, the assumptions that incentive to collect data always exist and that firms will continue to collect data not-withstanding governmental access to such data do not appear to be on strong footing. In some instances, an inverse relationship exists between incentives for collecting data and sharing it for governance.[64]
- Moreover, the incentives of data subjects to allow the collection of data by private entities might also change, thereby potentially affecting the efficiency of data-driven markets and, subsequently, data-driven innovation.[65] **Thus, we need to assess whether the government, regulator or a data trustee or any other authority is equipped to balance the data economy? Are there any used cases in this regard?**
- **Are smaller companies within the data economy equipped to leverage the opportunities presented on data sharing? Is there a first need to improve their efficiencies in terms of infrastructure, contractual practices, etc.**?
- It has been argued that digital markets are dynamic and consumers can switch to different platforms/ service providers with ease (multi-homing). Any ex-ante regulation to 'curb' network effects without an assessment of harms to consumers is ill-founded. In any case, competition law and economics have the right tools to conduct this analysis and step in and where it is required.

---

64 Niva Elkin-Koren and Michal S Gal, "The Chilling Effect of Governance-by-Data on Data Markets," *The University of Chicago Law Review*, 2019, 29,

65 Niva Elkin-Koren and Michal S Gal, "The Chilling Effect of Governance-by-Data on Data Markets," *The University of Chicago Law Review*, 2019, 29, https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=6126&context=uclrev.

- The legitimate state aim of reducing entry barriers for start-ups and reducing the networks of exploitation of data by large companies can be adequately addressed through competition law frameworks.[66] Therefore, we should consider, whether **existing laws related to competition and intellectual property are not adequate to address the concerns represented by the Report and it would be wise to create an excessive regulatory regime?**
- A study examining the European Commission principles on private data-sharing observed while holding consultations amongst industry stakeholders that most stakeholders were against any regulatory intervention for data sharing. They stated that the effects were seen in the market and the creation of data siloes are normal market behaviours in the data economy and do not indicate a market failure.[67] **Thus, it is pertinent to examine the concerns highlighted in the Report emanating from normal market dynamics rather than market failure?**
- Is there any evidence of uncertainty emanating for startups- MSMEs at this stage or in the immediate future - or is it too soon to think about uncertainty for MSMEs and benefit more from sharing and accessibility of public data and overcoming it other barriers to innovation?
- In a study examining the pre-requisites of digital information sharing in industry 4.0, it was pointed out that social capital, which constitutes a social network, trust, and shared vision amongst the actors of industry 4.0, is important to information sharing.[68]

  We should also examine this within the Indian context: whether we have existing social capital in the form of a transactional relationship, social interaction, and trust models based on which businesses will share the data to create more certainty for them?

3. Assumption on Policy Maturity

3.1 'Data Trustees' framework for processing data sharing requests

Details from the Report

- To govern the request for data sharing, the Report proposes establishing a new kind of intermediaries in the form of data trustees. Data trustees are responsible for the identification of HVD for public interest and processing requests for data

66 CUTS Comments on Non-Personal Data Framework, https://cuts-ccier.org/non-personal-data-what-the-govt-proposes-why-it-needs-reworking

67 Begoña Gonzalez Otero, Evaluating the EC Private Data Sharing Principles: Setting a Mantra for Artificial Intelligence Nirvana?, 10 (2019) JIPITEC 66 para 1

68 Julian M. Müller, Johannes W. Veile, and Kai-Ingo Voigt, "Prerequisites and Incentives for Digital Information Sharing in Industry 4.0 – An International Comparison across Data Types," *Computers & Industrial Engineering* 148 (October 1, 2020): 106733, https://doi.org/10.1016/j.cie.2020.106733.

sharing. Here Report assumes that the data trustee framework will work efficiently in Indian and will lead to the creation of trusts.

- **Moreover, it also assumes the sustainability and capacity of data trustees to process the data requests so that it leads to the creation of a 'public good' that would be used for 'public interest'.**

Concerns

- While the committee states that data trustees will be organically created through the coming together of community members, under Section 8, it negates the traditional challenges with the trusteeship framework. It has been observed that the Indian digital economy space is relying on setting up intermediaries to liaison between the end-users and the service providers —for example, iSPIRIT[69] and setting up of account aggregators or the proposal to establish consent managers. The proposal of these mechanisms seems novel, however, we should be mindful of the 'intermediary problem' in India.

  There is mistrust in intermediaries due to their tendency to tilting towards commercially beneficial interest, leaving the actual intended beneficiaries behind.

- These observations of the 'intermediary problem' were reflected in the Report of Vijay Kelkar regarding public-private partnerships (PPP). The Report stated a need for improvement in the existing PPP models. They should not be seen as a tool for evading its responsibilities and resulting in mismanagement, but to ensure service delivery to the citizen. Overall, the Report recommended that there should be better risk allocations with benefits to the citizen at the core.[70]
- Furthermore, data trustees may create competitive disadvantages for data custodians if they also have similar mission statements and objectives. For example, the National Payments Corporation of India, which is a Section 8 company and may be eligible to be a data trustee, at the same time, may also hold competitive interest in case of processing requests related to the fintech sector, which may not serve the community's interest as assumed by the report. This could create further conflicts and a new form of data monopolies. Thus, there is a need for examination **– how the creation of new intermediaries affects current market operations?**
- **Moreover, in cases where government entities could be appointed as data trustees and could also act as data requestors creates a risk of bias. They may**

69 https://ispirt.in/

70 PRS Legislative Research. Report Summary: Report of the Committee on revisiting and revitalising the PPP model of infrastructure

use data for governance purposes without due consideration to community interest and privacy concerns.

- These concerns have also been reflected in a research conducted on evaluating data trusts as a model for data governance, and it highlighted that it is essential to establish the fiduciary responsibilities of the data trustee; however, at the same time, we should be mindful that ensuring this may itself be costly. Furthermore, it also observed that grievance redress in the forms of additional safeguards, certification mechanisms, unbiased representation, interoperability of and amongst data trusts is also important, and their funding is also essential pre-requisites.[71]

  Thus, it is important to assess – how can these be ensured in the Indian context? What will be the role of the government in assuring these factors? What can be the potential funding models for data trustees to carry out their responsibilities?

3.2 Setting up of Non-Personal Data Authority (NPDA) and legal and economic analysis

Details from the arguments as presented in the Report

- The Report, in its legal and economic analysis, also states that data could be equated with 'material resources' and that regulations can be made for equitable distribution. Here, the Report assumes that data could be governed at par with other material resources such as forests, pastures, or land to formulated common property rights.[72]

  Along with this, while the Report recognises that intellectual property rights would exist in NPD,[73] however, in its legal analysis it concludes the no IPR would rest in HVD. No precedents in the Indian jurisprudence have indicated such an analysis. **Moreover, the Report assumes that companies that HVDs will not satisfy the threshold of creativity and skills that will validate attaching intellectual properties.**
- The Report proposes creating a new NPDA, which will have an enabling function ensuring unlocking of economic benefit for society, creating a data-sharing framework, and maintaining meta-data directories. Along with this, it will also

71 Anouk Ruhaak, "Data Trusts in Germany and under the GDPR" (Germany: Algorithm Watch, 2020), https://algorithmwatch.org/wp-content/uploads/2020/12/Data-trusts-in-Germany-and-under-the-GDPR-Anouk-Ruhaak-AlgorithmWatch-2020.pdf#page=1&zoom=110,-357,719.

72 Page 58 of the Report

73 Page 32 of the Report

have an enforcing function of addressing privacy concerns and misuse of data and adjudicating on sharing of HVDs.[74]

The Report also assumes that existing sector-level authorities or competition authorities may not be adequate to address the concerns within the data economy such that economic benefits can be unlocked for public and economic interest. Additionally, with the upcoming PDPB 2019 and e-commerce policy, more regulations are proposed to be created covering the aspect of privacy protection and competition.

Concerns

- In an analysis conducted on properties of data to be a common public good, it was observed that data has some unique properties such as – multiplicity through which particular dataset can be at multiple places at any given times and can be accessible from multiple points, thus it can be very difficult to establish meaningful collective control over data. Since data cannot be created in a vacuum and needs to be collected with specific knowledge of subjective information that needs to be obtained for a specific purpose, it requires a systemic application of a query-based response to cultivate data, making data creation a creative process, rather than a freely available public good.

  Moreover, it can be challenging to assess the relevance of data in the local context as it may represent varied diversity. And, foremost, the unit of governance such as the community cannot be defined appropriately due to the lucid nature of data and lack of boundaries. **Overall these differentiate data from other physical common pool resources.**[75]

  Thus, we need to question - whether the understanding of the equitable distribution of material resources for the common good under Article 39 of the constitution can be applied to data? If not, what are the points of departure and what are the risks associated with such interpretation?

- Furthermore, the 'case for regulating data', has an underlying tone of inequitable distribution. It stipulates for the appropriation of data for economic benefits such as starting new businesses and developing new value-added services with an underlying objective of equitable distribution due to the accumulation of data by Big Tech companies. It seems that the Report has indirectly tried to address, under the broadly defined 'public interest' purposes, the problems of inequitable

[74] Page 20 of the Report

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

distribution, market failure and citizen welfare, all of which could be covered under competition law.[76]

- Experts have argued that the NPDA may not have any new role to play in prescribing equitable distribution. These could be addressed by the Competition Commission of India, considering it may also essential facilities doctrine[77]. However, this capacity of regulators will have to be enhanced.

  Therefore, there is a need for a deeper examination of how the existing authorities may not facilitate data sharing for community benefit? What may be the differential aspect of the proposed authority from the existing regulators?

- It is also being pointed out that creating new authority would create overlaps and more regulatory costs for the companies. **Hence, considering the evolving regulatory space in the digital economy, will the creation of such an authority lead to the risk of over-regulation and over-laps, further stifling innovation and investment in the sector?** [78]

- In an analysis examining copyrights in metadata, it was observed that there is no copyright in metadata on its face as it may not satisfy the creativity threshold.[79]At the same time, it also acknowledged that there remains a possibility that certain compilations may involve creativity in their selection and coordination. Additionally, it may also be important to examine the commercial sensitivity of meta-data, as they may indicate important information about the relation between databases or important filters.[80]

  The CoE presumes that data that is extracted from 'pre-set fields' would not violate the copyright protection of the database, given that there is no originality involved in arranging data according to such 'pre-set fields'. This could be challenged, as the unauthorised/forced extraction of relational fields of data in aggregated form, even if done as per pre-set requirements, would still violate a data business' copyright in the said datasets because even the segregated data sets would potentially comprise a copyrightable database compilation.

  Similarly, the CoE assertion that aggregated datasets are not subject to IP protection could also be questioned as most aggregated datasets are prepared through the application of skill and creativity. Thus they may enjoy copyright protection under Indian law. To determine whether a dataset is 'original', a case-

76 https://www.cci.gov.in/sites/default/files/advocacy_booklet_document/CCI%20Basic%20Introduction_0.pdf

77 Ashok Kumar Gupta. http://www.cci.gov.in/sites/default/files/speeches/CUTS.pdf?download=1

78 Deepro Guha, "Regulatory Cobwebs Threaten to Scupper India's Data Economy: A Closer Look at Threats, Opportunities- Technology News, Firstpost," *Tech2*, 13:39:28 +05:30, https://www.firstpost.com/tech/tech/regulatory-cobwebs-threaten-to-scupper-indias-data-economy-8678321.html.

79 Krista L Cox, "Metadata and Copyright: Should Institutions License Their Data about Scholarship?," 2017, 9.

80 Kurt Cagle, "The Value of Metadata," *Forbes*, accessed December 11, 2020, https://www.forbes.com/sites/cognitiveworld/2019/02/26/the-value-of-metadata/.

by-case determination would be needed. **Thus, foregoing the copyrights and trademarks in metadata would require further examination and analysis generally and on a case-to-case basis.**

- Furthermore, although Indian law does not protect trade secrets, India is a signatory to the TRIPS agreement and is obligated to protect the secrecy of commercially valuable.[81]

  Thus, while the Report recognises that there may be trade secret protection for certain raw NPD, it further states that this may not be applicable because of eminent domain. **Here again, there is further need for analysis – what are the possibilities of regulatory overlaps and risks?**

- On the lines of argument stated above, it has also been pointed out that data provisions should be complemented with coherent incentive mechanisms and sustainable business models while acknowledging the limitations of (data) markets. This may require addressing uncertainties about data ownership and clarification of the role of privacy, intellectual property rights (IPRs), and other ownership-like rights, which ideally should be undertaken by appropriate expert agencies and organisations.[82]

  Thus, we need to examine the existing legal regimes pertaining to data ownership, intellectual property rights within existing business models in India?

3.3 Interface with Personal Data Protection Bill 2019

Details from the arguments as presented in the Report

- The Report has taken a step in the right direction of recognising and overlaps between the PDP Bill and the framework proposed by this Report. The Report suggests the removal of clauses 91(2) and 93(x) from the PDP Bill for addressing this.[83] However, this suggestion is based on the assumption that Joint Parliamentary Committee will consider this.
- The Report has also defined NPD as any data that contains personally identifiable information (PII). Here again, the Report seems to have borrowed from the understanding of the PDP Bill, however, the meaning of PII is still not clear.

---

81 https://www.remfry.com/wp-content/uploads/2017/11/the-viability-of-trade-secret-protection-managing-the-ip-lifecycle-sept-2013.pdf

82 "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 Library" (OECD, 2019), https://doi.org/10.1787/276aaca8-en.

83 Page 10 of the Report

- For making the data principals more empowered, the Report proposes a consent mechanism in the form of an 'opt-out' option for anonymisation. Here, the Report assumes that consumers are capacitated to make choices regarding the treatment of their data and assess its purpose of use.
- The Report states that as soon as the data will anonymise it will fall in the category of NPA and as such clear distinction could be made between personal data and NPD. Here again, the Report assumes that techniques of anonymisation will be full-proof and binaries could be created between personal and non-personal data through such techniques.

Concerns

- There remains uncertainty regarding the suggestion of removing clauses from the PDP Bill, especially as there are parallel yet conflicting narratives evolving around the PDP Bill and NPD. In an earlier press reporting on the PDP Bill, it was indicated that Joint Parliamentary Committee (JPC) is contemplating also include NPD within the ambit of data protection, contrary to the current Report. This required for examining – whether the first step should be to have a PDP Bill?
- In India, a consumer privacy perception survey conducted by CUTS International found that users do not read privacy policies (notices) due to their length, legalese, complicated and unfamiliar language.[84] This indicated that there are notice and consent fatigue amongst consumers. In light of this, it should be considered – how can effective consent be obtained from consumers without awareness and understanding about anonymisation?
- Studies[85] and a detailed analysis conducted by the Article 29 Working Party while establishing standards for GDPR[86] have indicated that the level of anonymisation differs with different techniques and tools; thus the susceptibility of re-identification also changes. On the other hand, it is also observed that the over-anonymisation of a dataset can render it useless for doing further analysis and innovation.[87]
- The European experience with the General Data Protection Regulation (GDPR) application has indicated that the definition of personal data itself is very context-specific. In this regard, the EU Court of Justice in the case of *Peter Novak*

84 https://cuts-ccier.org/cdpp/

85 https://www.theguardian.com/technology/2019/jul/23/anonymised-data-never-be-anonymous-enough-study-finds, https://theprint.in/opinion/india-has-to-toe-a-fine-line-in-defining-non-personal-data-between-public-interest-and-ipr/382149/

86 https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

87 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 (February 1, 2020): 11–36, doi:10.1093/idpl/ipz026.

v. Data Protection Commissioner[88] upheld the concept of 'relatability' in defining personal data, which broadens the scope of definition, making it more context-specific.

Similar observations were also made in the privacy perception survey conducted by CUTS, involving in-depth interactions with around 2400 users, highlighting the need to include user perception and perceived sense of users' intimacy and necessity of data in the test of 'identifiability'.[89]

More recently, researchers have warned that considering the technological and jurisprudential evolution of the definition of 'personal data', a lot more reviously considered NPD will come within the category of 'personal data'.[90]

Thus it is important to examine – whether binaries between personal and NPD can be made? What are the associated privacy risks for communities and data principals in doing so?

- The Economic Survey of India 2019-20 pointed out that data should be used for the public good and for that, public sector data should be opened. On similar lines, the Report also proposes strengthening the policies and infrastructures for public sector data sharing. However, the push in the Economic Survey was towards public sector data sharing. **In light of this, should the strategy for open data sharing be the first step for NPD sharing framework to follow?**[91]
- The report borrows many of its concepts from the personal data protection bill of 2019**. How can we reach requisite policy understanding and maturity before the implementation of PDPB 2019**?.
- More generally, it may also be important to learn from the policy developments that are happening at the sector level in India, specifically the fintech[92] and health sector,[93] wherein initial development for data sharing is taking place.
- **Thereafter, assessing if we need sectoral level frameworks before prescribing a cross-sectoral level framework. And, should the implementation and the progress of existing policies such as National Data and Accessibility Policy, National AI strategy, Strategy for National Open Digital Ecosystems (NODE) should be assessed before the implementation of an umbrella framework for data sharing in India.**

---

88 ECLI:EU:C:2017:994, para. 35.

89 Objective: Engage with consumers on a pan India level regarding data and privacy protection on both, online, as well as offline platforms, from the government and private players alike. Expected Outcome: Policy reforms empowering consumers for data privacy and protection. https://cuts-ccier.org/cdpp/ and https://cuts-ccier.org/pdf/survey_analysis-dataprivacy.pdf

90 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 (January 2, 2018): 40–81, https://doi.org/10.1080/17579961.2018.1452176.

91 https://www.indiabudget.gov.in/budget2019-20/economicsurvey/doc/vol1chapter/echap04_vol1.pdf

92 Data Empowerment and Protection Architecture, 2020, https://www.indiastack.org/depa/

93 National Health Data Management Policy , 2020 , https://ndhm.gov.in/stakeholder_consultations/ndhm_policies

- We also need to assess - **do public, private and community non-personal data have stakeholders operating at the same level of policy maturity? And, is there a requirement of interoperability models or standards required for the horizontal applicability of data sharing principles? Do these models exist in India?**

4. Assumption on Market Mapping and Maturity

4.1 General Observation

Details from the Report

- The Report states that the world is 'awash with data' and provides estimates for data generated this year and predicted data to be generated by 2025 across the globe. Due to the increase of AI mechanisms, mobile phones and cloud-driven apps lead to an increase in the generation of data.[94] **In this context, the report assumes that the quantity of data is a primary driver of growth.**

Concerns

No assessment has been conducted about the amount of aggregated NPD collected in India and the valuation of such data. This raises concerns such as –

- What is the value of the data economy in India? What has been the contribution of data empowered industry on the GDP?
- How many data companies currently exist in the economy? What is the existing level of innovation in the market?
- What are the existing business models for utilising data and social capital generation[95] in the data economy in India? How is data being valued in existing business transactions? How will these transactions be affected by the framework proposed by the Report?
- Are there particular sectors that are seeing more growth in data drives businesses than others?[96] Is it beneficial to first have a sector-specific focus on data sharing considering the sectors that see more growth?
- Highlighting the importance of requisite data infrastructure, studies have pointed out that data analysis can generate promised value only when adequate hardware systems are in place. They have also indicated some important shortcomings and challenges of current technology trends.

