Evaluating global data policies around non-personal data

Asmita Verma[1] and Anjula Gurtoo[2]

1. Introduction

The second decade of the 21st century has seen an unprecedented rise in data. Innovations in digital technology such as the expansion of internet, introduction of smart devices and cloud-based applications has led to the generation and consumption of data at an almost manic pace. This rapid production of large-scale data has created an enormous potential for the use of data for economic as well as social and public good (Ministry of Electronics and Information Technology (MeitY), 2020). The economic value lies primarily in providing insights and analytics that help to identify patterns and trends for better decision making, and operational efficiency and effectiveness that improves outcomes and reduces effort, risk and costs for stakeholders not only across industry but in other private and public spheres as well. The social and public value is derived from the technical prediction and policy creation capacities of data that enhance the performance of public services while simultaneously improving governance via increased accountability and transparency. Since data has the potential to generate so much value, governments around the world are increasingly looking to govern and regulate all aspects of data to ensure its smooth access and sharing.

Effective sharing, however, implies that these regulations and policies take into account the nature of data. It is now well known that defining data is complex as it is not a uniform entity but heterogenous and contextual (Taylor, 2013; Organisation for Economic Cooperation and Development (OECD), 2019). Depending upon the situation, data can have different implications for governments, academia, businesses and other organisations. Easy availability of data can help improve governance capacity and identify emerging governmental and societal needs by, for instance, increasing efficiency and reliability of transportation and improved supply chains and logistics. In academia, access to data can help researchers foster new methods of data driven research and exploration. Data can also lead to the development of new business models that can be used by big companies, small and medium enterprises as well as start-ups.

From a legal perspective, data is often classified into three categories: personal, and non-personal (NPD) and anonymised data (Surblyte, 2016). Personal data has been generally

[1] Post-Doctoral Fellow, Centre for Society and Policy, Indian Institute of Science, Bengaluru.
[2] Professor, Centre for Society and policy, Indian Institute of Science, Bengaluru.

defined as data which relates to a natural or legal person, who is identifiable either directly or indirectly, in terms of any characteristic, trait or attribute or a combination of these, and includes any inference drawn on this basis for the purpose of profiling (European Union (EU), 2016; MeitY, 2019). NPD is generally defined by exclusion i.e., consisting of data that is not personal in nature. This kind of data can arise form two sources. First is data that has always been non-personal in nature, such as weather and transport data, sensor data or data generated from Internet of Things (IoT). The second is anonymised data which can be defined as data that was initially personal, but has later been made non-personal by anonymising it (MeitY, 2020). Consequently, the governing frameworks for all these types have to be developed by considering a number of factors such as security, privacy and ownership.

Much of the data-driven insights and innovation that leads to economic growth and employment comes from data that has been non personal from the beginning, particularly machine generated data (Kerber, 2016). For example, Transport for London (TfL) is a local government body in the United Kingdom (UK) which is responsible for managing transport services across London. It has an open repository of 62 datasets which are comprised of real time data such as live bus arrivals and traffic disruptions as well as fixed data such as station facilities and bus time tables and are used by a number of companies, among other users (Hodge, 2016). It has been estimated that the total gross value added from using TfL data by these companies directly and across the wider economy is between £12- £15 million per annum (Deloitte, 2017). Additionally, the estimated value of time saved by the users of the applications developed by these companies based on TfL data is about £15- £58 million per annum (Hodge, 2016).

Another instance is in the context of smart cities in India where access to data related to fiscal health and waste management can address the problems of solid waste and at the same time generate additional revenues for cities in India (Indian Urban Observatory, 2019). This study was based on datasets related to urban finances and waste management systems of Urban Local Bodies of about 22 cities, which were accessed through the Smart Cities Open Data Portal. The aim was to investigate their waste management efficiency and financial health in order to develop cost effective solutions in the local context. Findings indicate that the city of Indore had achieved 100% efficiency in recycling and processing of solid waste while maintaining surplus municipal revenues over expenditure to the tune of 30%. This was closely followed by Nasik with an efficiency rate of 80% and surplus revenue generation of 26%. Thus, best

practises from cities like Indore and Nasik can help other cities with lower waste management efficiency to improve their processes.

There are numerous other examples of the use of NPD in other sectors such as healthcare, and its benefits for the society at large. This has prompted countries around the world to create regulatory and policy frameworks to tap into this vast economic and social resource. The primary goal is to ensure the existence of a free and fair digital economy, where privacy of individuals is respected, while simultaneously creating opportunities for innovation and data sharing to reap the aforementioned economic and social benefits.

In this context, there are two objectives of this paper: the first is to conduct a review of global rules and policy guidelines around non-personal data; and, second is to evaluate these based on certain criteria that allow the use of this data for economic and social good. There is an urgent need for doing this because governments and businesses around the world are increasingly going digital. Many countries are at different stages of the process of developing and implementing policies and regulations that will foster the development of the most optimal mechanisms for the access and sharing of NPD and subsequently its use for public welfare. Thus, it is imperative that these documents be reviewed and evaluated for their comprehensiveness, appropriateness and potential effectiveness. This will help to identify the gaps and issues, if any, around NPD governance and management, and enable law and policymakers to suggest suitable solutions. Since the discourse around the value of such data has only recently gained traction, most of the policy and regulatory frameworks have been (or are being) developed quite recently, and thus not much academic literature on evaluating these is available either locally, in the Indian context, or globally. This paper attempts to fill this lacuna and contribute to the research on the governance and regulation of NPD around the world.

The remainder of the paper is organised as follows: After discussing the methodology for the review of global laws and policy documents around NPD in section 2, we focus on discussing the conceptual framework from which we derive our criteria for the evaluation of the same in section 3. This is followed by a country/region-wise summary of the aforementioned documents in section 4, while an evaluation of those documents based on the criteria adopted from the conceptual framework is done in section 5. Finally, section 6 concludes our analysis.