---

94 Page 5 of the Report

95 Julian M. Müller, Johannes W. Veile, and Kai-Ingo Voigt, "Prerequisites and Incentives for Digital Information Sharing in Industry 4.0 – An International Comparison across Data Types," *Computers & Industrial Engineering* 148 (October 1, 2020): 106733, https://doi.org/10.1016/j.cie.2020.106733.

96 Use cases in India

These include a lack of intelligent Big Data sources, scalable real-time analysis capability; lack of support (in networks) for latency-bound applications; the need for necessary augmentation (in network support) for peer-to-peer networks; and rethinking on network support the cost-effective, high-performance storage subsystem.[97]

In light of this, we should examin**e the current technology trends in India with respect to the development of technology hardware (such as scalable real-time analysis capability, network support, cost-effective storage systems)?**

- Will the mechanism as proposed by the framework incentivise start-ups to participate in data sharing? Will they themselves make a data contribution towards this sharing mechanism without established incentives for data collection?

4.2 Mandatory Data Sharing Approach

Details from the arguments as presented in the report

- The report prescribes mandatory data sharing for the public interest, with the only caveat of adjudication by the Non-Personal Data Authority (NPDA). **Here again, the report assumes that imposing mandatory data sharing on data custodians will help in boosting the economic interests and lead to public benefits.**[98] **Furthermore, it disregards the risks of disincentivising in collecting data and investing in research and innovation.**[99]

Concerns

- Experts have highlighted the adverse impact of mandatory data sharing, stating if the market is not mature, it stifles competition and investment instead of promoting it.[100] **In light of this, we need to examine, has the Indian data economy (specifically concerning NPD) has evolved to adopt mandated sharing of data without any adverse consequences?**
- An OECD assessment conducted on the risks of data sharing pointed out that while regulation may impose data access, it may also undermine incentives to invest in data in the first place, particularly when data commercialisation and licensing are not viable options. For instance, for organisations and individuals,

97 Wasim Ahmad Bhat and S.M.K. Quadri, "Big Data Promises Value: Is Hardware Technology Taken Onboard?," ed. Professor Leroy White and Professor Xu Chen Dr Xiaojun Wang, *Industrial Management & Data Systems* 115, no. 9 (January 1, 2015): 1577–95, doi:10.1108/IMDS-04-2015-0160.

98 Page 21 of the Report

99 Mark MacCarthy, "Data Sharing: A Problematic Idea in Search of a Problem to Solve," *CIO*, August 30, 2018, https://www.cio.com/article/3301175/data-sharing-a-problematic-idea-in-search-of-a-problem-to-solve.html.

100 https://cuts-ccier.org/pdf/report-webinar-optimal-governance-of-non-personal-data-august11-2020.pdf

including researchers, which build their competitive advantage based on data lock-in, mandatory data access and sharing could undermine their ability to compete, to a point where their incentives to invest in data may be too low to enter a particular market. For some start-ups, this could mean that they lose their attractiveness as acquisition targets of larger firms and thus their economic value. [101]

Thus we need to examine - whether mandatory data–sharing leads to more innovation and are there any use cases indicating this? Is there a risk of anti-competitive effects of mandatory data-sharing? Will this also increase the risk of stifling innovation?

- Similarly, in the Copenhagen Expert Workshop conducted on issues of data sharing and re-use, it was indicated that mandatory data sharing would bring more unintended consequences for the smaller startups as sharing their data would lead for the bigger players to enter into niche markets, wherein startups tend to develop their competitive edge.[102]
- **Mandatory data sharing and loose criteria of 'public interest purpose' may create a misalignment of incentives and externalities through the chain of data sharing. This gap needs to be assessed in light of incentive misalignments and externalities that may emerge from mandatory sharing of HVD as a 'public good'. Thus, it may need to be further analysed the impact of mandatory data sharing on the data custodians?**
- A study conducted in the EU on the economic impacts of business on government data-sharing noted that data businesses might themselves strategize to expand business or provide chargeable public benefits, or the data business may be in the business of providing data to the government for carrying out its functions. In such cases, neglect of 'proportionality' between private costs (current and foreseeable) and public benefits could be detrimental.[103]

  Thus, there is a need to further examine how an appropriate balance can be struck between intended public benefits and costs of sharing?
- **More generally, maintaining other public goods such** as water or roads lies with the government; however, the current framework differs in this regard. Assuming that data is a public good, the transaction cost of creation, identification and fulfilment of the data sharing requests have been transferred to the data trustee (government or non-governmental organisation). However,

101 "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.

102 "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.

103 https://ec.europa.eu/jrc/sites/jrcsh/files/jrc119947.pdf

in a situation where proper and sustainable incentive structures are not specified for data trustees, it may prioritise private interests over public or community interests.[104]

- **Can mandatory data sharing lead to enhancement for public interest purposes? Are there any uses cases in India for this? What are the risks of possible misalignments and risks? Are there alternatives to achieving such purposes?**

5. Process Followed

Details from the arguments as presented in the report

- The Committee of Experts for formulating Non-Personal Data was constituted in 2019 under the chairmanship of Kris Gopalakrishnan. This was done after the Srikrishna Committee presented in its Report the Personal Data Protection Bill 2018 and also recommended that the government may come with a framework to protect community data where the community is identifiable.[105]

  The Kris Gopalakrishnan Committee constituted representatives from civil society, industry representatives, and experts.
- Committee met with representatives from various business-like health, e-commerce, not-for-profit/think tanks, technology service providers, and experts to formulate its recommendations. A literature review of the current status of such data-sharing frameworks across jurisdictions was also undertaken.

Concerns

- While the committee did undertake consultations, **the list of experts with whom it consulted was not made public. Furthermore, the evidence or the findings of the consultation is not disclosed.**
- The committee also **did not undertake an impact assessment of existing data sharing strategies and mechanisms, initiatives that exist in India**.
- The committee also did not undertake Regulatory Impact Assessment (RIA) and cost-benefit analysis.

104 https://nic.org.uk/app/uploads//Data-As-Infrastructure.pdf

105 https://www.meity.gov.in/writereaddata/files/constitution_of_committee_of_experts_to_deliberate_on_data_governance_framework.pdf

Conclusion

The BSA indicates that the Report has based its rationale to propose regulation for NPD sharing on assumptions that require further analysis and nuanced scrutiny. While the Report's intentions are noble, it needs to be appropriately embedded within economic, social and political contexts. The nature of the data economy is such that it grows at a fast pace. However, data is different from other resources and as such, implications of its sharing would be different, this presents a problematic situation that requires an evidence-based policymaking process.

In this regard, countries worldwide are trying to experiment and evolve their understanding of data-sharing. While there are no established best practices, it would be beneficial to consider appropriate nuances of the various dimension of data-sharing.

Some such dimensions as synthesised from the secondary literature review are presented below:

- **Purposive construction of value and nature of NPD:** Studies and research indicated that the nature of data is multi-faceted. Its value is lucid. As such, attribution to its nature and value cannot be done in a vacuum. There are multiple factors to consider, such as proprietary rights in data, power dynamics in data usage, quality of data, the cost involved in data management, interoperability challenges.

  Therefore, it is challenging to concretely state that data as an economic resource in the public good for public interest purposes can create the intended value. In this regard, the literature points towards a "purposive construction," i.e., to determine the purpose of sharing, intended beneficiaries and associated risks in a granular manner, to determine the nature of data to be shared and value intended to be achieved from such sharing.

  Thus, rather than making blanket claims and shying away from the problematic and lucid nature of data, data sharing should be approached through a step-by-step process, starting from clearly identifying the 'purpose of sharing'.

- **Clarifying the scope of "public interest purposes":** Research, case laws and policies indicate that no fixed meaning can be assigned to the "public interest purposes". The report has also talked about "public interest purposes" loosely, leaving ample room for ambiguities leading to speculations that anything can be presented in a way so that it falls under the ambit of such purpose.

  The secondary literature points towards unintended consequences such as conflicts between public and private interest, disincentive private players to collect data and the benefits of not reaching the targeted beneficiaries, from a broad interpretation of "public interest purposes".

In this regard, the 'theory of public interest regulation' advised for identifying market gaps; state capacity and infrastructural prerequisites before prescribing a regulation. Case laws and other studies also recommended that alternative approaches such as facilitating voluntary data-sharing should be explored to fulfil public interest purposes.

- **Tracing Market/Regulatory and Technical Gaps:** While the report has discussed the value and the benefits of data sharing, it hasn't presented evidence of the market, regulatory and technical gaps that it aims to address. As pointed out above, the mapping of market failure is important to provide validation for regulation for public interest purposes.

  Moreover, it should not be assumed that incentives to collect and data are given and will continue despite prescribing governmental access to such data. Thus, assessing the current market and regulatory failures is necessary to identify a clear problem statement for the proposed NPD framework.

- **Establishing Fiduciary Responsibility of Data Trustees:** The report has recommended a 'Data Trustee' framework to process the data sharing requests, creating a new form of intermediaries. However, the experience of India has warned us about the 'intermediary problem' that may exist. This may create biases and mistrust, leading the "Data Trustee" to act against the public interest. In this regard, experiences from other countries recommended having clear rules to establish the fiduciary responsibility of data trustees with appropriate safeguards, certification mechanisms, clear mission statements, unbiased representation, interoperability, and sustainability. The focus here should be building trust between the intermediaries, community and consumers.

- **Careful Examination of Existing Legal Regimes:** Literature in this regard pointed out that the nature of data is very different from other material resources. It has some unique characteristics such as - data can exist at multiple places at one point in time and can have diverse representations from multiple communities simultaneously.

  Thus, assigning similar ownership rights or community frameworks in data as other material resources would be sub-optimal. Moreover, it is also essential to assess IPR pertaining to datasets. Since the nature of data is lucid, a new form of creativity may be reflected in compiling datasets or even meta-data. Thus a blanket negation of IPR should be avoided.

  On the other hand, it is important to consider and analyse, the regulatory objective of the NPDA, as it may overlap with the jurisdiction of CCI if it is intended to address the problem of inequitable distribution and of the DPA in the case where there are no clear boundaries between personal and non-personal data. It might be important to take stock

of the regulatory landscape and tools available at this stage and examine their applicability within the data economy.

- **Adequate Policy Sequencing:** Regarding the interface between the NPD framework and PDP Bill, the literature cautions that the binaries between personal and NPD are very difficult to make as no anonymisation techniques are full proof. Thus, an over-reliance on anonymisation is a slippery slope. Moreover, there already exists issues regarding consent fatigue, and burdening consumers with knowing anonymisation in making consent decisions will not be fruitful.

  With these complications and lack of clarity, the experiences traced within the literature suggest that the focus should be first on opening public sector data and developing sector-wise approaches of data sharing targeted towards fintech or health sector, which have a more mature policy environment regarding data management policies. Along with proper data protection laws, cyber-security and appropriate legal mechanisms to recognised intellectual property rights in data.

- **Risks of Mandatory Data Sharing:** Research and experts have cautioned that mandatory data-sharing without assessing market maturity would increase the risk of stifling competition and investments in the data economy. And these consequences would be much graver for smaller start-ups, creating misalignment of incentives.

  In this regard, studies conducted in the EU emphasises that in determining incentives in data sharing, 'proportionality' needs to be struck between private costs (current and foreseeable) and public benefits. These incentives should be available for all actors in the data-sharing process, even the data trustees, to have a sustainable infrastructure and avoid leaning towards private interest due to lack of incentives.

Comparative Jurisdictional Analysis

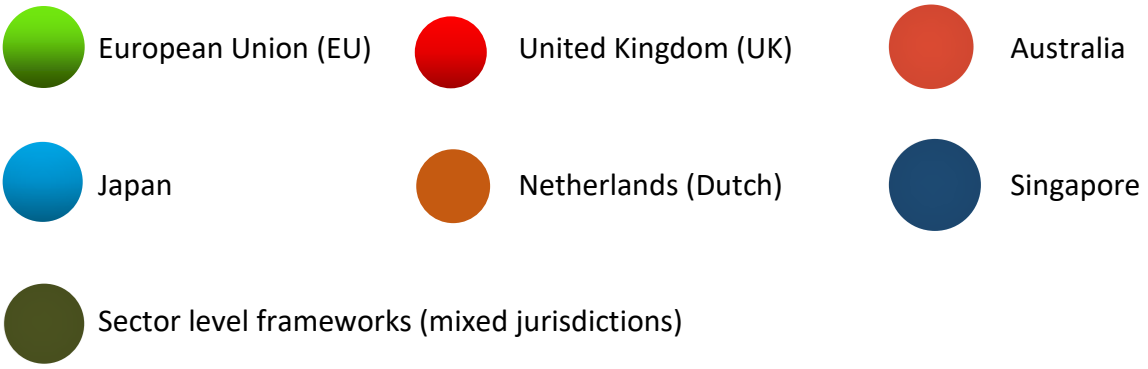
Introduction

Data sharing frameworks are still evolving in various jurisdictions and different approaches and methods of governance are being adopted by countries based on different contexts. As India is embarking on a similar journey, it would be beneficial to take stock of the relevant data-sharing framework to facilitate cross-learnings. Most countries are still in a nascent stage of the development of such frameworks. In light of this, a comparative jurisdictional analysis of such frameworks would shed light on some unanswered questions and concerns presented by the revised report on the Non-Personal Data Governance Framework (the Report).

To this end, the comparative jurisdictional analysis presents an overview of the rationale and objectives of the existing and proposed data-sharing framework/initiatives/strategies across jurisdictions at both sector and umbrella levels. The aim is to put the rationale and assumptions presented by the Committee of Experts (CoE) in the Report in contrast with other jurisdictions to derive comparative assessments.

The complete comparison matrix of nineteen **data sharing frameworks/policies/strategies (frameworks**) is available as Annexure I, while the section below focuses on its analysis. **In the forthcoming section on analysis of the comparison matrix, diagrammatic representation (matrix format) of the insights from the assessment is being depicted to highlight commonalities and differences between jurisdictions and frameworks.**

Please note the colour themes of jurisdictions as represented in the diagram in the analysis section below –



1. Assumptions on the Rationale and Goals

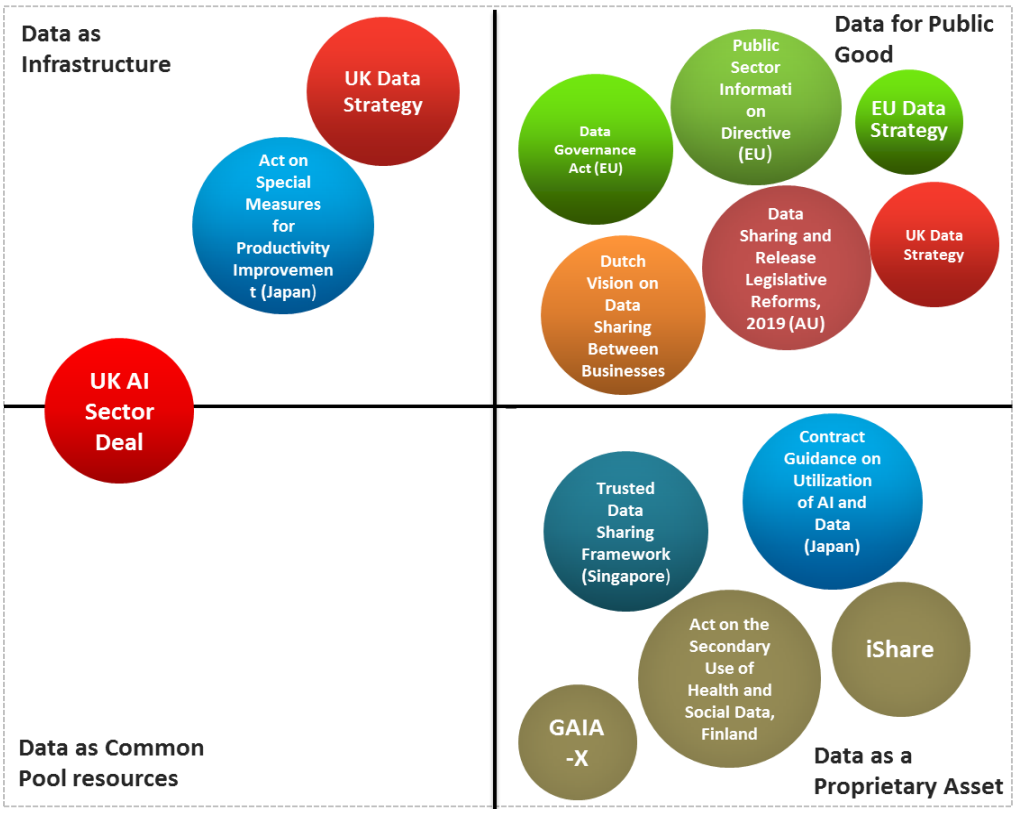
1.1 Nature of Data

Questions for analysis derived from baseline scenario assessment of the Report -

- ***Is data similar to other economic resources? And, in what form is the value derived from data?***
- ***Whether we should consider data as a 'public good' and in what circumstances? And, if so, what kind of data can be classified as a 'public good', and how do we separate data which is 'public' and 'non-public?***
- ***What are the costs involved in transitioning data from a 'club good' to a 'public good'?***
- ***Are there any existing proprietary rights in data? If so, will there be conflicts with existing proprietary claims in data and the envisaged idea of data as a 'public good' for public interest purposes and sovereign purposes?***

Figure Design: The matrix design is used for analysis to club the different data-sharing frameworks based on commonalities so that relational analysis of differences and similarities could be easily depicted and assessed.

Figure 1: Nature of data, authors' Analysis



In **Figure 1**, the buckets of commonalities have been understood as mentioned below:

a) Data as infrastructure: This bucket includes those data-sharing frameworks/policies/ strategies that interpret data as an underlying building block for developing other services. Therefore, these jurisdictions focus on data sharing and exchange to develop infrastructures of data and for data.
b) Data for the public good: This bucket includes those data-sharing frameworks which have internalised value of data for the public good. Notably, different jurisdictions have varied perceptions of 'what is a public good or public interest', depending on their experiences and needs.
c) Data as a proprietary asset: This bucket includes those jurisdictions which have to internalise the proprietary nature of data and facilitate data sharing, recognising the rights of business and enterprises in the data collected and maintained by them.
d) Data as a common pool resource: This bucket includes those jurisdictions in which rights in data are attributed to a group of entities or individuals which may take decisions on sharing of such data.

Analysis (insights derived from Rationale as presented in the comparison matrix)

Jurisdictions have attributed different interpretations to the nature of data depending on the value expectation. One key observation that can be assessed from **Figure 1** is that while some jurisdictions prescribe data to be used for the public good, unlike India, 'public good' attribution to data is not adopted by any other jurisdictions. Terminologies attached to describing the nature of data are also closely linked to the benefits of sharing and the policy maturity in the jurisdictions.

Within the Indian context, there is a basic confusion between data 'as' public good and data 'for' the public good. This is an essential differentiation because if data is considered a 'public good', it will attract economic interpretation, requiring a more nuanced approach to balance incentive and value distribution.

As shown in Figure 2, jurisdictions that consider data for the public good are also jurisdictions that recognise the benefit of data sharing for public interest and facilitate private sector data sharing. It should be noted that while these frameworks aim to promote private sector data sharing, they don't prescribe for mandatory data sharing.

This is specifically important as these jurisdictions tend to focus more on 'public sector data' sharing and are cautious of incentive problems of making private-sector data available for the public interest. Moreover, they acknowledge the differentiation of sharing "public sector data" for the public interest and private sector data to be available for public benefits and prescribe differential treatments accordingly, rather than adopting a "one-size fits all approach".

Figure 2: Benefits of sharing data, authors' analysis

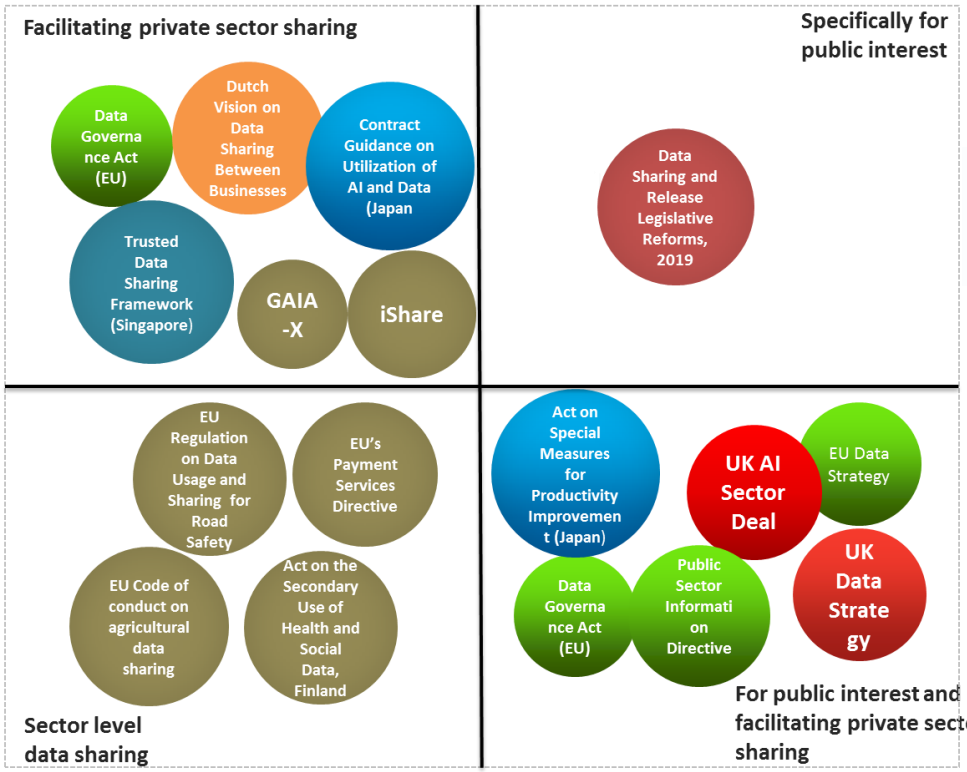
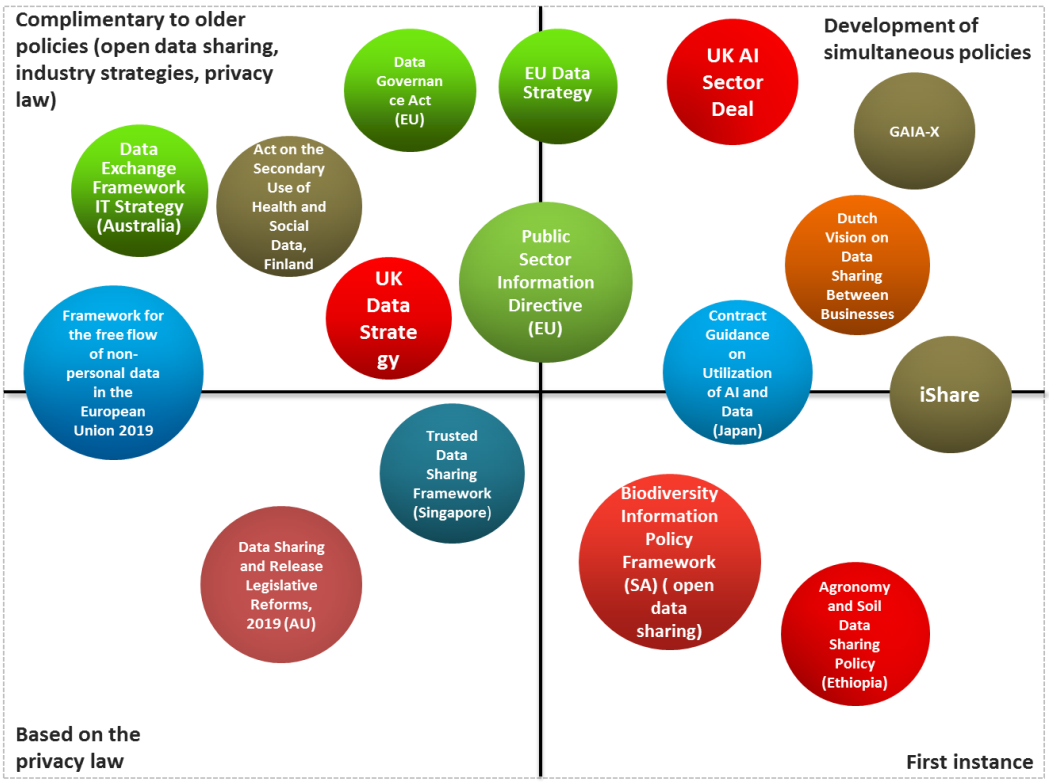


Figure 3: Policy maturity, authors' analysis



Furthermore, as can be observed from **Figure 3,** most of the frameworks which have internalized the value of data for public good have based their strategies and policies complimenting older policies or are in the process of developing simultaneous policies. For instance, the EU Data Strategy, while recognising the value of data for the public good, also recognises that EU-wide common data spaces will be formed based on sector-specific needs, for which simultaneous policies may come into place.

Moreover, the High-Level Expert Group on Private Sector Sharing in the EU has also recognized the FATEN principles[106] of 'business to government' data sharing for public interest to form a practicable and accountable basis for implementing data sharing. At the same time, most of these jurisdictions have functional open data sharing policy, which has also been a guiding factor in the understanding value of data for the public interest.

In the Indian context, while the Report recognises data as a 'public good', it seems to have merged a wide ambit of benefits within this 'public good', without forming differentiation between facilitating openness of private sector data amongst private players and achieving larger public interest. This negation can lead to dilution of associated costs involved in changing the nature of data. Additionally, the complimenting and simultaneous policy frameworks are slow to develop in India, specifically concerning the opening of public sector data.[107]

1.2 Value of Data

Questions for analysis derived from baseline scenario assessment of the Report:

- ***What are the pre-requisites for value creation through data sharing? And, who are the actors involved in such value creation?***
- ***Will the data trustees and communities be capacitated to identify the value of data for themselves?***
- ***Is there substantial evidence available to prove that regulation and access to data would enhance value realisation?***
- ***Will the presumed value creation incentivise a business to keep collecting and sharing data without considering the cost of data collection?***

In **Figure 4**, the buckets of commonalities have been understood as below:

a) Recognition of pre-requisites: This bucket represents those frameworks, which have recognised pre-requisites in terms of technical infrastructures, financial capacities and policy, which are necessary to realise adequate value from data.

106 Fairness , Accountability and Autonomy, Trust and Transparency, Equity and Efficiency, Non-Maleficence

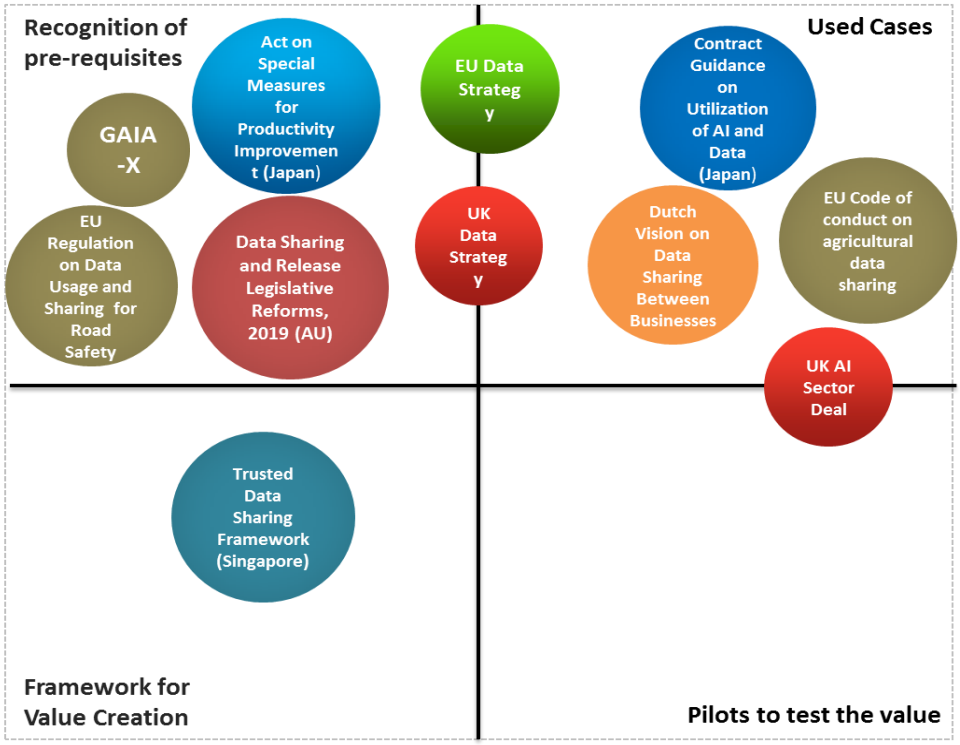
107 Simonti Chakraborty, "Open Data Policy of the Government of India: What Has It Achieved?" (Centre for Budget and Governance Accountability, August 22, 2018) , https://www.cbgaindia.org/blog/open-data-policy-government-india-achieved/.

b) Use Cases: This bucket consists of those frameworks which have relied on or identified used cases of data sharing in assessing expected value creation from data.

c) Framework for value creation: This bucket consists of frameworks that have specifically prescribed mechanisms through which value derived from data could be assessed.

d) Pilots: This bucket represents those frameworks, which are undertaking pilots to assess the value that could be derived from data.

Figure 4: Value of data, author's analysis



Analysis (insights derived from Rationale as presented in the comparison matrix)

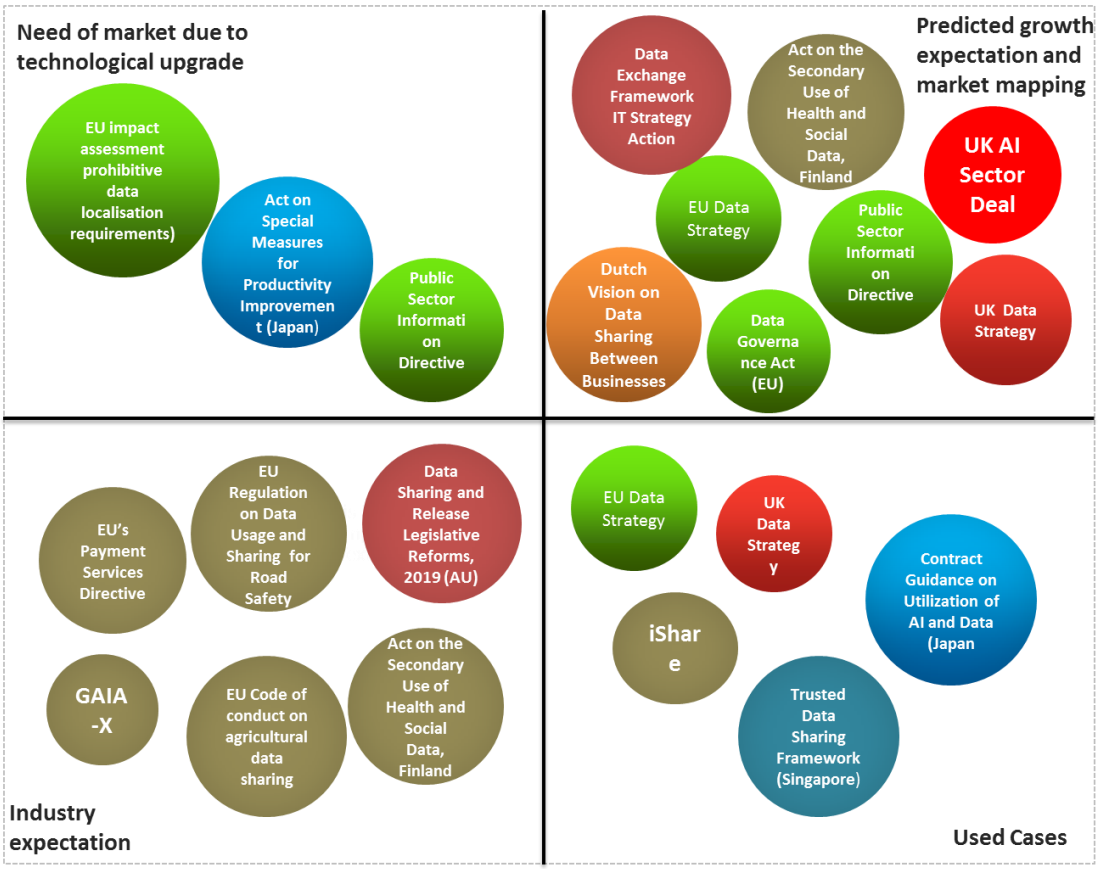
As shown in Figure 4, most jurisdictions have identified pre-requisites through which value could be generated from data. In the EU context, many of these requirements and expectations of value creation stem from regular impact assessments and used cases of data sharing. Singapore is the only jurisdiction that has also created a framework to assist the private sector in assessing its data value. This framework suggests taking stock of the data, assessing rights in data, and determining the benefits of sharing the data.

This recognition of pre-requisites of value creation is closely related to the policy maturity is derived from experiences with open data sharing policies or market

mapping, which depict used cases of data sharing to assess the challenges and associated risks in creating value out of data sharing.

This could also be observed from **Figure 5**, which represents parameters of market assessment, wherein the frameworks that lie in the 'used case' bucket are also those that lie in the bucket of pre-requisites of value creation. This indicates a close link between assessing the market's need and envisioning value created from sharing data.

Figure 5: Parameters of market mapping/ need assumptions, authors' analysis



In the Indian context, while the Report recognises the technological architecture required and other policy specifications which may develop over time, however, a detailed and evidence-based impact assessment or use cases of data sharing, where stakeholder concerns are taken into account would further help in solidifying the grounds of value creation as well as assessing whether such value creation would match the rationale and objective of the Report.

1.3 Benefits of Data Sharing

Questions for analysis derived from baseline scenario assessment of the Report:

- ***How are the benefits of data sharing addressing market and regulatory gaps in the data economy in India? Is the government capacitated to execute such regulation in favour of public interest?***
- ***What are the costs, benefits and prerequisites essential to achieve the envisaged public interest in data sharing?***
- ***What can be the tensions between data sharing for public interest and individual interests in data?***
- ***What is indicated by the used cases of data sharing for the public interest? What can be learned from those?***
- ***Whether we have adequate data infrastructures addressing interoperability, usability, and readiness of data to realise the envisaged benefits of sharing?***

In ***Figure 2***, the buckets of commonalities have been understood as below:

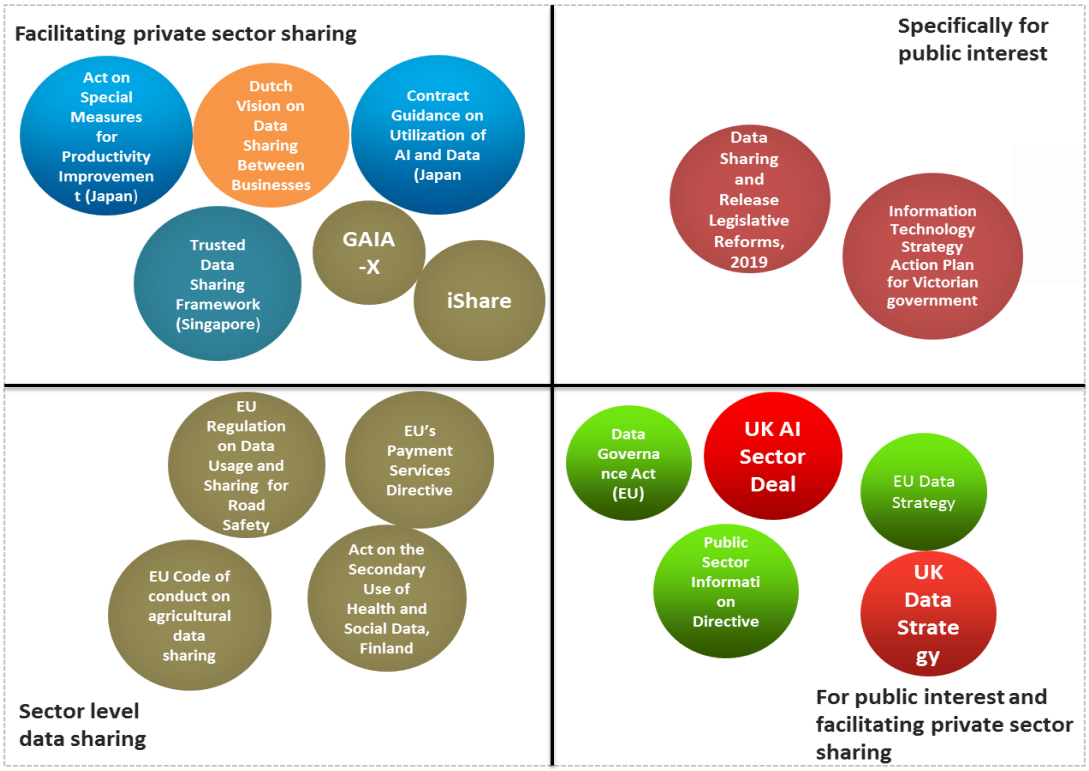
a) Facilitating private sector sharing: This bucket represents those frameworks that have a primary objective of enhancing or facilitating data sharing amongst private sector players. The purpose of such facilitation may differ depending on the identified regulatory or market failures.

b) Specifically, for the public interest: This bucket includes those frameworks which are specifically focused on utilising data just for public interest purposes. The jurisdictions that lie within this bucket have a specific focus on delivering public services, improving public administration and research to improve public policy.

c) Sector level data sharing: As the name suggests, this bucket includes those data-sharing frameworks which focus on only specific sectors such as transport, health, payments, etc., which can be both related to the public interest as well as for the private sector.

d) Public Interest and facilitating private sector sharing: This bucket includes frameworks covering the objective of achieving both public interest and facilitating private sector sharing.