2. Methodology

This paper draws from a review of academic and non-academic literature in the context of non-personal data around the world. This was done in two stages. In the first stage, we identified countries and regions that have some form of national/regional data access and sharing frameworks, whether proposed or in different stages of execution. This was done through a preliminary review of academic articles and reports via online databases such as google scholar, JSTOR as well as legislative databanks. Some of the key words used were 'non-personal data regulations', 'NPD access and sharing frameworks', 'data governance' and 'data economics and policy'. Based on this, we shortlisted six countries and one region viz, India, Australia, Japan, Canada, United Kingdom (UK), Estonia and the European Union (EU).

In the second stage, data policy proposals and papers, directives, guidelines, reports, frameworks, legislations and regulations were retrieved from various government sources such as the websites of ministries, departments, and parliaments of the different countries/regions. Consequently, thirteen documents were selected for analysis (See table 1). These were subsequently reviewed and evaluated.

3. Conceptual Framework

The evaluation of the documents was based on a set of criteria that were adopted from the OECD[3] report on 'Enhancing Access to and Sharing of Data-Reconciling Risks and Benefits for Data Re-use across Societies' (OECD, 2019). The report lays out several mechanisms available to policymakers, regulators and business and other organisations for establishing data-governance models that facilitate data access and sharing, particularly in the context of data that can be considered non-personal such as Artificial Intelligence (AI) and IoT. This includes various factors that should be considered while designing data policies or laws, ranging from data typologies, key actors and their roles in the data economy, data access and sharing strategies to the direct and indirect economic and social benefits of data availability and reuse and the major risks and challenges facing governments, businesses and other stakeholders in advancing the same (See figure 1 for a visual representation).

The first factor to be considered is the different types of data available and whether the policy or other regulatory documents take into account the nuances and the overlaps that are often

[3]The OECD is an intergovernmental economic organisation with 38 member countries, founded in 1961 to stimulate economic progress and world trade. https://www.oecd.org/about/document/ratification-oecd-convention.htm

associated with it. Depending upon the context of its use and its implications for different stakeholders, the framework categorises data along three broad dimensions. These include, firstly, the degree of identifiability of data which reflects the risk of possible harm. Accordingly, data is classified into personal and non-personal data, with personal data associated with a greater degree of identifiability and consequently greater risk.

The second is the overlapping domains of data, which reflects the interests of different stakeholders. Data is categorised into three domains viz personal, public and private. Personal domain includes data "relating to an identified or identifiable" individual, public domain covers everything not protected by IPR or other similar rights and thus free to access and reuse and private domain includes all proprietary data. It is important to distinguish between these domains for comprehensive development and application of data policies and regulations across countries. The third is the source of data origin, which reflects the contribution of different entities to data creation. Based on this, data can be provided, observed, inferred or derived.

The second factor is concerned with the main participants of the data economy and whether the documents have clearly defined their roles and responsibilities. These include data holders, users and intermediaries. A data holder is an organisation, institution, enterprise or individual administering or controlling the access and use of a data source. They are by far the most important actors because they agree to share data for the purpose of social and economic welfare. Their prime responsibilities include deciding the purpose for which and the manner by which data is processed. A data user is the entity responsible for acquiring the aforementioned benefits. These may include government organisations, civil society, academia and industry. Their responsibilities range from using the data for the purpose defined by the holder to ensure the safe management and use of datasets. A data intermediary is an entity that brings the data holders and the data users together to facilitate the smooth access and sharing of data. They may also provide added-value services such as data processing, payment and legal services. They are of different types ranging from data repositories and data brokers to data marketplaces such as online platforms and personal data stores.

The third factor is the ways and strategies by which data can be accessed easily and shared smoothly between the actors. There are three most popular ways of accessing data by users which includes, ad-hoc downloads, application programming interfaces (API) and data sandboxes, respectively in decreasing order of the associated risks. With downloads, the data

is stored in a commonly used format and stored online. Consequently, security and privacy risks are higher since the data holder can no longer control its use once the data is downloaded. APIs provide access to data via an internet web interface. This allows data holders to control what data they want to share with any user they want. Thus, they enable interoperability of different actors and their services, while limiting the potential risks. Data sandboxes are the most secure way of accessing data as they typically require that the data analysis takes place at the same physical location as where the data holder stores the data and thus offer the strongest level of control.

As far as sharing strategies are concerned, there are five different mechanisms available to data holders depending upon the degree of openness in ascending order. The first is via data portability which is the most restrictive form of data sharing. It allows the sharing of data only with either the person whose data it is or a third party chosen by that person. Thus, it restricts access of data to those that were involved in the creation and collection of data. This is followed by contractual agreements and data markets. Many organisations prefer to engage in bilateral or multilateral arrangements to sell or buy data. Others might use a variety of models to sell their data on the market, ranging from subscription based to freemium models.

Then comes other less restrictive arrangements such as data sharing partnerships and sharing data for social good benefit. In data partnerships, organisations enter into partnerships with each other to share and mutually enrich their datasets to derive insights which individually wouldn't have been possible. This typically happens when data is too sensitive to share openly or is restricted via privacy, intellectual property or national security concerns. These can also take the form of provision of data in the private sector to support societal objectives such as in the fields of healthcare or policymaking. Finally, the least prohibitive form of data sharing is via open data. Open data are data that are non-discriminatory in access and can be freely used, re-used and redistributed by anyone at anytime and at any place (OECD, 2005; International Open Data Charter, 2015) and is increasingly being recognised as extremely crucial for maximising the value of data.

The fourth factor to be taken into account is whether the document discusses the social and economic benefits accruing to data sharing and re-use. This is arguably the most important factor since the end goals justifying the increased access and sharing of data is rooted either in the public welfare and societal good or better economic outcomes in terms of reduced accounting and opportunity costs, increased productivity and Gross Domestic Product (GDP),

or both. The societal benefits include enabling increased transparency, accountability and empowerment of users. The economic benefits include creating new business opportunities for incumbents and new entrants, and smaller and larger firms alike, promoting cooperation and competition within and among sectors and countries, finding insights and using them to drive innovation and increasing efficiency by harnessing increasing returns to scope via data integration.