Figure 2: Benefits of sharing data, authors' analysis (same as presented previously)



Analysis (insights derived from Rationale as presented in the comparison matrix)

Mapping of the benefits of data sharing, as depicted in **Figure 2,** denotes that frameworks specifically aiming to facilitate private sector sharing are primarily in the form of contract guidance and are more recommendatory. On the other hand, most frameworks that are in fact in regulation focus on specific sectors. And other frameworks which have a diversity of focus are being proposed as a more holistic strategy, reliant on used cases and recognising pre-requisites for realising the benefits of sharing.

Even the proposed EU Data Governance Act stipulates 'data altruism' for facilitating private sector data sharing. This approach differs from the Indian Report, which proposes for regulatory framework mandating data sharing for public interest purposes without clearly linking the benefits of sharing to evidence of market or regulatory gaps to justify its approach for mandating data sharing.

The envisaged benefits of sharing data are also closely related to regulatory, market and technical gaps within the data ecosystem. (as depicted in **Figure 6***)*. As shown in Figure 6, most of the framework that lies in the last bucket focusing on facilitating private sector sharing and serving public interest purpose has conducted consultation and impact assessment to analyse the regulatory, infrastructural and market imbalances.

Moreover, those frameworks that are aiming to facilitate private sector sharing have identified particular gaps either of market imbalance or infrastructure or lack of trust amongst private players rather than identifying broad market imbalances without adequate evidence.

This is to emphasise that the expected benefits of sharing are identified based on the gaps in the data economy or the ecosystem. For instance, the industry feedback in Singapore recognised a lack of trust amongst private players to share data, which led to a trusted data-sharing framework that specifically aims to provide guidelines for building greater trust in data sharing.

Similarly, the impact assessments and consultations conducted for the EU Data Strategy identified market imbalances and supply-side infrastructure gaps. Therefore, the strategy focuses on making EU-wide sectoral data spaces that act as infrastructures and assesses the mechanisms and principles to facilitate private sector data sharing for businesses and engage with the infrastructure.

Another analysis is also related to the dimension of value assessment (as depicted in **Figure 4**). The frameworks that have envisaged benefits in sharing data for the public and private sector have also relied on used cases, established pre-requisites, and principles essential for achieving these benefits. For example, the Australian data-sharing framework focuses specifically on public interest and has also identified the need to build trust, ensure consistency, and stipulated specific areas for deriving benefits of sharing.

In the Indian context, while the benefits of data sharing have broadly been identified, its corresponding assessment of the regulatory, market and infrastructural gaps is missing. Moreover, the value assessment and identification in used cases or essential pre-requisites are not adequately approached.

Figure 6: Assessment of gaps and authors' analysis

2. Targeted Market, Regulatory and Infrastructure Failures

Questions from the Baseline Scenario Assessment of the Report:

- ***Are there any pieces of evidence of market, regulatory and infrastructural gaps which the report aims to address?***
- ***If there is any evidence if the government can tap into the value of data, and as such, the focus should be now on private-sector data sharing?***
- ***Is the government, regulator or a data trustee, or any other body equipped to create balance in the data economy? Are there any used cases in this regard?***
- ***Are smaller companies within the data economy equipped to leverage the opportunities presented on data sharing? Is there a first need to improve their efficiencies in terms of infrastructure, contractual practices, etc.****?***

Market, technical and regulatory gaps being addressed by different frameworks are represented in the form of a grid, in **Figure 6,** as represented above, to highlight analyses which have been in the form of impact assessments, consultation or industry feedback, and the kind of specific or multiple gaps that have been highlighted through. These gap assessments are then used to formulate benefits and the targeted objective of sharing the data. Each gap has been assigned a different colour scheme and the hexagons with multi-coloured borders represent multiple gaps identified by the assessments.

Analysis (insights derived from gaps as presented in the comparison matrix)

Each framework is attempting to address multiple and varied gaps within its data ecosystem. In most of these jurisdictions, a prior impact assessment was undertaken to determine the problems concerning inefficiencies in the data markets and the potential loss of value and market distortions that the policy needs to address. The assessment further exhibited that consultations, observations, surveys, and industry feedback were needed to determine the appropriate intervention, varying from soft regulation to more stringent regulation.

In the EU, policies introduced in context or complementary have tried to address the gaps with previous policies. Jurisdictions have also left room for voluntary initiatives like GAIA-X in the EU and iShare in the Netherlands to address the market failures amongst themselves. The responsible legislative authorities support these initiatives like the European Commission and the Dutch Parliament. However, they are not explicitly regulated as they follow the principles established by the existing policies.

The EU has, however, more directly made interventions for specific and targeted market failures. The Framework for Free Flow of NPD has targeted the high prohibitive costs imposed by the cloud service providers based on data localisation and compliance

requirements. The framework recognises that this can stifle innovation amongst small and Medium Enterprises (SMEs) and therefore target it accordingly.

Similarly, the European Strategy for Data ascertained that data accumulation by a handful of companies results in creating market imbalances while also creating interoperability issues. The Strategy, therefore, proposes standards and tools to be developed for the data principals to exercise their data rights in compliance with other European Data regulations like GDPR.

In Japan, the Contract Guidance on Utilisation of AI and Data aims to propose best practices as a survey conducted by a think tank indicated that a significant portion of the companies is facing problems related to leakage of trade secrets unauthorised use of data. The policy has therefore sought to address these leakages.

Similarly, the Japanese government also introduced the Act on Special Measures for Productivity Improvement, which relied on the analysis conducted by the Economic Policy Package, indicating the slow growth in specific sectors. To bring the growth rates back up, the Act calls for an increase in investments in specific sectors, including the use of IoT and AI. And to achieve those higher investments, the act proposes incentives for voluntary data sharing to foster innovation in these sectors.

In the Indian context, while the Report assumes that there are market imbalances and data accumulation, it fails to justify this assumption based on evidence. Thus, it eventually fails to link mandatory data sharing for the private sector to achieve public interest purposes. The Data economy is volatile and is growing rapidly.

As indicated by the experience of the EU, it would be beneficial to start from a holistic strategy that lays out plans for subsequent years and, in that process, build its policies and laws on evidence. This would also adequately highlight the gaps in government and regulatory capacities which need to be addressed at this stage.

3. Market Mapping and Need Assessment

Questions for analysis derived from baseline scenario assessment of the Report:

- ***What is the value of the data economy in India? What has been the contribution of data empowered industry on the GDP?***
- ***How many data companies currently exist in the economy? What is the existing level of innovation in the market?***
- ***How is data being valued in existing business transactions? How will these transactions be affected by the framework proposed by the Report?***
- ***Is it beneficial to first have a sector-specific focus on data sharing considering the sectors that see more growth?***

- ***Whether mandatory data–sharing leads to more innovation and are there any use cases indicating this? Is there a risk of anti-competitive effects of mandatory data-sharing? Will this also increase the risk of stifling innovation?***
- ***What are the risks of possible misalignments and risks in mandatory data sharing for markets? Are there alternatives to achieving such purposes?***

In **Figure 5**, the buckets of commonalities have been understood as below:

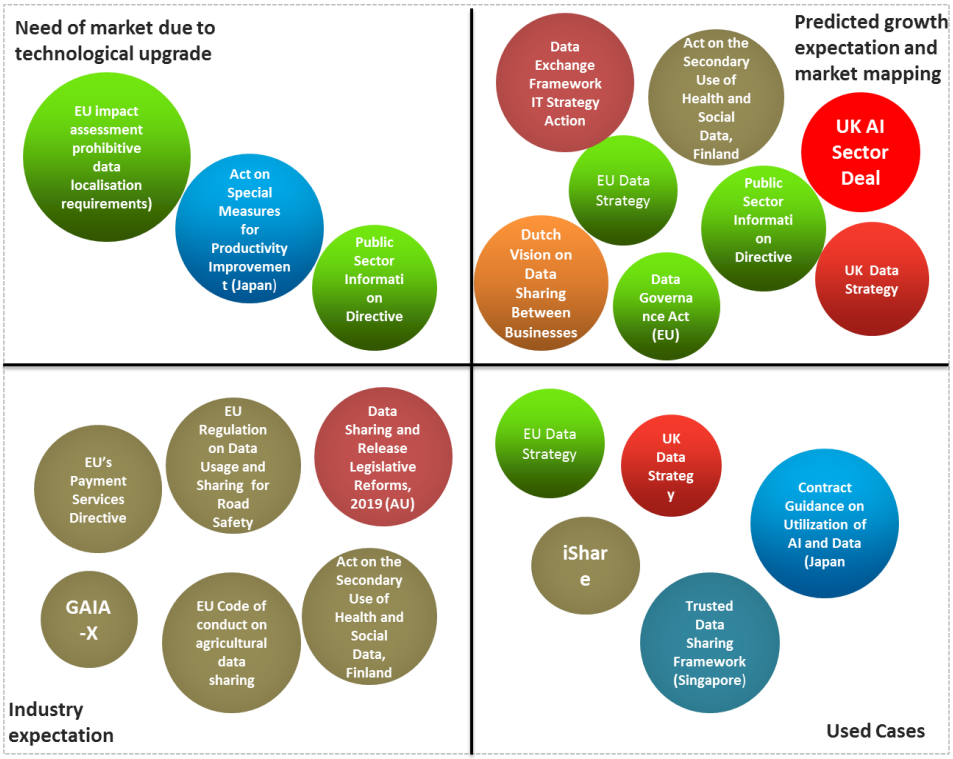
a) Need for Technological Upgrade: This bucket represents those frameworks that have come about after assessing the need to revise regulations or adopt new regulations to internalise technological upgrades in the market.

b) Market Mapping: This bucket includes those frameworks after assessing expected value generation from data sharing in quantifiable terms (such as contribution to the Gross Domestic Product or revenue). This assessment also involves mapping the current data usage and sharing trends within the market.

c) Industry Expectation: This bucket represents those frameworks that have resulted from the industry itself realising the value in data sharing and wanting to develop a shared understanding of data sharing practices.

d) Use Cases: This bucket includes those frameworks which have attempted to assess market readiness and usefulness of data sharing based on 'use cases' where data sharing has added positive economic value.

Figure 5: Parameters of market mapping/ need assumptions, authors' analysis



Analysis (insights derived from market need as presented in the comparison matrix)

One of the key observations from **Figure 5** is that most regulatory frameworks have resulted from market needs being pushed by the industry itself or explicitly targeted on technological development aspects. For instance, the Public Sector Information Directive, more popularly known as Open Data Directive, came about due to digital transformation in business models, which necessitated real-time access to dynamic data, which led to the directive to introduce technical and API guidelines for making public sector data available.

Along with this, the impact assessments also highlighted the expected value creation from such changes in the directive. On the other hand, the frameworks which have based their market readiness and need an assessment on the market mapping are more holistic and recommendatory. For many of them, for instance, UK Data Strategy, consultations are still underway to assess industry expectations and demand.

Market mapping and need assessments are again closely related to the assessment of market failure, which also highlights the technological trends and gaps in the market. In this regard, the consultations and impact assessment on the EU Data Strategy and also the report of the High-Level Expert Group on Private Sector Data Sharing for Public Interest highlighted that mandatory data sharing can result in disproportionate effects and therefore it is important to have specific guidelines in place and every case should be judged on its own merits to assess if data- sharing would be required. Additionally, even if data sharing would be required, companies would be compensated for sharing such data.

In the Indian context, parameters for assessing market trends and any quantifiable assessment of expected economic value creation are missing. Additionally, it might also be relevant to have more consultations with the industry to assess their expectations and the need for sharing data for public interest to avoid unintended consequences. In this context, it may be beneficial to take a more staggered approach as the EU, i.e., first to have a more comprehensive framework, and then assess specific needs based on impact assessment and wider industry consultations.

4. Policy Maturity

Questions from the Baseline Scenario Assessment of the Report:

- ***How will the new regulator in the form of the Non-Personal Data Authority (NPDA) and new intermediaries in the form of data trustee facilitate data sharing?***
- ***Will the risk of over-regulation and over-laps further be stifling innovation and investment in the sector?***
- ***Should the strategy for open data sharing be the first step for NPD sharing framework to follow? And should the implementation and the progress of existing policies such as the National Data and Accessibility Policy, National AI strategy,***

Strategy for National Open Digital Ecosystems (NODE) should be assessed before the implementation of an umbrella framework for data sharing in India.

- ***Do we need sectoral level frameworks before prescribing a cross-sectoral level framework?***
- ***Do all stakeholders in the data ecosystem operate at the same level of policy maturity? Is there a requirement of interoperability models or standards required for the horizontal applicability of data sharing principles?***

In **Figure 3**, different policies across jurisdictions are plotted to segregate the kinds of policy maturity they have. As observed, it is difficult to assign a level of policy maturity to most frameworks as they have relied on multiple levels of policy understanding and maturity, therefore, intersecting across the plotted axis. The levels of policy maturity have been divided into four buckets:

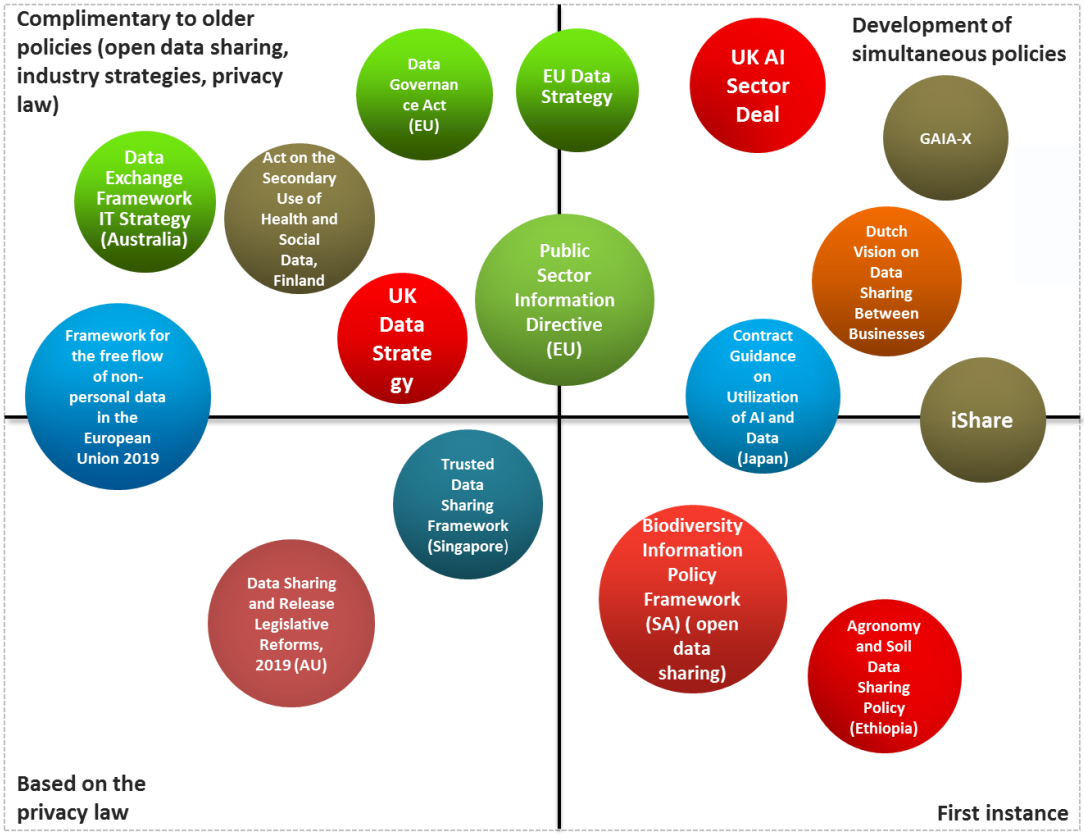
a) Complementary to Existing Policies: In this bucket, policies and frameworks based on existing laws and policies are placed. Examples of existing laws and policies include open government initiatives or industry-level strategies for data sharing.

b) Simultaneous Policies: Policies that were developed in coordination with other data-related policies are placed into this bucket. These include cross-sectoral regulations, inter-jurisdictional sharing, and industry-specific sharing frameworks. It can be observed between these and the last bucket; umbrella multi-jurisdictional EU policies are placed. These policies allow for sectoral level and individual jurisdictional frameworks based on existing and complementary policies.

c) Privacy Law basis: In this bucket, data sharing policies and initiatives that rely on privacy laws or privacy principles already in force are placed. There are a few policies that have not only relied on privacy first principles but also on other data policies to come up with new data sharing frameworks.

d) First Instance: Policies that have not relied on any existing laws and policies related to data, or privacy, are placed in this bucket. As observed, all these policies are either open government initiatives or voluntary sectoral level frameworks. Despite this, a few of these frameworks have relied on the development of new data policies simultaneously.

Figure 3: Policy maturity, authors' analysis



Analysis (insights derived from policy maturity as presented in the comparison matrix)

It can be observed in the multi-jurisdictional assessment that there is an evolving regulatory space for data and digital policies. While some jurisdictions are also at the nascent stages of regulations as India, other jurisdictions have had a substantial history of regulating parts of data or the digital economy.

The most common policy which acts as a precedent in all these jurisdictions is privacy legislation, policy, or guidelines. A privacy framework acts as a precondition to jurisdictions across the board. It can be observed that all jurisdictions that have introduced non-personal data sharing policies have robust privacy laws and principles. Jurisdictions like Australia have had a Privacy law since 1988.

Using the privacy law as the basis for data sharing principles, the Data Sharing and Release Legislative Reforms were introduced in the parliament in 2019, providing data sharing standardisation across the board to allow for interoperability and wider access.

Similarly, the European Union has had some kind of privacy regulation since 1995 in the form of the Data Protection Directive, which was later amended, and repealed with more comprehensive privacy regulation in the General Data Protection Regulation.

In cases where data sharing has been introduced as a first instance, such policies have been aimed towards sectoral level open data framework, which is highly specific, and for

a specific public purpose. This is apparent in both the Biodiversity Policy in South Africa and in the Agricultural Data Sharing Policy in Ethiopia.

In the Japanese Contact Guidance on Utilisation of AI and Data, which has been seen as a first instance for data regulation, the policy has been designed in tandem with privacy legislation and broader data legislation. Further, the policy is aimed towards the particular sectors of AI and provides incentives for data sharing among the participants.

Policies in other jurisdictions have either relied on previous data policies in addition to privacy laws, or the jurisdictions have come up with simultaneous policies to govern data in different sectors for highly specific purposes or provide breathing room for voluntary data sharing frameworks. Reliance on complementary policies as well as simultaneous policies is apparent in Australia and the European Union.

Australian Privacy Law and the Legislative Reforms led the way for Data Exchange Framework. Similarly, the GDPR, Data Directive, and the Data Governance Act have led to the EU Data Strategy and the Free Flow of Non-Personal Data.

These jurisdictions have had the upper hand in regulating data and facilitating sharing because of their experience and learning from the previous policies. These policies have also served as a backbone for data sharing models. They have also helped these jurisdictions place checks and balances and grievance redress mechanisms appropriately, constituting a prerequisite for data sharing. The policy maturity achieved from existing data policies also serves as use cases for future regulations, the practice that has been heavily relied on by the EU.

If the NPD Framework as proposed by the CoE were to be placed in this matrix, it would be placed in the bottom right corner of this bucket, based on its unfounded policy maturity. However, the framework would be one of its kind even in this bucket. All the other policies are either open government initiatives or voluntary and sectoral, unlike the NPD, which seeks mandatory and cross-sectoral umbrella sharing.

As has also been highlighted in another section, policy maturity forms an essential basis for determining the value of data and the nature of its use. In this context, many of the questions and ambiguities regarding the role of the new authority and data trustee could be adequately addressed if we could learn from the policy evolution in the data ecosystem.

Conclusion

This Comparative Jurisdictional Analysis has examined the rationale and objectives of existing and upcoming data sharing policies, frameworks, and initiatives across six jurisdictions, including nineteen data sharing frameworks. These data-sharing frameworks include umbrella as well as sectoral frameworks. Overall, the comparative analysis indicates that the Report in India requires further analysis and evidence gathering to make the 'problem statement' and objective clear.

One of the prominent gaps within the Indian Report is the lack of market mapping and policy maturity, which has led to vagueness in prescribing value and benefits intended to be derived from data sharing.

Some of the prevalent practices in other jurisdictions and their usefulness within the Indian context are mentioned below:

Nature of Data: Under the rationale and goals, the CoE places reliance on data's economic value and has prescribed High-Value Datasets to be shared as a public good for broadly stipulated public interest purposes. In doing so, the Report, while intending to achieve public interest purposes, gives data a 'public good' treatment to forge due consideration to unintended incentive misalignment.

The comparative analysis shows that data in jurisdictions wherein data is attributed as Public Good to be available for the public interest, there is no mandated data sharing. In cases where sharing is voluntary or sector-specific, data has been treated as a proprietary asset. It is notable that frameworks across jurisdictions realise the multifaceted nature of data and have considered that.

The comparative analysis highlights that prescribing nature of data is closely linked with expected benefits and value creation in data. Along with this, jurisdiction which prescribes data sharing 'for' public interest purposes distinguish the treatment of public-sector and private-sector data.

This distinction is made to internalise the incentive problem which emerges from prescribing private sector data to be treated as a public good for public interest purposes. The nature of data should be closely related to expected benefits and value rather than approaching it through the lens of the "one size fits all".

Value of Data: To assess the value of data derived from data sharing, frameworks have relied on various dimensions. These dimensions include *pilot tests*; *use cases of data sharing*; developing mechanisms to quantify the value of data; and *recognising pre-requisites* in the form of infrastructure, financial capacity, and policy framing, all of which can help identify the value of data.

While the technological architecture required for data-sharing has been recognised in the NPD report, other parameters like pilot studies or use cases have not been considered. In this regard, it would be beneficial to assess the impact of the open-data initiative in India in creating value (social, economic and public) along with assessing other private sectors initiatives.

Benefits of Data Sharing & Related Market and Regulator Gaps: Jurisdictions have been trying to determine and develop frameworks that facilitate the benefits of sharing data. Thus, the frameworks have clearly identified the intended benefits of data sharing linking them to market and regulatory gaps, which have also led them to define their approach to data sharing.

These benefits range from facilitating *private sector and voluntary sharing*; promoting *sector-specific sharing*; specifically*, for the public interest*; and in some cases, for both *public interest and facilitating data sharing in the private sector*. In the Report, while there has been a brief discussion on the benefits of data sharing, the specificity in prescribing these benefits cannot be achieved due to a lack of market and regulatory gaps and impact assessment.

The comparative analysis shows that the frameworks have first identified the specific market failures they have intended to target, based on the impact assessments and evaluations, and then following up with consultations, observations, surveys, and industry feedback which was used to determine the appropriate policy intervention, ranging from soft to stringent regulations.

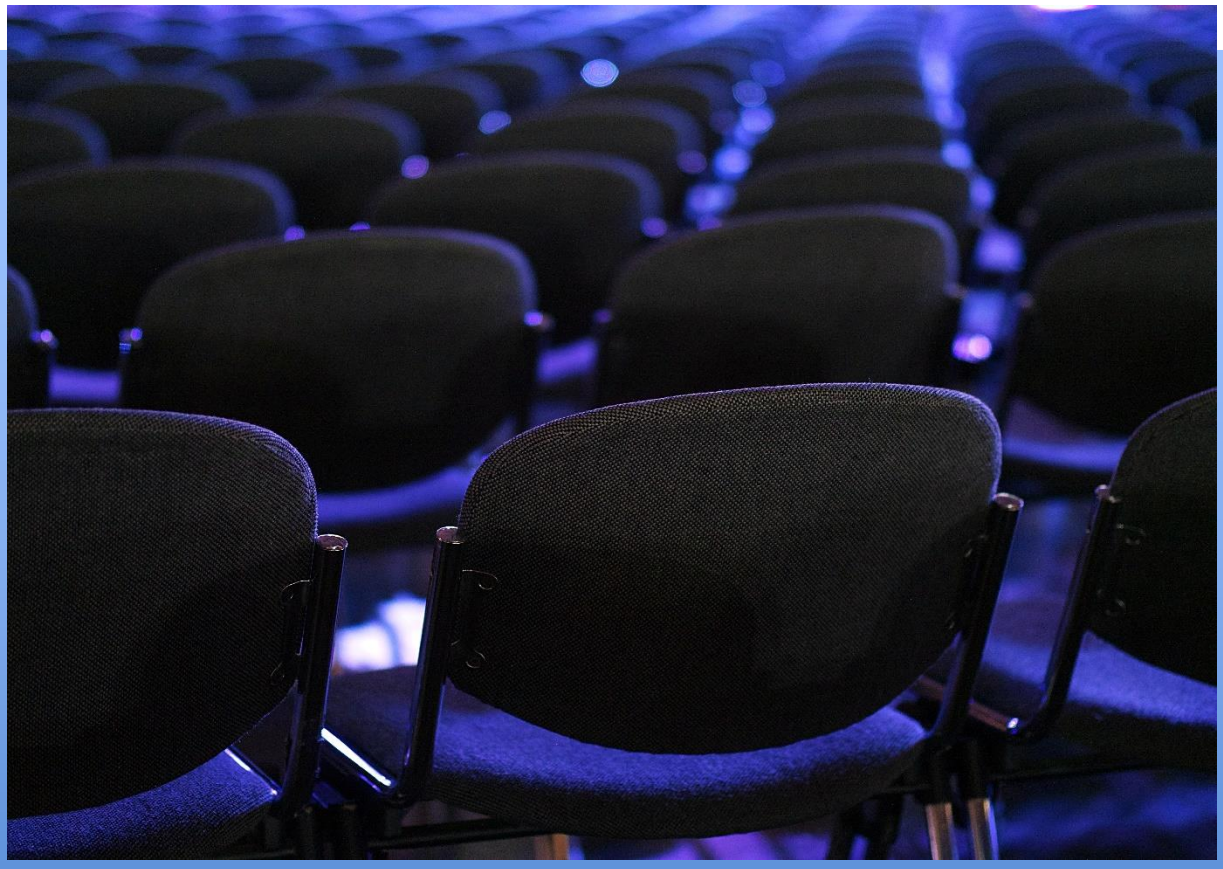
Market Mapping and Need Assessment: The comparative analysis further exhibits that a comprehensive market mapping and need assessment were relied on by data sharing frameworks across jurisdictions. These assessments were further used to eliminate risks of misalignments with the needs of the markets and the proposed policy intervention. These assessments recognised several factors indicating the need for regulation. These include the *market needs due to technology upgrades; industry expectations, use cases; and quantifiable market mapping*. However, the Report does not rely on any such assessment or mapping to highlight market needs and avoid possible misaligned incentives for data markets.

Policy Maturity: All framework assessed proposes a data-sharing approach that has been built upon experiences gained from a certain level of policy maturity. The only exceptions to this are the proposed policies for *open data initiatives*. In all other instances, the data sharing policies and frameworks have relied on *privacy laws, complementary policies* in the form of aforementioned open data initiatives and industry strategies; or more commonly in harmony with other data-related policies *simultaneously designed*. Even the developing countries such as South Africa and Ethiopia have also initiated building open data policies.

In the Indian context, the suggested framework cannot fit into any of these definitions. There are also explicit concerns regarding the distinction between personal and NPD. While the proposed framework by the Report is being deliberated at the same time as the PDP Bill, it is essential to note that all the evaluated cases already have a basis for privacy principles before they approach data sharing.

Based on this multi-jurisdictional dissection of data sharing policies and frameworks from across the world, it is clear that the Report has several parameters, dimensions, and perspectives to first identify, assess, and evaluate before recommending any form of data sharing. It is pertinent to note that while India does have an open public sector data policy in the name of the National Data Sharing and Accessibility Policy (NDSAP), it has not been properly enforced. Had this policy been widely applied, the learning from this could be used as a use case, but this would have also provided the level of policy and market maturity and the level of understanding of the data market that is missing in the Report.

Analysis of Stakeholder Consultations



Introduction

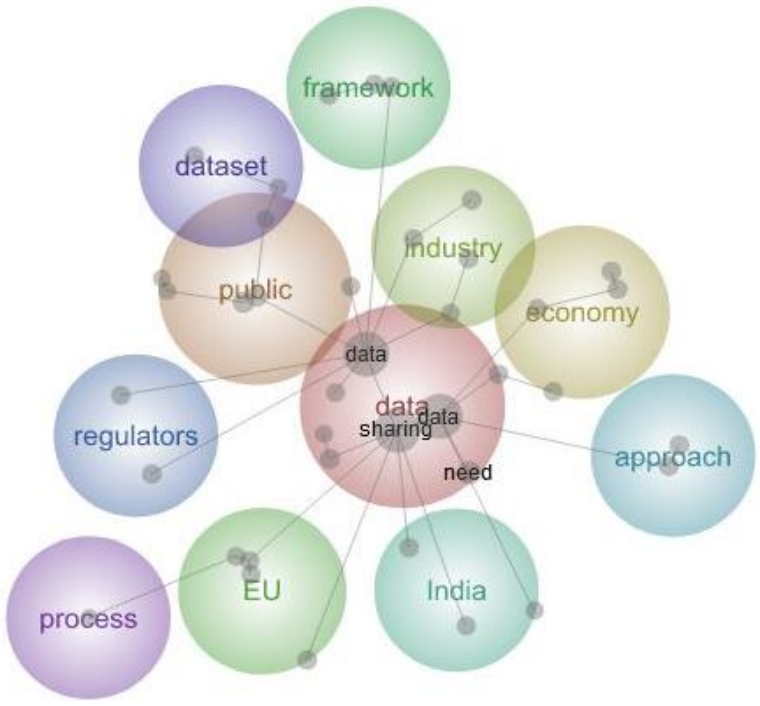
The report of the CoE on the Report will impact a wide array of stakeholders, including start-ups, small and medium-sized enterprises working with data, consumers, data processors, data trustees and government entities. Through about 20 key informant interviews and a roundtable discussion with national and international experts, industry representatives, consumer rights and civil society representatives, we were able to gauge the perspectives of these stakeholders regarding the rationale, assumptions and approaches to the NPD sharing framework proposed in the Report.

These consultations were conducted as semi-structured interviews and inputs were also taken from the online roundtable discussion conducted by Consumer Unity & Trust Society on Future of Data Governance from a Consumer Perspective, including national and international experts and policymakers.[108]

Method for Analysis

The analysis of these consultations was conducted using Leximancer,[109] a text analysis software to identify relevant concepts and assessing their linkages.[110]

Figure 7: Concept analysis, derived from Leximancer text analysis



[108] https://cuts-ccier.org/webinar-on-future-of-data-governance-in-india-a-consumer-perspective/

[109] https://www.pacifictranscription.com.au/2013/08/30/using-leximancer-for-qualitative-analysis-of-transcripts/

[110] Sotiriadou, P., Brouwers, J. and Le, T.A., 2014. Choosing a qualitative data analysis tool: A comparison of NVivo and Leximancer. Annals of Leisure Research, 17(2), pp.218-234, Haynes E et al., "Semiautomated Text Analytics for Qualitative Data Synthesis.," *Research Synthesis Methods* 10, no. 3 (July 9, 2019): 452–64, https://doi.org/10.1002/jrsm.1361.

Figure 7 is a pictorial representation of the text analysis of comprehensive notes of stakeholder consultations. It indicates different concepts highlighted by the stakeholders, with the lines indicating connecting thoughts and comments related to different concepts. For example, data-sharing and economy; the nature of data as 'public-good for public-interest, etc. These concepts are identified based on the number of hits a particular word has in the conversation (in this notes). The authors have further coded the identification of the concepts in a way that words such as "take", "and" "need" can be combined with other concepts rather than them appearing as part of the analysis.

Summary of the Analysis

Stakeholders identified that prescribing the "one size fits all" approach adopted by the Report may not be well suited for the data economy, as different types of data attract different benefits, sensitivity and value. Moreover, the compliance for any data sharing rules can only be applied if there is an industry consensus and citizen impetuous, thus rather than rule prescription, soft nudges would be required for internal commitment with industries to come on the same page. On the other hand, even for consumers, their role as "data citizens" needs to be understood and internalised.

Furthermore, as highlighted in the consultations, the first step should be to map the kinds of data being collected by companies; how they treat their data; and existing data-sharing practices; so that a specific problem statement can be formulated for identifying the right infrastructure for addressing problems within the data-sharing ecosystem. Additionally, even for prescribing a regulatory mechanism, the questions regarding 'what kind of regulator' and 'what is expected out of the regulator' need to be identified.

Analysis of Specific Concepts

- **'Public good' framing of data and sharing data for public interest**

  It was indicated in stakeholder consultations that the report has confused between data "as" public good and data "for" public good. The good public term for non-personal data (NPD) may not be apt as excludability can be created in data through existing intellectual property rights, which would create incentive problems. There also exist differences between how companies treat data. For some, it may just be a by-product, while for others, it can be an important resource.

  In such cases, fixing incentives can be complicated and would vary on a case-to-case basis. This indicated that the Report interprets the data "as" a public good or "for" the public good in a vacuum. Moreover, such framing seems like just a mechanism through

which legal hurdles and proprietary interest in data could be bypassed, such that its economic value can be leveraged.

On the point of data sharing for the public interest, it was indicated that unintended consequences that may accrue from mandatory sharing of data for public interest need to be accounted for. If a certain dataset is already available for the public interest, it will discourage business-to-business (B2B) sharing, which could discourage the development of data markets. Another critical observation was on the lack of clarity in identifying beneficiaries for the public interest.

Since the value of data changes very quickly, a dataset that may not be valuable now could become high-valued datasets (HVD) in the future and would attract sharing obligations. In such cases, organisations which work on sensitive political and social issues, sharing such data before their actual analysis could be very problematic and may attract undue scrutiny from the government. Even for some start-ups, this could raise flags with investors, as they could now compare and question companies on publicly disclosed data and the data that has been made available to them, which may lead to possible loss of investments.

- **Regulatory Mechanisms**

  Stakeholders highlighted that the current regulatory capacity might not be well suited for governing data-sharing, which is evident from experience with telecom and broadcasting regulators. Moreover, it was emphasised that there is a need for a more holistic regulatory approach and open data policies.

  This presents two options, i.e., to either develop existing regulators' capacities or establish a new regulator. However, stakeholders cautioned that while having a new regulator could be an option, having multiple regulators within the data space would lead to overlaps and confusion.

  On the issue of regulatory overlaps, it was pointed out that conflicting narratives are appearing from the Joint Parliamentary Committee on Personal Data Protection Bill (PDP Bill) and CoE on the NPD governance framework. Thus, we need to wait to assess how the debate will evolve after a Personal Data regulation is implemented.

  Regarding identifying the role of the competition authority and sectoral regulators, the objective of data sharing needs more clarity. This is because if the objective is to open up data for the public interest, then a separate regulator would be required; however, if the objective is to have equitable distribution of data for economic interests, competition authority could play a role.

  On the other hand, while sectoral regulators may not be capacitated to take every perspective into account, they can shed light on the nuances of differing business models within a particular sector. Overall, evolving effective approaches to regulatory

collaboration and being sensitive towards compliance burden is essential as the data ecosystem is complicated and involved multiple layers.

- **Industry Dynamics**

  Shedding light on the current data management practices, stakeholders indicated a lack of data exchange. There also exists a grey market for data exchanges, where while there is the availability of data, quality and security standards in such datasets are questionable. This indicates that while data availability is not a problem, its under-utilisation and maintaining the quality standard is an issue.

  The industry representative pointed data sharing as a remedy cannot be prescribed in a vacuum without assessing the quality of data and its security. At the same time, they also cautioned that how the Report envisaged data sharing seems like a wish-list, with a lot of uncertainty and vague concepts. This requires the CoE to broaden its scope and assess the data management practices of companies to understand the nuances of data exchanges and the differential impact of data access on different entities. For instance, while it would open opportunities for data processors but would be challenging for data custodians.

  Additionally, with limited understanding and uptake of data exchange practices, most companies would find it difficult to integrate data-sharing within their operations. Therefore, the starting point of understanding these nuances should be to map the dynamics of the data economy, organisational posture and existing data use.

- **Incentives**

  Stakeholders indicated that prescribing mandatory sharing of HVDs would create incentive problems disincentivising businesses from investing in data collection. This problem stems from making data available for the public good, devoid of intellectual property recognition such as trade secrets held by companies in datasets.

  Furthermore, it was also emphasised that fair and reasonable charges as prescribed under the Report would not be adequate and would come off as an obligation. International experts indicated that this is a contrast from the EU, which has stayed clear from the mandatory sharing path as EU laws recognise trade secrets protection. So industries are confident, they would have a *suo-moto* legal safeguard. Industry representatives pointed out that data custodians would fight this approach if incentive alignment is not ensured.

  Furthermore, datasets may reveal business strategies and competitive insights, which would also disincentivise businesses from sharing data.

  Geopolitically, this stance on mandatory data sharing could also set a tone for other developing countries, which may rely on India's data governance framework; this puts an

additional responsibility to closely evaluate the realities of the industry dynamics to ensure adequate incentives and pre-requisites are available.

- **Privacy Risks and Implementation Constraints**
  One of the pertinent concerns throughout the conversations with stakeholders was the blurred categorisation between personal and NPD. International experts highlighting the European experience cautioned that the differentiation between the two is becoming thin, which presenting regulatory challenges. Additionally, stakeholders also indicated doubts in adopting anonymisation techniques as there always exists risks of re-identification and without an adequate privacy framework, such concerns may become difficult to address. Moreover, even with anonymisation, datasets can reveal information about businesses and their strategies and when these datasets are combined with other datasets, they may present risks of profiling. Thus, the aspect of sensitivity attached to NPD cannot be negated.

- **Data Trustees & Community Rights Framework**
  Stakeholders pointed out that they envisage the relationship between data trustees and data custodians would be adversarial. The framework would emerge as self-regulatory mechanisms for the industry player to come together for data-sharing for mutual benefits. And there may emerge a tendency in the industry wherein they may be more reserved for sharing data with consumer groups or civil society in cases that may not serve their mutual benefit.

  For the community rights framework, it was highlighted that there is a lack of clarity on the criterion on which community could be defined. Along with this, stakeholders cautioned that conflicting interests might emerge amongst communities creating difficulties for data trustees to adjudicate and balance such interests. For example, in payment services, one community can be merchants and other consumers, which may present rivalrous interest. Thus, without a clear definition of concepts and criteria for community interaction amongst each other and data trustees, data chaos would emerge.