The fifth and final factor focuses on recognising and addressing the risks and challenges that policy makers, legislators and other decision makers face in facilitating ease of access and sharing of data. These risks depend on the nature of data and affect the actors, and data access and sharing mechanisms. These have been grouped around three major issues. The first issue deals with balancing the advantages of increased access and sharing of data with the potential risks. This includes taking into account digital security risks, particularly confidentiality breaches, privacy, intellectual property and other rights related issues, adopting a risk management approach and address barriers related to cross-border data access and sharing.

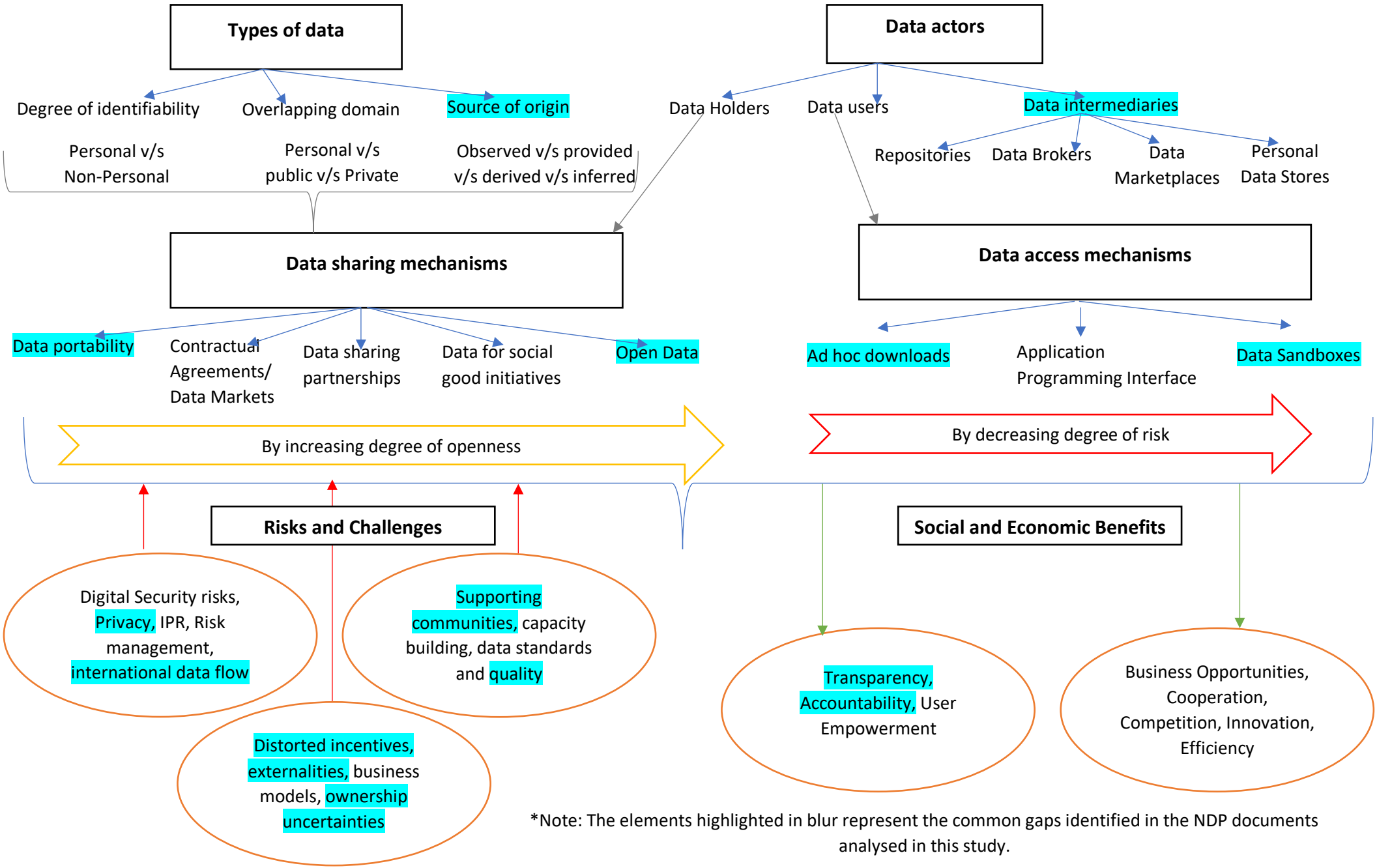
The second issue focuses on building trust and empowering data users and customers for the optimal re-use of data across society. This involves creating, supporting and engaging communities of stakeholders around data access and sharing, and taking advantage of their varied approaches to the same for the better management of risks and incentives structures. It also considers developing and improving data related skills and infrastructure to derive maximum value from the enhanced availability of data for social and economic welfare. Further, it involves setting common data standards that facilitates interoperability and ensuring that quality of data is up to mark.

The third issue involves recognising that effective sharing of data requires methodical incentive systems and sustainable business models. This entails understanding externalities of data sharing and limitations of current business models/markets, analysing the advantages and disadvantages of mandatory access to data which is increasingly becoming a preferable way of easing and enhancing access to data, and clearing uncertainties regarding ownership of data.

4. Overview of rules and policies on non-personal data around the world

In this section, we describe the key highlights of the various policy, regulatory and other documents developed by the requisite countries in the sphere of non-personal data. Table 1 presents the type of document, summarises its broad objective and the type of data it covers.

Figure 1: Data Access and Sharing Framework
(Source: Author's creation adapted from OECD)*



*Note: The elements highlighted in blur represent the common gaps identified in the NDP documents analysed in this study.

India

In December 2020, a committee of experts constituted by the Ministry of Electronics and Information Technology released a governance framework for non-personal data (MeitY,2020). This was a revision of the earlier version which was released in July 2020, after taking into account comments and suggestions from different stakeholders including businesses, research institutions and civil society organisations. The key idea behind the framework is that increased access to non-personal data can be used to generate social, public and economic value, but at the same time can also cause several harms. Thus, the framework proposes, among other things, a new regulation for NPD which is governed by an independent regulatory authority. Key highlights include describing in detail what constitutes NPD, given the difficulty in reaching a consensus globally, the roles of the key actors including data principal, data custodian, data trustee and data business, and establishing community rights over NPD. It also talks about controlled data sharing in the context of high-valued datasets which have several societal benefits such as policy making and innovation.

Australia

The Data Exchange Framework-Streamlined Performance Reporting, 2019

In 2019, the Department of Social Services of the Australian government introduced a data exchange framework (DEX) to streamline reporting of outcomes of program performance in funding agreements (Department of Social Services (DSS), 2019). The objective was to describe clearly the outcomes achieved as a result of the services provided by different organisations to aid their clients. The framework lays down the key governing principles of its design and implementation. It also specifies data requirements including priority data which is mandatorily required to be shared by all organisations and an extended dataset under a partnership process, which organisations may share with funding agencies to receive relevant reports on their performance.

Data Availability and Transparency Bill, 2020

The data availability and transparency bill was introduced in the Australian Parliament in late 2020 (Parliament of Australia, 2020). The objective behind this was to enable government and selected private organisations to share data for the purposes of delivering government services efficiently, improved policy making and program implementation, and facilitating research and innovation. The bill describes the key actors that would benefit from increased data sharing

such as data custodians, accredited users and accredited data service providers. It discusses the appointment of a national data commissioner who will act as an independent regulator and will provide guidance on data related issues. It also clearly describes the purposes for which and the ways in which data can be shared to achieve the aforementioned benefits.

Japan

Japan implemented the basic act on the advancement of public and private sector data utilisation in 2016 (Government of Japan, 2016). The purpose of this was to utilise effectively and appropriately the large amount of data that the internet has allowed to generate to achieve societal outcomes such as safety and development of a comfortable living environment for the citizens. It lays down the basic principles on which the act functions, the responsibilities of the State, local public entities and companies in the optimal usage of data and the development of a basic plan which ensures the smooth circulation of data.

Canada

The government of Canada launched the Canadian Data Governance Standardisation Collaborative to take forward the goals of the Digital Charter Implementation Action proposed by the government in November 2020, of developing and strengthening a national data strategy to enhance the flow of data for deriving economic value (Standards Council of Canada (SCC), 2020). The objective behind the collaborative was to ensure the development of common compatible data governance norms which are essential for the smooth functioning of a data driven economy and become an international influencer in the sphere of data governance. The report highlights thirty-five key issues related to data governance that needs addressing including but not limited to data ownership, data sharing, interoperability, user rights, data quality and ethics, and AI and machine learning.

United Kingdom

National Data Strategy, 2020

The National Data Strategy was launched by the government of United Kingdom to encourage the utilisation of digital services in response to the Covid-19 pandemic in 2020 (Department of Digital Culture, Media and Sports, 2020). The policy paper highlights the main goal of the strategy which is to promote the effective utilisation of data to increase economic growth, create new jobs and drive innovation. Priority areas include harnessing the social and economic

value of data, improving public service delivery and efficiency by optimal use of public sector data, promote capacity building and ensure smooth cross border flow of data.

Data Ethics Framework, 2020

The UK government first published the Data Ethics Framework in 2018 and subsequently updated it September 2020. The prime goal of this document is to provide a set of guidelines for data practitioners, policymakers and anyone working with public sector data to promote its appropriate and responsible usage (Government Digital Service, 2020). It lays out the principles on which the framework is based and steps to achieve the effective use of data at each stage of the user projects, which include defining the public benefit, creating a skilled team, complying with the law, understanding the quality and limitations of data and finally, considering the policy implications of user projects.

Estonia

The Information Systems Data Exchange layer, also known as X-Road, was first implemented by the Estonian Ministry of Economy and Communications in 1998 and was last updated in 2016 (Government of the Republic of Estonia, 2016). The prime objective behind its creation was to enable both public and private organisations to engage in online exchange of data. It provides a platform to use data in a secure and standardised way to produce and consume digital services. It is based on the principles of confidentiality, integrity and interoperability to ensure smooth flow of data.

European Union

The EU has by far the most comprehensive set of policy guidelines, directives, proposals and regulations around the access and sharing of non-personal data.

Regulation (EU) 2018/1807 on a Framework for the Free Flow of Non-Personal Data in the European Union, 2018

The Regulation on a framework for the free flow of non-personal data in the European Union came into effect in May 2019 (European Parliament, 2020). The prime objective of this regulation was to ensure the operationalisation of goods and services related to the non-personal data and its management in the EU and to remove any barriers for the same. The document discusses several issues including the basics of non-personal data and its regulatory

framework, shedding of data localisation requirements, data availability for competent authorities and cooperation among each other, and data portability for professional users.

Regulation on European Data Governance (Data governance Act), 2020

The Data Governance Act is a legislative proposal set forth by the European Commission in November 2020 (European Parliament, 2020). Its main aim is to promote the re-use of public sector data which particularly includes data that may be protected by intellectual property rights and trade secrets, and personal data. This will enhance the availability of data for use, in order to provide solutions in the fields of AI, healthcare, smart cities and more. Key features of the proposal include fostering a business model for data intermediation, promoting data altruism and establishing safeguards for the uncomplicated cross border transfer and access of NPD.

Directive (EU) 2019/1024 on open data and the re-use of public sector information, 2019

The Directive on Open Data and the Re-use of Public Sector Information was adopted by the European Commission in June 2019 (European Parliament, 2019). The idea was to develop a regulatory framework which kept pace with the rapid progress in digital technologies and facilitate the re-use of publicly funded information across the EU. The document discusses the need to make public sector data available to users in the form of open data and dynamic data[4], and the obligation on public authorities to provide public documents as well as high valued datasets free of cost to the users.

European Commission Staff Working Document on Guidance on Sharing Private Sector Data in the European Data Economy, 2018

The European Commission released a Staff Working Document providing guidance on sharing private sector data in April 2018 (European Commission, 2018). The objective was to present a set of guidelines to harness the use of large- scale data that has been generated due to the emergence of the IoT, big data analytical tools and AI applications to facilitate data driven innovation and consequently growth in GDP and employment opportunities. The document discusses a set of principles revolving around data sharing and contractual agreements for Business-to-Government and Business-to-Business transactions. It further discusses the legal and technical aspects of data sharing including data donorships and partnerships as well as data platforms.