Conclusion

The stakeholder consultations as indicated above highlight some challenges with rationale, assumption and approaches concerning different aspects of NPD sharing. Overall, there was a consensus that governing data-sharing is a complicated process. There is a need for broader consultation and a comparative assessment of Indian realities before taking umbrella measures and prescribing stringent conditions of mandatory sharing. Some of the key recommendations that emerged from the consultations are:

- **Mapping industry practices** – Before diving into formulating a regulation for NPD sharing it is necessary to map industry practices of data management and sharing. This would shed light on organisational practices and help identify nuances related to anonymisation techniques; identifying costs in segregation and collection of data; quality standards required for data usage, and securing data, etc.

  Mapping these practices could help identify particular sectors, activities and benefits, which could be a better starting point for prescribing data-sharing rather than conflicting it with business interests. For example, starting with weather and geospatial data collected by the public sector.

- **Need for a 'balancing act'** – The data economy is moving at a very fast pace. For the Indian economy to hop on this boat, a balancing act would be required throughout the policy formulation process. It would be beneficial to start by identifying wherein the data value chain there is breakage. Data access is being hindered so that appropriate solutions can be devised to address that, rather than take a "one-size fits all" approach.

  At the same time, justifications for having tradeoffs between data sharing and privacy; private and public interest; mandatory and voluntary sharing need to be clearly stated. For this, the framing of such policies should be an evidence-based and continuous process, which should be ramped up step by step considering technological advancement.

- **Purpose Framing** – As is evident from stakeholder consultations, there seems to be a consensus that public interest and community rights of framing in NPD are to bypass checks and balances. In light of this, it is important to bring in more specificity to purpose framing such that certain limitations or safeguards could be in place for justifying the purpose of sharing NPD. For this, evidence of data usage for observed problems, such as using data for COVID prevention measures or disaster management, should be identified.

  Thus, it was recommended by stakeholders that instead of submitting all kinds of HVDs to the data trustee, data custodian may be required to submit only a few datasets which qualify to be highly relevant for public interest purpose for identified problems. Along with this, there is a need for precise identification of the entities which would define these 'public interest purposes' for qualifying datasets to be HVDs. For this, a more consultative policymaking process should be evolved through which sector regulators and experts who have more expertise in data utilisation could be involved. And lastly, clarity is needed in clearly identifying intended beneficiaries i.e., to define communities adequately.

- **Building Regulatory Capacity –** As highlighted in the previous section, there are two parallel views on regulatory capacity. While some stakeholders pointed out that

introducing a new regulator would increase compliance costs, the focus should be on increasing the capacity of existing regulators.

On the other hand, it was stated taking a holistic approach and having one regulator for data governance would be beneficial. However, these views must be interpreted in light of the debatable distinction between personal data and NPD, which inadvertently makes it difficult to have a regulatory demarcation and therefore, the overlaps amongst regulators should be avoided.

Another consensus that emerged was to focus on increasing the capacity of existing regulators and build an efficient mechanism of regulatory collaboration, such that sector-level expertise and nuances could be internalised in the adjudication process.

- **Policy Roadmap for Implementation –** Stakeholders recommended that an essential step in forming a roadmap for data-sharing is to first have laws and safeguards that can guarantee data protection and privacy for consumers and adopt a rights-based data sharing approach.

  To rectify the problem of the blurred distinction between personal and non-personal data, it was recommended that data should be categorised as personal and aggregated data, which can make the implementation of data protection regulations clearer.

  Thus, while these constraints should not deter in facilitating data sharing, India should adopt an adequate policy roadmap that starts from making public-sector data available and then moving to private sector data sharing, as public sector data is collected on taxpayers' money in a usable manner. Along with this, having appropriate data protection legislation in place and strong linkages to existing policies is also a necessary prerequisite.

ANNEXURE I

Comparison Matrix Examining Rational and Assumption to Data Sharing

The comparison matrix presents an overview of the rationale and objectives of the existing and proposed data-sharing framework/initiatives/strategies (primarily non-personal data, and at times also including personal data) across jurisdictions in a similar format. The aim is to compare this for further assessing the rationale and objectives of the Non-Personal Data Governance Framework for India, as suggested by the Committee of Experts chaired by Kris Gopalakrishnan. To conduct an in-depth assessment of the rationale, the matrix follows the following design –

Vertical Axis: Mapping in Different Jurisdictions

Cross-sectoral and umbrella frameworks/initiatives/strategies/ laws by public authorities for data sharing across jurisdictions (countries or recognised coalition of countries such as the European Union)

(This will cover those initiatives which are applicable across sectors for data sharing and access and are not limited to one sector or area. This covers the initiatives which are undertaken by the government bodies or public institutions.)

Sectoral frameworks/ initiatives/ strategies/ laws by public authorities for data sharing across jurisdictions.

(*This covers sector-level frameworks of data sharing such as agriculture, payment systems, and transport, etc. This also includes the sector level frameworks which are undertaken by public institutions, authorities or government bodies.)*

Cross-sector and sectoral framework/ initiatives/strategies undertaken by the private sector.

(*This will include initiatives by the private sector in data sharing which covers both cross-sector and sectoral level within and across jurisdictions, thus, may also include global initiatives*)

Horizontal Axis: Parameters of Synthesis

Rationale and goals for establishing the framework
(Specifying the broader aim and objective of the framework)

Assertions behind the rationale
(Specifying the assertions, assumptions on basis of which the broader rationale was reached)

Targeted policy, market and regulatory gaps and failures
(Specifying the market or regulatory gaps, failures, and issues that the framework is trying to address)

Policy Maturity
(Specifying existing policies/ laws or frameworks existed during the development of current framework, and their role)

Market Mapping
(Specifying mapping of data economy, expectation of value economic contribution or value generation from data)

Process followed and timelines
(Specifying the consultation or impact assessment undertaken in the process of formulation of the framework)

Comparison Matrix

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
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[111]	**The framework's objective is to achieve data mobility, across member states** in Europe, which is currently inhabited by data localisation practices and mandates of the member states to enhance data economy and competitiveness in the industry. The framework also encourages the industry to come up with self-regulatory codes of conduct for sharing of non-personal data.	The Framework is established on the assertion that the development of the data economy in the EU is hampered due to - **imposing technological and legal requirements for storing of data in the geography of specific member states; lack of trust; legal uncertainty**[112] **and other vendor lock-ins (cloud service providers) practices. These restrictions were also imposed by the private sector through contractual and legal requirements to switch service providers.**	The impact assessment reports also revealed problems concerning inefficiencies in the data centre sector, loss of growth and innovation potential, loss of operational efficiency, and market distortions that need to be addressed. **The assessment illustrated the high prohibitive costs that were imposed by cloud service providers for SMEs for switching data which was further aggravated due to localization requirements,**[113] **market**	The rules proposed by this **framework were made complementary to the established provision of the GDPR, through which free movement of data across borders could be facilitated.**	The factsheet prepared for the Framework revealed the current and predicted functioning of data flows and their impact. It noted that lower cost of data services and more flexibility to companies could boost EU GDP by 4 % until 2020. **It also identified, predicted additional revenue in other sectors if the data localisation restrictions were removed.**[114] **Thus, indicating the effect**	**Before introducing the framework, impact assessment studies were conducted.**[115] The impact assessment specifically considered different scenarios from no- legislation, soft- legislative intervention to strong legislative intervention and its impact in the context of data flows across jurisdictions on social and environmental concerns. The impact assessment was conducted through

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

112 The impact assessment identified that there is a perceived existence of data localization requirements by businesses and public sector organizations, which limits there choice of location for data processing

113 http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=46844

114 http://ec.europa.eu/newsroom/document.cfm?doc_id=47000

115 https://ec.europa.eu/digital-single-market/en/news/facilitating-cross-border-data-flow-digital-single-market-study-data-location-restrictions

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
		These assertions were supported by the public consultation conducted on building a European Digital Economy, in which 62 % of survey respondents (businesses and organisations) supported removing data localisation restrictions and 55 % said that legislative action was necessary for doing so.	**economy and legal compliance.**		**on the markets from the existing baseline**.	wide consultation with stakeholders.
GAIA- X[116] (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 to turn them into a homogeneous, user-friendly system. The resulting federated form of data infrastructure will strengthen both access	The assertion behind this initiative is **the need to create a trustworthy, sovereign digital infrastructure for Europe**. It also aims to create a pan-European network of partners to ensure a reliable ecosystem for innovation, providing independence from relying on external infrastructure.[117]	This project aims to achieve **data sovereignty at the European level to decrease reliance on Chinese and American companies** for cloud infrastructures.[118] This has specifically come about as Europe follows higher privacy and security standards compared to its Chinese and American counterparts. However, this initiative is just seen as a starting and is envisaged to grow gradually in scale. [119]	Considering the legislation progressive **data regulations by the EU, the initiative comes timely, allowing it to claim protection under EU data laws like the GDPR while promoting its own guiding policies in harmony with the EU Digital Single Market strategy**.[120]	There are no European rivals to tech giants like Google and Amazon, the proposed interface seeks to create a unified ecosystem to link existing cloud services from European companies. Founding members have **conducted their internal research and concluded that the initiative will help launch competitive**	The project **is initiated with the support of 22 firms from across EU member states.** The project is also in alignment with the European Commission President Ursula von der Leyen's priority. It is expected to **launch a prototype in December 2020 and launch the**

116 https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html
117 https://www.data-infrastructure.eu/GAIAX/Redaktion/EN/Downloads/gaia-press-release-september-15th-en.pdf?__blob=publicationFile&v=3
118 https://www.zdnet.com/article/meet-gaia-x-this-is-europes-bid-to-get-cloud-independence-from-us-and-china-giants/
119 https://tech.newstatesman.com/cloud/gaia-x-cloud-project-peter-altmaier
120 https://www.techrepublic.com/article/what-is-gaia-x-a-guide-to-europes-cloud-computing-fight-back-plan/

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
	and share data securely and confidently. This initiative has also come as fostering the goals for EU Strategy for Data.				**offerings from Europe out onto the world market.**[121]	**completed project in 2021.**[122]
European Strategy for Data 2020 [123]	The **measures laid out in this strategy contributes to a comprehensive approach to the data economy to increase the use and demand for data and data-enabled products and services throughout the Digital Single Market**[124] **in Europe.** The strategy is designed to **facilitate access and re-use of data in the economy and society** while keeping those who generate the data in	The strategy emphasises the economic and public value of data. **At the same time, the strategy interchangeably uses the notion of public interest**[126] **and data for 'public good'. For achieving the objective of making private data available for the public interest, an expert group was established which proposed certain principles for such sharing.**[127] While assessing this value, the strategy accounts for the currents trends of usage of data in various sectors in the EU	The strategy highlights that there is not enough data available for innovative re-use, which is reliant on data holder and user and also the nature of data involved. **The strategies identify the following key issues based on their consultations and observation - e accumulation of data in the hands of few companies creating market imbalances; data interoperability issues within and across**	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, Database rights, Open Data Directive, Cybersecurity Act.** While introducing this strategy there was parallel guidance issued on private-sector data	The data market assessment tool which mapped the data economy in the EU indicated that currently, the **data economy accounts for 2.4 % of EU GDP, which is projected to grow to 5.8% in 2025.**[132] **Additionally, this strategy came after the European Digital Single Market policy was initiated in 2014, under which data economy and market**	This strategy **came about as a result of consultations, studies and assessments, which have been ongoing in Europe for a few years now**. The creation of the portfolio on 'A Europe Fit for the Digital Age' was created along with the vision of the European Digital Single market being discussed in 2015.

121 https://www.bmwi.de/Redaktion/EN/Publikationen/Digitale-Welt/das-projekt-gaia-x-executive-summary.pdf?__blob=publicationFile&v=6

122 https://www.euractiv.com/section/digital/news/gaia-x-group-signs-corporation-papers-in-bid-for-eu-digital-sovereignty/

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

124 A Digital Single Market (DSM) is one in which the free movement of persons, services and capital is ensured and where the individuals and businesses can seamlessly access and engage in online activities under conditions of fair competition, and a high level of consumer and personal data protection, irrespective of their nationality or place of residence. The 2014-2019 Commission had identified the completion of the DSM as one of its 10 political priorities. https://ec.europa.eu/digital-single-market/en/shaping-digital-single-market.

126 The basis of processing data for public interest has been stipulated in the GDPR. https://edps.europa.eu/sites/edp/files/publication/20-06-16_opinion_data_strategy_en.pdf

127 https://www.euractiv.com/wp-content/uploads/sites/2/2020/02/B2GDataSharingExpertGroupReport-1.pdf

132 https://ec.europa.eu/commission/presscorner/detail/en/fs_20_283

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
	control. Through this strategy, the EU seeks to empower the businesses and the public sector to make better decisions using data. [125]	such as areas of public interest, IoT, could and quantum computing. Moreover, the strategy bases its assertion on serving the need for the individual through creating value for the economy and society. **The strategy based its logic on used cases in the EU of data sharing** [128] **and their contribution to the increasing efficiency and saving labour costs**. Additionally, it also emphasises that the proposed strategy should complement the broader industrial strategy to create a data agile economy.	**sectors;**[129] **and supply and demand-side problems with data infrastructures** specifically with cloud providers which need to be addressed. Additionally, it appropriately identifies - the requirement of developing standards and technical tools for exercising rights of data principals as prescribed in GDPR and the e-privacy legislation; developing cybersecurity framework to ensure security within data value chains to develop trust amongst stakeholders; and promoting digital skills so that professional expertise can be built within EU.	sharing, which specifically notes the outcome of the public consultation indicating that at this stage the horizontal legislation for private sector data sharing is not necessary and the same should be proposed at a later stage.[130] There also has been sector-specific legislation and guidance in place for business to business and government to business data sharing.[131]	**assessment and release of periodic review of investments in digital infrastructure, artificial intelligence was undertaken** [133] The models of data sharing as proposed in the strategy are also inspired by the existing member state-led initiatives such as Finish and French Data Health Hub.	
Public Sector Information	The objective of this directive **is to make public sector data available for commercial**	The premise of the directive is based on the need to update the directive based on technological (emergence of	**The impact assessment revealed the following issues in public data sharing - dynamic data**	This **directive comes as part of the larger EU Digital Single market plan and**	The directive **has been introduced taking into account the evolving markets**	This directive is a revision of a 2013 directive and a replacement to the

125 https://ec.europa.eu/commission/presscorner/detail/en/fs_20_283
128 https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy#a-single-market-for-data
129 https://datalandscape.eu/data-driven-stories/what-limits-data-sharing-europe
130 https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0232&from=EN
131 https://ec.europa.eu/digital-single-market/en/guidance-private-sector-data-sharing
133 https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=53056

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
Directive (PSI) 2019[134] (also called the open data directive)	**and non-commercial purposes**. The framework establishes an open data sharing mechanism for sharing and re-use of public sector data to all entities and individuals. The Directive's key focus is on the economic aspects of the re-use of information rather than on access to information by citizens.	IoT and Big Data) and legal developments (GDPR and consultation of European digital economy and development of Digital Singles Market).[135] **The directive basis its assertion on the value of public data in developing internal market and act as a resource for the development of applications for consumers and legal entities.** It also adequately **identifies the capacity of the public sector to collect, reproduce, produce and disseminate a wide range of information in many areas in a machine-readable form**at. This would also promote transparency and accountability through feedback received from re-users and end-users on the data. To exploit the potential of public data, the focus is on - **providing real-time access**	**sharing through APIs is limited; overcharge on the re-use of public data which acts as a market barrier for new entrants; data from transport, utilities and data generated from public sector funding was not covered in the previous directive of 2013**; and lock-in arrangements between public and private sector which only benefits bigger companies.[137] Consultations further revealed the following [138] – non-uniformity of data use and sharing practices amongst member states; there exists a variety of licensing conditions which hampers effective re-use; costly redress procedures; the possibility of database right being used to restrict access and re-use; misunderstanding on	**builds on the existing Directive of 2013 for public data sharing**. The idea being the harmonisation of data governance and sharing mechanisms when the EU digital strategy comes into effect. The directive also aims to harmonise its provision with the database directive, GDPR and the ongoing consultations on the European Digital economy	**due to the digital transformation of businesses across sectors and is therefore focused on economic aspects of data use and re-use**. **The impact assessment noted that the direct economic value of PSI is to increase from a baseline of EUR 52 billion in 2018 to EUR 194 billion in 2030.**[139] This was based on assessing how much value each PSI component has added or is expected to add, which includes net effects of research institutions, APIs, para-public bodies, exclusive agreements and charging.	2003 directive which covered only specific public institutions. This revision in the directive comes at a time when Europe is gearing up for its data strategy. The proposal for the revision of the directive was made due to its periodic review, after which impact assessment and consultations were undertaken, making the directive updated on the consultation of European Data Economics.

134 https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L1024&from=EN
135 https://ec.europa.eu/digital-single-market/en/news/impact-assessment-support-study-revision-public-sector-information-directive
137 https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51645
138 https://ec.europa.eu/digital-single-market/news-redirect/621219
139 https://ec.europa.eu/digital-single-market/en/news/impact-assessment-support-study-revision-public-sector-information-directive

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
		to dynamic data via adequate technical means; increasing the supply of public data, and public undertakings and research organisations to tackle new forms of arrangements for sharing data. The directive encourages member states to ensure transparency in data sharing and public information and to ensure open access to publicly funded research data at all levels.[136]	appropriate techniques to be used for pseudonymisation or anonymisation processes; and lack of clarity on the meaning of public interest.			
Proposal for a Regulation on European data governance (Data Governance Act) 2020[140]	The proposal aims to **produce the draft of the Data Governance Act**[141] **, which aims to foster the availability of data (both personal and non-personal) for reuse and facilitate trust amongst the private sector through trusted data intermediaries by introducing a horizontal governance framework.** For this, the proposed act	The assertion behind the proposed regulation's objectives stems from the objectives highlighted in the National data strategy. The objectives rely on the broader assertion that an increase in data reuse and availability will add social and economic value based on the industry report in the EU.[142] **The impact assessment also asserts that the**	**The impact assessment identified three problem brackets i.e., low trust in data sharing, issues around reuse of public sector data and collecting data for the common good, technical obstacles to data use.** Within these problem brackets, it was specifically identified that the current practices are leading to consolidation of dominant market actors'	The proposed regulation flows from the **National Data strategy and the aims to European Digital Singles Market**. Moreover, it takes into account the existing policies such as the open data directive, GDPR and existing sectoral level frameworks and builds	The factsheet for the proposed regulation predicts the economic value of data for the market. It states that the annual economic value of data sharing will increase up to €7-11 billion by 2028. It also noted that there will be an increase of 1.3 trillion increase in productivity through IoT data by 2027 and	Consultation on the current proposal was initiated after the adoption of the National Data Strategy in February 2020. The consultation involved 219 business organisation and companies, of which 43.4 % were SMEs. Along with this, 10 workshops with experts from different

136 https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information
140 https://ec.europa.eu/digital-single-market/en/news/proposal-regulation-european-data-governance-data-governance-act
141 https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=71222
142 https://www.vodafone.com/content/dam/vodcom/files/public-policy/Realising_the_potential_of_IoT_data_report_for_Vodafone.pdf

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
	prescribes re-use of public sector data which was not permitted earlier, with technical assurance capability to ensure data protection, privacy and confidentiality, data altruism, certification and labelling framework for data intermediaries to facilitate trust.	**exchange of data amongst the private sector from diverse sources to benefit their value change may be greater than monetary incentives for data sharing. Furthermore, the proposed regulation also aims for facilitating regional and local level data sharing amongst member states and businesses** so that they can formulate holistic consumer insights and have an alternate business model from the big-tech. Moreover, the assertion brings in the view that since the public sector data has been generated at the expenses of the public and must be fully utilised for the benefit of society. The regulation also formulates its basis for facilitating the single market for data ad increasing flexibility, which can only emerge through harmonising and plugging gaps in data sharing practices amongst stakeholders and member states.	power due as data intermediation is being provided by dominant forms are non-EU, lack of competing offers leading to increase in transaction costs **which may act as a burden for the SMEs. Additionally, it emphasises that limited availability of data will affect Ai innovation and stagnation of data professional and companies. Moreover, the problem statement, specifically recognises that internal markets are not fully developed enough to achieve economies of scale which results in dependencies on third countries.**	from it and ensures harmonisation with them. **Furthermore, in setting up of governance body, also relies on a GDPR kind of data protection board**.	will also add to societal benefits, for instance, in the saving in the health sector amounting to 120 billion euros.	sectors were conducted to gather issues and problems before setting up a horizontal data governance framework.