[4] Dynamic data is data that is updated in real time and frequently such as weather and traffic data.

Table 1: Summary of global rules and policy guidelines on Non-Personal Data

Country	Regulation/Legislation/ Other Policy Document	Type	Broad Objective	Type of data
India	Report by the Committee of Experts on Non-Personal Data Governance Framework, 2020	Framework	Establishing a data sharing framework that unlocks economic benefits from non-personal data as well as provides certainty of regulations	Non-personal data
Australia	The Data Exchange Framework-Streamlined Performance Reporting, 2019	Framework	Promotes a two-way focus on both smarter and more efficient ways of collecting data from organisations, and more useful reporting about the outcomes achieved for individuals, families and communities.	Public and private sector; demographic
	Data Availability and Transparency Bill, 2020	Bill	Promoting better availability of public sector data and establishing institutional arrangements of sharing public sector data	Public sector data
Japan	Basic Act on the Advancement of Public and Private Sector Data Utilisation, 2016	Regulation	Determining the responsibilities of the State, local public entities, and companies by providing for basic principles with respect to the advancement of the appropriate and effective use of public and private sector data	Public and private sector data
Canada	Canadian Data Governance Standardisation Roadmap, 2018	Report	Coordinating the development and compatibility of data governance standards to harness the economic benefits of data while limiting its potential harms.	Personal and industrial/commercial data
United Kingdom	National Data Strategy, 2020	Policy Paper	Boosting the better use of data across government, businesses, civil society and individuals for the benefit of the society at large.	Personal and non-personal data; public and private sector data
	Data Ethics Framework, 2020	Framework	Guiding appropriate and responsible data use in the public sector	Personal and non-personal data
Estonia	Information Systems Data Exchange Layer, 2016	Regulation	Establishing a technical infrastructure and environment that enables secure exchange of internet data between members	Non-personal data
Europe	Regulation (EU) 2018/1807 on a Framework for the Free Flow of Non-Personal Data in the European Union, 2018	Regulation	Establishing a clear, comprehensive and predictable legal framework for the processing of data other than personal data and ensuring its free flow within EU to develop a strong data-based economy and enhancing the competitiveness of industry	Non-personal data
	Regulation on European Data Governance (Data governance Act), 2020	Proposal	Ensuring the availability of data for use by increasing trust in data intermediaries and by strengthening data-sharing mechanisms across the EU.	Public and private sector data; personal data
	Directive (EU) 2019/1024 on open data and the re-use of public sector information, 2019	Regulation	Promoting the use of open data and stimulating innovation in products and services by establishing a set of rules governing the re-use of public sector and publicly funded information across EU	Open data; public sector data
	European Commission Staff Working Document on Guidance on Sharing Private Sector Data in the European Data Economy, 2018	Report	Laying the foundations for a competitive advantage for European business actors to make the most of data technologies by defining key principles to govern data interactions in business-to-business and business-to-government situations.	Private sector data
	A European Strategy for Data, 2020	Report	Enabling data availability to all to foster an environment of innovation and	Non-personal data

	empower citizens to based on insights provided by non- personal data

A European Strategy for Data, 2020

The European Commission released a strategic report on a European Strategy for Data in February 2020 (European Commission, 2020). The key objective behind the report was the development of a framework that makes it easier for stakeholders to access data for accelerating growth and creating social and economic value. Key highlights include creation of European Data Spaces in public and economic spheres, a single digital market and underlines the importance of open data, data altruism and international data sharing.

5. Evaluation criteria

Based on the OECD report, we evaluate these policies and regulations based on 32 shortlisted criteria in our study. These are organised according to the five factors (which are discussed above in section 3) which should be taken into account while developing data policy and regulatory frameworks.

Types of data and different actors and their roles in the data economy (See table 2)

While most documents do discuss the varied types of data according to their degree of identifiability and overlapping domains, the source of data origin is not adequately discussed. Similarly, the roles of data providers and users have been defined and standard operating procedures have been set, but that of data intermediaries has been missing in half of the documents.

Data access control mechanisms and sharing approaches (See table 3)

We find that majority of documents studied did not adequately discuss the various mechanisms that data users have at their disposal to access data. The use of APIs is acknowledged in about half of the documents but downloads and data sandboxes hardly find a mention.

Majority of the documents do advocate for the establishment of data markets and contractual agreements for sharing data with stakeholders in the wider data economy. They also discuss the benefits of data sharing partnerships as well as making data available for citizen welfare. However, issues around open data and data portability are explored in just half of the documents.

Table 2: Evaluation Criteria: Whether the document discusses different types of data, actors and their roles in the data economy

Country	Regulation/Legislation/ Other Policy Document	Degrees of identifiability of data	Overlapping domains of data	Source of data origin	Data providers (holders and controllers)	Data users	Data intermediaries
India	Report by the Committee of Experts on Non-Personal Data Governance Framework, 2020	✓	✓	✓	✓	✓	✓
Australia	The Data Exchange Framework-Streamlined Performance Reporting, 2019	×	×	✓	✓	✓	×
	Data Availability and Transparency Bill, 2020	✓	✓	×	✓	✓	✓
Japan	Basic Act on the Advancement of Public and Private Sector Data Utilisation, 2016	×	✓	×	×	✓	×
Canada	Canadian Data Governance Standardisation Roadmap, 2018	✓	✓	✓	✓	✓	✓
United Kingdom	National Data Strategy, 2020	✓	✓	✓	✓	✓	×
	Data Ethics Framework, 2020	✓	✓	✓	✓	✓	×
Estonia	Information Systems Data Exchange Layer, 2016	×	×	×	✓	✓	✓
Europe	Regulation (EU) 2018/1807 on a Framework for the Free Flow of Non-Personal Data in the European Union, 2018	✓	×	×	✓	✓	×
	Regulation on European Data Governance (Data governance Act), 2020	✓	✓	×	✓	✓	✓
	Directive (EU) 2019/1024 on open data and the re-use of public sector information, 2019	✓	✓	✓	✓	✓	×
	European Commission Staff Working Document on Guidance on Sharing Private Sector Data in the European Data Economy, 2018	✓	✓	×	✓	✓	✓
	A European Strategy for Data, 2020	✓	✓	×	×	✓	✓

Social and economic benefits of data access and sharing (See table 4)

Most regulatory and policy documents have acknowledged the role of data in enhancing transparency and accountability of public services and the benefits of data access and sharing for the empowerment of users. At the same time, they lack properly defined frameworks that would enforce transparency and accountability. As far as the economic benefits are concerned, all documents talk about the advantages of data availability for data driven innovation and its contribution in increasing efficiency and productivity, but do not clearly discuss ways of crowdsourcing new insights that would bring about those improved economic outcomes. Most documents also discuss the use of data for making existing businesses more profitable as well as the creation of new business opportunities for small and medium enterprises and start-ups. Further, the contribution of increased access and sharing of data to enhance cooperation as well as competition among sectors and countries has also been mentioned in more than half the documents.

Risks and challenges of data access and sharing (See table 5)

Although increased access to data and its sharing is crucial for the economic and social good, it is equally important to weigh the potential harms for development of comprehensive regulations and policies around data. Consequently, the evaluation criteria is based on three themes focused on understanding the risks and challenges of data access and sharing as discussed in the OECD framework.

The first theme focuses on balancing the risks and benefits of greater availability of data. We find that almost all documents in this study do acknowledge the presence of digital security risks and potential confidentiality breaches, but most of them do not lay down explicit rules or guidelines to deal with them. We also find that the policy and lawmakers take personal privacy, intellectual property and other rights into consideration while designing the regulatory and policy frameworks. They recognise that data can be exposed or breached and this can have several direct and indirect costs to the government, businesses and other organisations.

This is reflected in most documents recommending for adoption of a risk management approach which is based on the assumption that some level of risk is always present in accessing and sharing data and aims to identify, forecast and reduce those risks. Finally, given the increased proliferation of data, with most of it being collected, processed and stored outside national boundaries, cross border access and sharing of data has become more important than

Table 3: Evaluation criteria: Whether the document discusses different data access control mechanisms and data sharing approaches

Country	Regulation/Legislation/ Other Policy Document	Ad hoc downloads	Application programming interfaces	Data sandboxes	Contractual agreements and data markets	Open data	Data portability	Data sharing partnerships	Data for social good initiatives
India	Report by the Committee of Experts on Non-Personal Data Governance Framework, 2020	✓	✓	✓	✓	✓	✓	✓	✓
Australia	The Data Exchange Framework-Streamlined Performance Reporting, 2019	×	×	×	×	×	×	✓	✓
	Data Availability and Transparency Bill, 2020	×	×	×	✓	×	×	✓	✓
Japan	Basic Act on the Advancement of Public and Private Sector Data Utilisation, 2016	×	×	×	✓	×	×	✓	✓
Canada	Canadian Data Governance, 2018 Standardisation Roadmap	×	✓	✓	✓	✓	✓	✓	✓
United Kingdom	National Data Strategy, 2020	×	×	×	✓	✓	✓	✓	✓
	Data Ethics Framework, 2020	×	×	×	×	✓	×	✓	✓
Estonia	Information Systems Data Exchange Layer, 2016	×	×	×	✓	×	×	×	×
Europe	Regulation (EU) 2018/1807 on a Framework for the Free Flow of Non-Personal Data in the European Union, 2018	×	✓	×	✓	×	✓	✓	✓
	Regulation on European Data Governance (Data governance Act), 2020	✓	×	×	✓	×	✓	✓	✓

	Directive (EU) 2019/1024 on open data and the re-use of public sector information, 2019	✓	✓	×	✓	✓	×	✓	✓
	European Commission Staff Working Document on Guidance on Sharing Private Sector Data in the European Data Economy, 2018	×	✓	×	✓	✓	✓	✓	✓
	A European Strategy for Data, 2020	×	✓	×	✓	✓	✓	✓	✓

Table 4: Evaluation Criteria: Whether the document discusses the social and economic benefits of data access and sharing

Country	Regulation/Legislation/ Other Policy Document	Transparency	Accountability	Empowerment of users	Business opportunities	Cooperation across sectors and countries	Competition across sectors and countries	Crowdsourcing new insights	User driven innovation	Increased efficiency
India	Report by the Committee of Experts on Non-Personal Data Governance Framework, 2020	✓	✓	✓	✓	×	✓	✓	✓	✓
Australia	The Data Exchange Framework-Streamlined Performance Reporting, 2019	✓	✓	✓	✓	×	×	✓	✓	✓
	Data Availability and Transparency Bill, 2020	✓	✓	×	×	×	×	×	✓	×
Japan	Basic Act on the Advancement of Public and Private Sector Data Utilisation, 2016	×	✓	✓	✓	✓	✓	×	✓	✓
Canada	Canadian Data Governance	✓	✓	✓	✓	✓	✓	✓	✓	✓

	Standardisation Roadmap, 2018										
United Kingdom	National Data Strategy, 2020	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	Data Ethics Framework, 2020	✓	✓	✓	×	×	×	✓	✓	×	
Estonia	Information Systems Data Exchange Layer, 2016	✓	✓	×	×	✓	×	×	✓	×	
Europe	Regulation (EU) 2018/1807 on a Framework for the Free Flow of Non-Personal Data in the European Union, 2018	✓	✓	✓	✓	✓	✓	×	✓	✓	
	Regulation on European Data Governance (Data governance Act), 2020	✓	✓	✓	✓	✓	✓	×	✓	✓	
	Directive (EU) 2019/1024 on open data and the re-use of public sector information, 2019	✓	✓	✓	✓	✓	✓	×	✓	✓	
	European Commission Staff Working Document on Guidance on Sharing Private Sector Data in the European Data Economy, 2018	✓	✓	✓	✓	×	✓	✓	✓	✓	
	A European Strategy for Data, 2020	✓	✓	✓	✓	✓	✓	✓	✓	✓	

ever. However, in about half of the documents, no discussion has been done regarding this aspect.