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
Netherlands						
Dutch Vision on Data Sharing Between Businesses 2019[143] (Under Dutch Digitalisation Strategy)[144]	The strategy aims to facilitate data sharing between businesses where the market failed to do so and is unable to organise data sharing or their abuse of dominance by few players in the market. The strategy 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 ensure 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.	The strategy recognises that data is a resource for the 21st century and its re-use and sharing will benefit the business. As a part of the bigger Dutch Digitalisation Strategy to get the Netherlands ready for the digital future, the report deals with the data sharing processes and frameworks between businesses. The strategy covers data sharing for innovation and increasing competition. It also recognises that compulsory 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.	Based on the consultations reasons as to why and in what capacity government - intervention would be required- unwilling to share (competitive advantage), being unable to share due to lack of technical standards, not being permitted to share due to legal obligations.	The strategy recognises the data market growth and the important developments in data regime by the EU like GDPR, initiatives like iSHARE, and non-profit tools like My Data Done Right. The strategy takes these developments into account before building up a progressive data-sharing regime on these developments, ensuring harmony and market inclusion with existing policies.	Dutch businesses and public authorities have invested a significant amount in ICT capital (EUR 26 billion out of a total of EUR136 billion invested by the Netherlands in 2015) with the view that data will be an important resource. The strategy realises its role in encouraging data sharing in the growing market and ensuring that the market organises productive data sharing itself, or if dominant providers or platforms that have unique access to specific data are unwilling to share it with other parties. The strategy also talks about its role in limiting risk in areas such as privacy and cybersecurity	The strategy discusses several case studies on different approaches towards data sharing. Based on these approaches, the strategy fills the gap between them. The strategy is somewhat flexible as it hasn't been converted into law or regulation yet.

143 https://www.government.nl/documents/reports/2019/02/01/dutch-vision-on-data-sharing-between-businesses
144 https://www.government.nl/documents/reports/2018/06/01/dutch-digitalisation-strategy

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
Singapore						
Trusted Data Sharing Framework[145]	The Framework is **aimed to provide guidance on key considerations to enable data sharing** (such as valuation mechanism for data, structuring legal relationships to enable data sharing, technical amongst private sector to address trust security concerns considerations, consent mechanism and other operational considerations). **This Framework is just a guide for industry and not for compliance**	The framework states that **data forms a valuable asset for developing Artificial Intelligence (AI) and companies and will add 5-6 % to their outputs**. The objective of the framework **also supports Singapore's Digital Economy Framework for Action, which emphasises developing infrastructures to build capacities** specifically in the Infocom and Media for tapping in data flows and AI.[146] Additionally, **the framework recognises that it is difficult for the business to value their data, thus along with this framework, a guide on the valuation of data was also released**.[147]	The framework notes that based on industry feedback, the data-sharing ecosystem is still in a nascent stage and guidance is still very much required to help organisations, including professional data service providers, overcome the concerns of data sharing.	**The framework comes when Singapore already has a robust data law like the Personal Data Protection Act 2012.** In fact, consultations were undertaken on improving the framework to balance data sharing and data protection concerns. At the same time, Data Protection Trustmark certification scheme was also introduced. **This framework also came as the backdrop of Singapore's intent to participate in the APEC Cross-Border Privacy Rules system and APEC Privacy Recognition for Processors System**.[148] Moreover, along with the current framework,	**It draws from the use cases of data sharing such as bilateral information sharing between banks and telcos to increase customer service and experience; information sharing between Credit Information Bureau in Singapore and banks**, which gives better information on credit risks, data sharing in the real estate sector between property service provider and real estate companies.	The framework is reliant on external studies and cites the Personal Data Protection Act 2012 for a majority of its approaches. In 2020, the PDPA Act was amended to ensure ease in data sharing in cases of contractual necessity and legitimate interests.[150]

145 https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Trusted-Data-Sharing-Framework.pdf
146 https://www.imda.gov.sg/-/media/Imda/Files/SG-Digital/SGD-Framework-For-Action.pdf
147 https://www.imda.gov.sg/-/media/Imda/Files/Programme/Data-Collaborative-Programme/Guide-to-Data-Valuation-for-Data-Sharing.pdf?la=en
148 https://www.nas.gov.sg/archivesonline/data/pdfdoc/20170727002/Developing%20a%20Trusted%20Data%20Ecosystem%20to%20Support%20Singapores%20Digital%20Economy.pdf
150 https://fpf.org/2020/11/18/singapores-personal-data-protection-act-shifts-away-from-a-consent-centric-framework/

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
				a guide to data valuation, proposing a different approach to a valuation was also introduced. It also fills in for IP and Copyright laws which do not address this. [149] Thus, the framework is merely providing guidance to data partnerships in an already existing policy regime		
Australia						
Data Sharing and Release Legislative Reforms, 2019[151]	**It introduces the standards for legislation that will empower government agencies to safely share public sector data with trusted users for specified purposes**. It aims for streamlining and modernising data sharing, overcoming complex legislative barriers and outdated secrecy provisions. It forms the basis of the new	**The Australian government released this framework based on the Productivity Commission's (PC) report which aimed for increasing public sector data availability for better delivery of digital services.** The main objective is to make public sector data available for the delivery of government services, improving public administration and research helping to create better public policy.	The Productivity Commission's report identified the areas and organisation in the public sector where most data is retained and made a comparative assessment with other countries regarding the availability of open data. **In this assessment, it was found that Australia was lagging in data availability resulting from inconsistent practices and no single approach**	Before the current framework, **Australia had an 'Open Government National Action Plan' which it plans to complement to increase access to public sector data.**[153] Australia also previously introduced Consumer Data Right to facilitate data flow from the private sector. **The Consumer Data Right (CDR) and the Data Sharing and**	The maturity of the **market for data access was indicated through the responses from the consultation in which research institutions and civil society indicated that access to the public sector data is necessary** and could be an important driver of innovation at this stage.	The reforms have been derived from the Productivity Commission's findings on the use and recommended data reforms to unlock the full potential of public sector data. In 2018, the Office of the National Data Commissioner was established within the Department of the Prime Minister and

149 https://www.imda.gov.sg/-/media/Imda/Files/About/Media-Releases/2019/Factsheet-on-Trusted-Data-Sharing-Framework.pdf?la=en

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

153 https://www.pc.gov.au/inquiries/completed/data-access/data-availability-use-government-response.pdf

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
	regulation to be introduced for purposes of sharing such data.	This reform framework also asserts that access to data **will increase citizens' access to information through one contact point; will reduce time and access, and will help in assessing sectors-wise funding priorities** Additionally, to inform guidelines on the data release it also identified case studies where data sharing could be implemented and related challenges and applicable laws and frameworks.[152]	**to public sector data sharing.** PC **also noted that their exits lack of trust by both data custodians and users in existing data access processes and protections creating numerous hurdles to sharing and releasing data are choking the use** and value of Australia's data.	**Release legislation are both part of Australia's efforts to reform data legislation.** While the CDR relates to private sector data, the Data Sharing and Release legislation is focused on government-held data. The current framework has taken into consideration the finding of the **privacy impact assessment conducted in 2019.**		Cabinet to oversee the reforms to improve data sharing and use across the Australian public sector.
Data Exchange Framework IT Strategy Action Plan 2017-18[154]	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. **The framework primarily focuses on sharing**	The framework came about as support Victorian Centre for Data Insight's (VCDI). Data Reform Strategy, API (application programming interface) gateway, since the existing policies did not account for all data classifications.	The Data Exchange framework is a part of the larger Information Technology Strategy 2017-18 Action Plan. The **action plan states that the government systems are an amalgamation of old and new, thus; a uniform infrastructure is needed so that the government**	Before the introduction of the Data Exchange Framework, the Victorian government had created 'Victorian Data Insights' which gave clear policy design, protection and safeguards establishing data analytics technology, along with also creating an	The framework **will also be complemented by the 'Service Victoria initiative' also identified by the IT strategy identifies that in 2017-18, which will allow Victorians to undertake** an initial range of transactions with the government,	This framework forms the part of the Action of the IT strategy 2017-18. The strategy came about through various studies and consultations.

152 https://www.pc.gov.au/inquiries/completed/data-access/report/data-access.pdf

154 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

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
	between government and third parties and between government		**can deliver services more efficiently.** [155]	Information management system.[156]	all from the one internet portal, with the range of transactions increasing over time. This again facilitates the building of the Data Exchange Framework.[157]	
Japan						
Contract Guidance on Utilization of AI and Data by Ministry of Economy Trade and Industry 2018[158]	These guidelines aim to give standards and details that should be included in be **formulating terms while contracting for data sharing**.	The guidance is based on the assertion that 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 utilization of data or AI technology **due to a lack of sufficient experience in contract practices and the gaps in understanding between the parties involved.**	**A survey conducted by a think tank on this issue of data sharing revealed that 15.2 % of companies out of the 562 responses believed that data utilization contracts helped them achieve efficiency and reduced costs**, however, the contracts signed in for this played a significant role in a smooth transaction. **At the same time, 15% of cases also highlighted problems related to leakage of trade secrets**	The guidance intends to aid the overall strategy of the Japanese government intended to promote data sharing and innovation. Based on this, **the government had formulated a study group focusing on the fourth industrial revolution and intellectual property systems. The report of the study group has informed the provisions of this guidance.**	**It must also be noted that this guidance has also been based on the finding of the IoT Acceleration consortium which analysed and provided guidance on used cases of data utilization for IoTs.**[161] Thus, it indicated the data usage practices followed by the stakeholders.	The first version of the contract guidance was created **in 2017; however, after receiving comments from industry and associations, a revised version was published in 2018**, on the recommendation of the committee under Professor Toshiya Watanabe.

155 https://www.vic.gov.au/sites/default/files/2020-08/Information-Technology-Strategy-Action-Plan-2017-18.PDF
156 https://www.vic.gov.au/information-technology-strategy-2017-18-action-plan
157 https://www.vic.gov.au/information-technology-strategy-2017-18-action-plan
158 https://www.meti.go.jp/english/press/2018/0615_002.html, https://www.meti.go.jp/press/2019/04/20190404001/20190404001-1.pdf.
161 https://www.meti.go.jp/english/policy/mono_info_service/information_economy/index.html#two

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
			and unauthorised use of data utilization.[159]	**Previous policies related to this guidance involved the establishment of the Personal Information Protection Commission.**[160]		
Act on Special Measures for Productivity Improvement, 2018[162]	The **act is aimed at attracting investment and facing international competition through increase productivity in the IoT, big data** and artificial intelligence. Notably, the provision under this act are subject to the Basic Act on the Advancement of Public and Private Sector Data Utilization[163] and Act on the Protection of Personal Information	This act had been enacted at the backdrop of **Japan's 'Economic Policy Package of 2017', to address the disruptive changes and international competition which is generated by emerging technologies. The act will help gain investment and increase market productivity**. In an OECD assessment, on markets in Japan, it was noted that the government will enhance the development of **smart infrastructures trying to achieve the priority vision of the government for data-driven society 5.0.** This act was enacted as one	The **Economic Policy Package of 2017 recognized that Japan had been facing sluggish growth in sectors such as automated driving, health sector, agriculture, construction and financial sector.** For this, within the policy, Japan has laid down its target for achieving goals for Society 5.0, with investment in IoT and AI to facilitate growth in the above sectors.[164] For achieving these targets, the proposed act aims for businesses to increase and incentivising data sharing.	As a support to the New Economic Policy Package 2017, the act ensures that the regulations do not lag behind the planned objectives of the policy package.	The act establishes a "regulatory sandbox" and provides tax breaks to encourage IoT investment for facilitating data sharing, and to encourage SMEs to invest in business facilities, in turn driving the market growth.	The act was drafted based on consultations under the cabinet office and was introduced in June 2018.

159 http://www.hitachi.com/rev/archive/2019/r2019_03/pdf/P057-063_R3a03.pdf
160 https://www.ppc.go.jp/en/
162 https://www.meti.go.jp/english/press/2018_06/0606_001_00.html
163 http://www.japaneselawtranslation.go.jp/law/detail/?id=2975&vm=02&re=
164 https://www5.cao.go.jp/keizai1/package/20171208_package_en.pdf

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
		such initiative towards building a data-driven economy.				
United Kingdom (UK)						
National Data Strategy 2020 (Under Consultation)[165]	This strategy is an all-encompassing data strategy that aims to leverage the existing strengths of the UK 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 choice for consumers and small businesses.	This strategy comes at the backdrop of used cases of data sharing by private companies and various sectors, inspiring the parameters and focus of this strategy. The strategy broadly identifies areas for data sharing and include the usage of data for the public sector and for a fairer society, however it does not specifically rely on 'public interest purposes' and recognises the necessity of an appropriate legal framework for the usage of data for the public interest. At the time, it relies on the Report of the National Infrastructure Commission – Data for Public Good to developed policies to unlock the value such that balance could be created. For this, UK has come with	Before the introduction of the strategy, consultations were conducted which recognised the following barriers **for the data economy - lack of governance and senior leadership buy-in on data issues; a lack of agreed standards and poor data quality impacting the effective use and interoperability of data; a data skills gap, both specialist and baseline** in the workforce; a fear of privacy issues and negative thinking leading to datasets being 'closed by default'; lack of clarity on data access rights; legacy infrastructure and software; a culture of 'working in silos' with data across both the public and private sectors.[168]	The strategy acts as an enabler to the **UK AI Sector Deal discussed next and builds on the existing frameworks such as the Research Powers of the Digital Economy Act (2017), which has already enabled data usage in various sectors for research purposes.** Moreover, the strategy relies on the findings of the Competition Market Authority (**CMA)'s report on online platforms and digital advertising, which highlights that the limited access to data by some companies comparing to tech giants may limit innovations**.	The UK data market is the largest in Europe, with high investments in tech, the strategy, therefore, seeks to capitalise and prepare for the futures markets. The **strategy notes that the data economy grew about twice as quickly as the rest of the economy during 2010, making up about 4% of UK GDP in 2020.** Further, as per the estimates noted by the **report in 2018, the UK exported £190 billion in digitally delivered services (67% of total UK services exports) and**	Before the strategy was introduced evidence collection and roundtable consultations were undertaken by the Department of Digital Culture, Media and Sports. The summary findings are available in the public domain. The questions for call for evidence were on the themes of people, government and economy. 105 responses were received from various organisation and sectors including ICT, education organisations, businesses, the public sector, etc. Through this exercise existing

165 https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy

168 https://www.gov.uk/government/publications/uk-national-data-strategy/call-for-evidence-and-roundtable-engagement-summaries#national-data-strategy-roundtable-engagement-summary

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
		Gemini Principles'[166] for sharing data for the public good. The strategy also notes from its consultations that - **views vary markedly on a range of broader questions based on different perspectives** – for example, on legislation, burdens on business, burdens on the public, privacy and trust, the roles of players in the ecosystem for setting standards etc., **and there is unlikely to be 'right' answer on these questions, or their solutions**, The **trade-offs between objectives will be inevitable** – for example, the benefits of making data open and the costs of maintaining it, **there is a wide disparity in engagement with issues around data across the economy and society**.[167]	It is also intending to fill in the gaps left by the Data Protection Act 2018 which focused on personal data similar to the EU GDPR.[169]		**imported £90 billion digitally delivered services (52% of UK services imports**). This estimate indicates the maturity and the predicted growth of the data economy of the UK The strategy aims for businesses to embrace technology, leading to job creation by data use, opening up whole new markets and drives demand for a highly-skilled workforce.	good practices of data access and sharing were also identified. [170] Currently, the strategy is open for public consultation and is proposed to get the views of the stakeholder on the kind of government intervention that might be apt.

166 https://www.cdbb.cam.ac.uk/system/files/documents/TheGeminiPrinciples.pdf

167 https://www.gov.uk/government/publications/uk-national-data-strategy/call-for-evidence-and-roundtable-engagement-summaries#national-data-strategy-roundtable-engagement-summary

169 https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted

170 https://www.gov.uk/government/publications/uk-national-data-strategy/call-for-evidence-and-roundtable-engagement-summaries#national-data-strategy-call-for-evidence-summary

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
		The strategy also identifies that opening up of every dataset may not be the solution. In this context, it states that it will also be important to consider various costs and to ensure that data access is wide enough to reach all the sectors to maximise its value adequately.				
UK AI Sector Deal 2018-19 (Data Sharing Infrastructure)[171]	This 'Sector Deal' sets out actions to promote the adoption and use of AI in the UK and **deliver on 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**. In this regard, the deal also forms interlinkages with the goals of the UK's industrial and digital strategy within the data economy	The UK is home to some of the biggest names in AI innovation and training AIs need a vast amount of data, skilled employees, and innovation enablers like testing availability. The deal is intended to benefit the economy and society by attracting investment, creating jobs, and reaping AI's benefits, all at once. The independent **AI review also accounts for the findings of Royal Society's Machine Learning which indicated that large datasets would be needed**	The sector deal sets out actions to promote the adoption and use of AI in the UK and delivers on the recommendations of the independent AI review.[172] The review highlighted that there is a lack of know-how to proceed in formulating agreements and establish trust between parties and manage the data-sharing practice. **Furthermore, apart from building trusts, it points at times procedural and legal costs of data access, which may hinder such**	The AI Sector Deal proposed a data trust model for un-tapping the datasets from both public and private sectors. In terms of existing policies and practices regarding access to data, UK was ranked first in the world on Government performance on open data. The independent AI review which led to this strategy was also contextualised with parallel reports on data management	The studies conducted and the report on which the deal is based has talked about the approaching maturity of the UK data market. High investment with 33% of European investments captured by the UK, is to be combined with progressive and supportive policies for market growth. The review had estimated that AI could add an additional USD $814 billion (£630bn)	Studies and several projects were taken into account in the preparation of this Deal. Based on the multi-fold recommendations of the 2017 independent report on 'Growing the artificial intelligence industry in the UK', stakeholders from academia, market players, and authorities were contributory to the deal. It was presented in 2018 and finalised in 2019.