The second theme discusses ways in which trust can be developed which empowers governments, businesses and other organisations to make effective use of data for economic and societal benefit. Supporting and engaging communities of stakeholders is an important ste towards achieving trust as it can help define responsibilities and acceptable risk levels. Subsequently, we find that all but two countries take this into account in their policy and regulatory documents. However, where different communities are involved, there can be significant challenges in managing competing interests and expectations and we find a gap in policies and regulations to address this issue amongst several countries.

Capacity building is another important factor in enabling users to derive value from data. Most documents have, thus, acknowledged the role of fostering data related skills and infrastructure in maximising public and economic utility from data. In addition, having common data standards to ensure interoperability and maintaining data quality is another challenge to optimal use of data. While most documents do discuss the significance of the former, the latter is not adequately discussed in more than a third of them.

The third and final theme is centred on understanding the distorted incentive structures and shortcomings of current business models and markets that may inhibit data access and sharing. In this respect, the OECD framework recognises that significant costs of investment in collecting, processing, storing and analysing data has positive externalities and may generate a free rider problem, thus disincentivising stakeholders to share in the first place. However, most documents in this study fail to not only acknowledge this but to find useful solutions to this major challenge. Misaligned incentives that hamper voluntary sharing of data often give way to policymakers and regulators suggesting mandatory access and sharing. Although this may be beneficial and even required in some contexts, it may be risky in others. However, this issue hasn't found a mention in many of the documents considered above as well.

Also closely related to the sharing of data is the idea of data ownership. While ownership maybe easy to define and establish in the case of personal data, it is complicated in the case of non-personal data. Uncertainties regarding who owns NPD can hamper it access and sharing. In our study, more than half the documents do flag this as an issue, but no document provides clear and concrete guidelines or regulations to establish ownership. What we find is ownership viewed form the lens of existing IPR like copyrights and trade secrets laws. Lastly, most policy

guidelines and regulations recognise the limitations of existing business models and markets for facilitating data access and sharing and encourage the creation of data markets.

6. Discussion and Conclusion

In this paper, we presented a review of thirteen policy and regulatory frameworks around non-personal data spread across diverse jurisdictions globally and evaluated their features based on a set of 32 criteria which were adapted from the OECD's framework on enhancing data access and sharing. While the policymakers have covered a number of crucial factors, we have identified few areas that need urgent attention.

The first is the role of intermediaries. Evidence suggests that end users, such as consumers and businesses, typically do not have the know-how to use the data available to them and increasingly rely on intermediaries that provide the information in more user-friendly ways (OECD, 2019). Not only this, the intermediaries also provide added value services such as advanced data analytic services. Given the increasing demand for them, it is imperative that the policy and regulatory documents do take into account the different types of intermediaries and define their roles and responsibilities just as they do in the case of data holders and users.

Second is a description of the different ways of data access and sharing mechanisms. Knowledge about data access control mechanisms not only empowers users to access data via different techniques but also presents ways in which data holders, including individuals and organisations, can safeguard their interests and rights. Some frameworks have discussed about APIs as a popular data access mechanism with a comparatively lesser risk factor than downloads, but countries should also explore the concept of data sandboxes in enhancing access to data. This is because they are the safest mode of accessing data, particularly sensitive and proprietary data.

Coming to data sharing mechanisms, we find that while contractual agreements and data markets have been emphasised greatly, not much attention has been paid to open data and data portability. Although open data initiatives have been taken by certain countries such as London, Canada and India, the quality of the data is one of the biggest concerns that plagues these initiatives. Thus, a lot more needs to be done in making data freely available to all. Further, even a discussion about data markets is cursory in most of the frameworks as they do not adequately discuss data monetisation and the appropriate means to achieve it, like establishing buying and selling norms or discussing the pricing and licensing of data. The report on sharing private data by the European Commission in the EU is the only document which considers

Table 5: Evaluation criteria: Whether the document discusses the risks and challenges of data access and sharing

Country	Regulation/Legislation/ Other Policy Document	Digital security risks and confidentiality breaches	Privacy, IPR and other interests	Risk Management	Cross border data access and sharing	Supporting and engaging communities of stakeholders	Capacity building	Common data standards	Data quality	Externalities and misaligned incentives	Limitations of current business models/ markets	Risk of mandatory access to data	Uncertainties about data ownership
India	Report by the Committee of Experts on Non-Personal Data Governance Framework, 2020	×	✓	✓	×	✓	✓	✓	×	✓	✓	✓	✓
Australia	The Data Exchange Framework-Streamlined Performance Reporting, 2019	✓	✓	×	×	×	✓	✓	✓	×	×	✓	×
	Data Availability and Transparency Bill, 2020	✓	✓	✓	×	×	×	×	×	×	×	×	×
Japan	Basic Act on the Advancement of Public and Private Sector Data Utilisation, 2016	✓	✓	×	×	×	✓	✓	×	×	×	×	×
Canada	Canadian Data Governance Standardisation Roadmap, 2018	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	✓
United Kingdom	National Data Strategy, 2020	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	✓
	Data Ethics Framework, 2020	✓	✓	✓	×	✓	✓	×	✓	×	×	×	×

Estonia	Information Systems Data Exchange Layer, 2016	✓	×	×	✓	✓	✓	✓	×	×	×	×	×
Europe	Regulation (EU) 2018/1807 on a Framework for the Free Flow of Non-Personal Data in the European Union, 2018	✓	✓	✓	✓	✓	×	✓	×	×	✓	×	×
	Regulation on European Data Governance (Data governance Act), 2020	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	×	✓
	Directive (EU) 2019/1024 on open data and the re-use of public sector information, 2019	✓	✓	✓	✓	✓	×	✓	✓	×	✓	✓	✓
	European Commission Staff Working Document on Guidance on Sharing Private Sector Data in the European Data Economy, 2018	✓	✓	✓	×	✓	✓	✓	✓	×	✓	×	✓
	A European Strategy for Data, 2020	✓	✓	×	✓	✓	✓	✓	✓	✓	✓	×	✓

some of the revenue models available to stakeholders in the data economy and can help other countries to address this gap in their respective documents (European Commission, 2018)

Third, most of the documents discuss at length the social and economic benefits of enhanced availability and sharing of data. However, even though they mention principles such as transparency and accountability, most documents do not define standards to achieve them and lack structured mechanisms that may deal with potential breaches. Perhaps, the most comprehensive document that discusses accountability mechanisms is the Canadian Data Governance Standardisation Roadmap, 2019 (SCC, 2020). It lays out the obstacles that hamper the development of a rigorous accountability framework while also suggesting ways to address them.