171 https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal
172 https://www.gov.uk/government/publications/growing-the-artificial-intelligence-industry-in-the-uk

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
		to improve the outcomes of AI. Furthermore, the review notes that some sectors have generated such amounts of data that it could only be processed by AI. For achieving these goals, the deal focuses on enhancing the UK's existing data infrastructures considering the use cases of open data and data sharing such as CityMapper, focus on sharing geospatial data.	**access for smaller companie**s.	addressing ethical (trust and accountability) and governance questions specifically concepts of data stewards and data trusts.[173] Additionally, the review notes that there are existing best practices and data sharing frameworks amongst individual companies which can inform in building trusted data sharing frameworks.	to the UK economy by 2035, increasing the annual growth rate from 2.5% to 3.9%. Additionally, the market maturity of the AI sector was also contextualised through the Industrial Digitalisation Review, which covered the benefits of deploying robotics and AI to improve industrial processes indicating the potential of data infrastructures to add value.	
Sectoral Frameworks/ Initiatives/ Strategies of Data Sharing						
European Union						
Commission Delegated Regulation (EU) No 886/2013 for data and procedures for the provision, where possible, of road safety-	This directive **aims for the traffic data to be made easily available for exchange and reuse for the provision of information services, public and/or private road operators and service providers.**	The development of the directive was based on the assertion that **citizens must be properly informed about traffic incidents and situations.** **Data must be made available via the same format to achieve**	This directive is based on the **Commission's report which followed the 2010's directive to undertake an impact assessment**.[176] The report identified that there was a lack of data protection; **lack of interoperability;**	This **regulation is supplementing the Directive 2010/40/EU** with regard to data and procedures for the provision, where possible, of road safety-related minimum universal traffic	Within the consultation in ex-post evaluation, the stakeholders indicated that the **2010 Directive has set some very relevant objectives for the market stakeholders, however more**	The initiative for traffic data sharing in the EU was initiative with Its directive in 2010 aiming to create National Access Points (NAP) [179] to create a Single European Transport

[173] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf

[176] https://ec.europa.eu/transport/sites/transport/files/legislation/swd20190368-its-ex-post-evaluation.pdf

[179] National Access Points can take various forms, such as a database, data warehouse, data marketplace, repository, and register, web portal or similar depending on the type of data concerned and provide discovery services, making it easier to fuse, crunch or analyse the requested data sets

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
related minimum universal traffic information free of charge to users[174]	The directive thereby establishes the specifications necessary to ensure compatibility, interoperability and continuity for the deployment and operational use of data and procedures for the provision, where possible, of road safety-related minimum universal traffic information free of charge to users on EU level.	**compatibility, interoperability and continuity.** These objectives have facilitated the goals mentioned in the communication towards a '**European road safety area: policy orientations on road safety 2011-2020'. It identifies that 'Intelligent Transport Systems (ITS Communication**)' have many roles in improving traffic safety. Thus, in order to develop the ITS systems, it was necessary to processing traffic-related personal data (which would be anonymised) The report also asserts that developing interoperable information systems would rely on existing technical solutions and open standards which have been provided by the EU and other international bodies.	**fragmented development of the technological landscape at different levels.** Notably, there were also problems of lock-in effects, hindering competition and limiting opportunities for innovation. The targeted purpose of such data sharing would give real-time access to the public regarding road safety.	information free of charge to users. After the 2010 directive, the EU also formulated specific technical standard **to be followed the DATEX II (CEN/TS 16157) format, to ensure compatibility, interoperability and continuity for the deployment and operational use of Intelligent Transport Systems (ITS) for** the provision of EU-wide safety-related traffic information (SRTI) and real-time traffic information (RTTI) services.[177] While this is arguably a barrier to innovation (since all data must fit into the existing standard), it does support data sharing. In addition, the risk is minimised since	**technical guidance for data sharing interoperability was require**. [178]	Area. Thus, starting from the 2010 directive delegated legislation was introduced to facilitate data sharing through these access points. As a result the 2013 directive to lay down procedures and provisions for making data available free of charge to the users. This was further followed by directives in 2015[180] and 2017.[181]

174 https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013R0886&from=EN
177 https://www.datex2.eu/datex2/specifications
178 https://ec.europa.eu/transport/sites/transport/files/legislation/swd20190368-its-ex-post-evaluation.pdf
180 Adds provision of EU-wide real-time traffic information services
181 Adds provision of EU-wide Multimodal Travel Information Services.

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
		An EU harmonised profile has been created for the purpose of making road safety-related traffic data available in the same format.[175] This regulation is based on an accompanying report of the 2010 directive contains an analysis of the market and regulatory requirements, which were evaluated in the context of the growing data-sharing regime and in light of other upcoming data legislations like the GDPR, which came later.		the referenced standard is an EU harmonised profile. This can be considered a best practice to facilitate data sharing. The current directive was also followed up by the 2015 directive, which further aimed to assist in improving data sharing efficiencies		
EU Code of conduct on agricultural data sharing by contractual agreement 2018[182]	**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	This code of conduct is a contractual agreement between a coalition of associations from the EU agri-food chain that intends to promote the advantages of sharing agricultural data and enabling agri-business models, including agri-cooperatives and other agri-businesses, to swiftly move into an era of digitally enhanced farming.	The EU aims to promote **data sharing in many contexts and since legislative frameworks are not always capable of providing sufficiently detailed yet flexible, appropriate and commercially reasonable mechanisms of data sharing**. The Code **comes as an occupier of the vacant regulatory space after the**	After adopting the regulation on free–flow of non-personal data, which also referred to the free flow of agricultural data, the agricultural sector in the EU took the initiative to formulate a code of practice. The code has inspired the code of data use in arable farming in the Netherlands. However,	The code was agreed upon by the association and agricultural co-operatives, which considered it vital to share agricultural data for a move towards digital farming. They highlighted that big data, precision farming, drones, robots are part of the	The Code of Conduct was designed with the assistance and data provided by the European Council. The Code relied on and presented multiple case studies for different kinds of data sharing. The code was drafted by and in consultation with representative associations which

175 https://ec.europa.eu/transport/themes/its/road/action_plan/traffic-information_en
182 https://www.ecpa.eu/sites/default/files/documents/AgriDataSharingCoC_2018.pdf

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
	to be addressed appropriately. To this end, the framework gives guidelines on what components are to be considered while formulating data-sharing contracts.	**Compliance with the Code is voluntary and acts as a trust-building mechanism for the stakeholders who are party to the code or which to adopt the code in the future**.[183][184] The agricultural **cooperatives took the initiative to formulate this code for making the sector more** competitive and sustainable.	**introduction of GDPR and the non-personal data frameworks in specific cases.** The code aims to address the following issues: attribution of the underlying rights to derive data (data ownership); data access, control and portability; data protection and transparency; privacy and security of data; and liability and intellectual property rights.	this forms the first EU-wide initiative.	farming vocabulary and are a reality in many farms and cooperatives across Europe, promising to revamp the sector.[185] Digital technologies have given new momentum to the traceability of food and consumer information; however, the current landscape is fragmented.[186] At the same time, they also recognised problems with broadband in remote areas and expressed the need to develop those infrastructures.[187]	represented different consumer and stakeholder groups working in the agro-goods industry. The timely response from the agri-sector led to the official launch of the EU Code on 23 April 2018, which was signed by eleven major organizations representing EU agri-businesses.
Payment Services Directive 2015[188]	The **directive stipulates rules for sharing customer's payment data across service providers**.	To ensure that consumers, merchants and companies enjoy the full benefits of the European Single Market, there is a need to facilitate secure, efficient, competitive	The PSD2 replaces an older directive from 2007. The PSD1, while being appropriate at the time, could not foresee the future growth, the needs, and the	The PSD2 comes after well-established frameworks and agencies which were initiated by its predecessor. **The PSD**	The already existing PSD 1 Directive led to the notable development of new Payment Institutions.	Payments and Insurance (EGBPI), the Payment Systems Market Expert Group (PSMEG), and the Payments Committee.

183 https://copa-cogeca.eu/img/user/files/EU%20CODE/EU_Code_2018_web_version.pdf
184 https://link.springer.com/article/10.1007/s10676-020-09543-1
185 https://copa-cogeca.eu/Download.ashx?ID=1388629&fmt=pdf
186 https://copa-cogeca.eu/Download.ashx?ID=1748305&fmt=pdf
187 https://copa-cogeca.eu/Download.ashx?ID=1838642&fmt=pdf
188 https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015L2366&from=EN

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
	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.	and innovative electronic payments. **The PSD2 seeks to increase pan-European competition and participation in the payments industry also from non-banks and to provide for a level playing field by harmonizing consumer protection** and the rights and obligations for payment providers and users. For this, the directive asserts that sharing of data by banking institutions, with the consent of consumers through APIs. This is also based on the development in the **'Fintech market' wherein new players with emerging technologies such as AI or blockchain, which are not currently covered in any regulatory framework.** This also hinders innovation, as they are not able to get access at par with banking institutions.	digital transformation in the payments market after the widespread mobile adoption of the internet. PSD2 has tried to update the directive by taking into account the evolving payments market. This involves setting guidelines monitoring frameworks, and reporting mechanisms for the ease of consumers while providing guidelines for new players entering the market. The impact **evaluation by the European Commission revealed that technological developments have given rise to significant challenges in the payments ecosystem from a regulatory perspective, with many innovative payment products or services do not fall, entirely or in large part, within the scope of the previous directive**.	**1, therefore, already led to an overall integration of the retail payments markets leading to a single market for the payments system**. It has already led to the creation of a common framework for the conduct of business rules of payments services. It also provided a lighted regime to payment institutions which lead them to innovate better. The PSD2, thus, built upon this framework and adjusted them to contemporary needs, while also creating new regulations and mechanisms for data sharing. The directive also relied on other data sharing frameworks that had been initiated or planned by the European Commission. The PSD2 also added aspects which the first	It also made it easier for banks to have established in other markets with high entry costs. This further led to innovation in the financial sector and the emergence of non-bank institutions whose business models are consumer-oriented and based on new technologies such as blockchain.[190] Since the implementation of the PSD2 new payments firm has emerged that operate out of and provide pan European services. The wide adoption of open banking because of the support of the European Union, through well-established frameworks, is leading to a new maturity of the 'fintech' market.	These committees encompassed all stakeholders. The directive was introduced in 2016 and came into effect in 2018, providing ample time to market to prepare and adjust as needed.

[190] https://eur-lex.europa.eu/resource.html?uri=cellar:906ed6d3-f509-11e2-a22e-01aa75ed71a1.0001.04/DOC_2&format=PDF

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
		The objective is the creation of a more integrated European payments market, making payments safer and more secure and protecting consumers. Key targets included: replacing the PSD1, providing supplementary information for the card systems, regulating interchange fee, streamlining payments governance, and explaining interdependencies in the market.	The previous directive in some cases was also found to be too ambiguous, too general or simply outdated, taking into account market developments. Therefore, the need for a new updated policy considering the expansion of the Eurozone and the evolution of payment systems. The growing payments market with different rules for each country also made it difficult for firms to expand to another market within the Eurozone.	directive did not have, such as the rights of consumers it has established a mechanism through which they access all their information. This gave a war for the PSD 2 directive to introduce data sharing obligations to which they adapt to considering their ongoing integration and consumer awareness.[189]	The importance of an open business environment (through increased availability of data) has increased and offering regulatory sandboxes has proven to effectively support the development of the PayTech sector.[191]	
Finland						
Act on the Secondary Use of Health and Social Data, Finland 2019[192] (the Act)	The objective of this act is to facilitate effective and safe processing and access to the personal social and health data for steering, supervision, research, statistics and development in the health and social sector. A second objective is to	The new Act codifies the relevant legislation and broadens the possibilities to, under certain conditions, utilize and combine for secondary purposes personal data collected in relation to public or private social and health care operations.	The act is intended to remove the fragmentation of data sharing regulations and rules which are scattered across different regulations and sectors such as the Patient's Rights Act (1992/785), Act on Electronic Processing of Social and Health Care	The act was complemented by the Health Sector Growth Strategy for Research and Innovation Activities Roadmap for 2018-18[194] and Information to Support Well-being and Service Renewal: eHealth and	The eHealth strategy observes that there was a regional development of availability of patient information since the mid-2000s in the public sector, however, it was not as widespread in the	The Act requires compliance with GDPR and was changed significantly during its proposal phase by the parliamentary committees.

189 https://eur-lex.europa.eu/resource.html?uri=cellar:906ed6d3-f509-11e2-a22e-01aa75ed71a1.0001.04/DOC_2&format=PDF
191 https://www.sciencedirect.com/science/article/pii/S0167268120302328
192 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
194 https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/75145/MEE_guidelines_8_2016_Health_sector_growth_strategy_17062016_web.pdf

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
	guarantee an individual's legitimate expectations as well as their rights and freedoms when processing personal data. The act aims to create an IT ecosystem that both the supplier and consumers of data would access based on licensing terms.	The rationale being ensuring full compliance with the applicable data protection legislation while processing sensitive social and health care data for secondary purposes, while also facilitating better cooperation between the public sector and the private sphere.	Customer Data (2007/159), BioBank Act (2012/688) and Medicines Act (1987/395). The Act also aims to address the administrative burden for the secondary users of social and health care data and parallel and slow licence procedures with various authorities.[193]	eSocial Strategy 2020 (eHealth Strategy). Standards for the contents of electronic medical records have been developed since the 1990s, and technical data transfer standards since the 2000s. Active efforts have been made to standardise the content and technology of information management in the social welfare sector since the mid-2000s. Following this, National Kanta Services were formulated consisting of patient data repositories which were also accessible to the citizens.[195] In 2013, The European Health Telematics Association on evaluation of Kanta Services stated that	private sector. [197] Over time, Finland has defined as a priority the development of tools for health professionals, that will enable sharing of distributed patient information securely, leading to innovation in non-profit eHealth and private eHealth providers who work regionally in partnership with the public system.[198]	Since 2011, after a series of policies and public discussions, a national consensus has been reached through multiple strategies and programmes about the importance of knowledge-based decision-making and linking information and knowledge management to digitisation, experimentation, openness and integration of services. A working Committee was set up to Formulate the Act. The Act was proposed to the government in 2017, on which expert hearing and debates were conducted resulted in suggestion for amendments in 2018. After which the

193 https://blogs.dlapiper.com/privacymatters/finland-parliament-approves-new-act-on-the-secondary-use-of-social-and-health-care-personal-data/#:~:text=The%20Finnish%20Parliament%20has%20approved,effective%20within%20the%20following%20weeks.

195 https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/74459/URN_ISBN_978-952-00-3575-4.pdf?sequence=1&isAllowed=y

197 https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/74459/URN_ISBN_978-952-00-3575-4.pdf?sequence=1&isAllowed=y

198 https://www.tandfonline.com/doi/pdf/10.3402/ijch.v63i4.17749

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
				Finland is a model country for eHealth.[196]		new Act was proposed and passed in 2019. The Act, therefore, in a series of regulations to transform Finland into a significant platform-based economy came into force in May 2019, with a steering committee to oversee the beginning of its implementation stage till June 2019.[199]
Netherlands						
iSHARE[200]	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. The iSHARE uniform set of agreements for identification, authentication and	To enable all players in the logistics industry to connect based on mutual trust, irrespective of type, size, modality and jurisdiction, iSHARE provides a uniform set of agreements or scheme that enables organizations to give each other access to their data. Since they all work with the same identification, authentication and authorization methods, they	The iShare's primary role in filling the market gap is that of an intermediary that eliminates the need for costly and time-consuming integrations to share data with both known and unknown partners, while also allowing businesses to have full control of their data at all times.	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, an industry-government partnership model praised by the EU.	By building upon scalable trust principles from the payment and identity industry, it is laying the foundations for a revolution in simple, cost-effective, and safe data sharing, even with previously unknown partners. In effect, taking the logistics sector to the	In January 2017, logistics-related public-sector and private-sector organizations set to work in co-creation working groups to develop a uniform set of agreements – also called a 'scheme' – for identification, authentication and authorization that could be used and

196 https://blogs.dlapiper.com/privacymatters/finland-parliament-approves-new-act-on-the-secondary-use-of-social-and-health-care-personal-data/#:~:text=The%20Finnish%20Parliament%20has%20approved,effective%20within%20the%20following%20weeks.

199 https://www.sitra.fi/en/publications/a-finnish-model-for-the-secure-and-effective-use-of-data/#abstract

200 https://www.ishareworks.org/en/ishare

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
	authorization 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.	do not need to keep making new agreements every time they want to share data.			next level of data exchange maturity.[201]	applied by everyone in the sector.
South Africa						
Biodiversity Information Policy Framework[202]/ SANBI Data Sharing Agreement[203]	The South African National 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.	SANBI was mandated to collect, generate process, 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.	At a more global level, Open Access to Information has also been addressed as a collective of 34 governments including South Africa. The OECD declared their commitment to Openness: balancing the interests of open 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.	South Africa was one of the 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 while simultaneously protecting sensitive data and maintaining	Both the framework and the sharing agreement come at a time when there is a demand for data sharing in the market. The demand is mainly for 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.	The policy took shape by building upon the open government policy adopted by the South African government in the early 2000s. This was followed by several other policies at both national and regional levels based on the demand, which were brought forward. Eventually, the policies evolved and contributed to forming a national framework to share biodiversity based on set standards.

201 https://www.innopay.com/en/publications/how-ishare-revolutionising-data-sharing-logistics-sector

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

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

Data Sharing Frameworks	Rationale and goals	Assertions behind rationale	Targeted policy, market and regulatory gaps	Policy Maturity	Market Mapping/ Industry Demand	Process followed and timelines
				intellectual property rights.		
Ethiopia						
Agronomy and Soil Data Sharing Policy, 2020[204]	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 Agriculture established a national task force to develop a soil and agronomy data-sharing policy for Ethiopia.	At heart, Ethiopia is still an agriculture-based society and economy. But low-fertility soils and uncertain climate conditions have threatened smallholders and poor farmers across the country. This means that Ethiopia is struggling to produce more food for a quickly growing population. And on increasingly degraded lands. Part of the solution lies in determining the appropriate type and amount of fertiliser for a given location.	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 government had been exploring policies to improve the agricultural outputs of the country.[205]	A civil society-led team collected data from published literature, organisations, researchers, and students on "crop response to fertiliser application" to make this information more accessible. The result was a "coalition of the willing" (CoW) created by soil and agronomy experts eager to share their data or support data access.	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.	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 way forward for the CoW based on the evidence presented by the civil society and the CoW.[206] A draft was presented at several CoW meetings, with a finalised policy launched in June 2019.

204 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)

205 51050623-b954-46cf-bea3-aaefece29408 (moa.gov.et)

206 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).