Fourth, the frameworks also fail to establish concrete norms for breach of confidentiality and personal and collective privacy. An individual may not be singled out by aggregating non-personal data, but this data may provide 'invasive insights' on a communities' behaviour (Nanda and Kapoor, 2021). Further, where engaging different stakeholder communities is emphasized, and more so, the ownership of NPD by communities is encouraged, it is crucial to firstly define what a community is and secondly, understand how 'social dynamics and power inequalities' within communities will affect the exercise of these ownership rights (Marda, 2020).

Fifth, it is important to stress the safe and secure flow of cross border data in order to access large-scale high-quality data (which may be present in different jurisdictions) and harness its value in the fields of international trade, e-commerce and cloud computing (World Economic forum, 2020). Thus, regulatory and policy documents should take into consideration measures which encourage international data movement such as data portability, interoperability and integrity.

Finally, effective data access and sharing is hampered by distorted incentives and the free rider problem. Depending on the nature of data and context, different incentive mechanisms may be required to achieve optimal sharing and there may not be one standard solution to this issue (OECD, 2019). Thus, law and policy makers need to take this into consideration when finding ways to align current incentives structures and encourage the use of sustainable business models for the unhindered access of data.

References

Deloitte (2017), *Assessing the Value of TfL's Open Data and Digital Partnerships*, http://content.tfl.gov.uk/deloittereport-tfl-open-data.pdf.

Department of Digital Culture, Media and Sports. (2020). National Data Strategy, 2020. Government of UK. https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy

Department of Social Services. (2019). The Data Exchange Framework. Australian Government. https://dex.dss.gov.au/sites/default/files/documents/2019-09/data-exchangeframework-2019.pdf

European Commission. (2018). COMMISSION STAFF WORKING DOCUMENT Guidance on sharing private sector data in the European data economy. https://digital-strategy.ec.europa.eu/en/news/staff-working-document-guidance-sharing-private-sector-data-european-data-economy

European Commission. (2020). A European strategy for data. Accessible at https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy#documents

European Parliament. (2018). Regulation (EU) 2018/1807 of the European parliament and of the council of 14 November 2018: On a framework for the free flow of nonpersonal data in the European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807&from=EN

European Parliament. (2019). Open data and the re-use of public sector information. https://eur-lex.europa.eu/eli/dir/2019/1024/oj

European Parliament. (2020). Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act), 2020. Accessible at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020PC0767

European Union (2016) General data protection regulation. Off J Eur Union 49:L119 https://gdpr-info.eu/

Government Digital Service. (2020). Data Ethics Framework, 2020. Government of UK. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/923108/Data_Ethics_Framework_2020.pdf

Government of Japan. (2016). Basic Act on the Advancement of Public and Private Sector Data Utilization. https://japan.kantei.go.jp/policy/it/data_basicact/data_basicact.html#:~:text=Article%208%20(1)%20The%20government,of%20measures%20for%20the%20advancement

Government of the Republic of Estonia. (2016). Information Systems Data Exchange Layer. https://www.riigiteataja.ee/akt/127092016004

Hodge, B. (2016). United Kingdom's Transport for London. Get Set, Go! https://odimpact.org/case-united-kingdoms-transport-for-london.html

Indian Urban Observatory. (2019). Leveraging Big Data in Smart Cities. https://iuo.mohua.gov.in/portal/sharing/rest/content/items/fa743e2971154a18a58d1da28990d879/data

International Open Data Charter. (2015). *Principles*, https://opendatacharter.net/principles/

Kerber, W. (2016) A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis. Gewerblicher Rechtsschutz und Urheberrecht, Internationaler Teil (GRUR Int), 11/2016, 989-999, Available at SSRN: https://ssrn.com/abstract=2858171

Marda, V. (2020, October 15). Non-personal data: the case of the Indian Data Protection Bill, definitions and assumptions. https://www.adalovelaceinstitute.org/blog/non-personal-data-indian-data-protection-bill/

MeitY. (2019). The Personal Data Protection Bill, 2019. Government of India. http://164.100.47.4/BillsTexts/LSBillTexts/Asintroduced/373_2019_LS_Eng.pdf

MeitY. (2020). (Revised) Report by the Committee of Experts on Non-Personal Data Governance Framework GOI. https://static.mygov.in/rest/s3fspublic/mygov_159453381955063671.pdf

Nanda, A. & Kapoor, A. (2021). Understanding Non-Personal Data Sharing. A Principle First Approach. Aapti Institute. https://thedataeconomylab.com/wp-content/uploads/2021/07/Aapti-Understanding-NPD-Sharing-Report.pdf

OECD (2019), *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, OECD Publishing, Paris, *https://doi.org/10.1787/276aaca8-en*.

OECD (2005), *Principles and Guidelines for Access to Research Data from Public Funding*, OECD Publishing, Paris, http://www.oecd.org/sti/sci-tech/38500813.pdf.

Parliament of Australia. (2020). Data Availability and Transparency Bill, 2020. https://parlinfo.aph.gov.au/parlInfo/download/legislation/bills/r6649_first-reps/toc_pdf/20174b01.pdf;fileType=application%2Fpdf

Standards Council of Canada. (2020). Canadian Data Governance Standardization Roadmap. https://www.scc.ca/en/system/files/publications/SCC_Data_Gov_Roadmap_EN.pdf

Surblyte, G. (2016). Data as a Digital Resource. *SSRN Electronic Journal*, (16). https://doi.org/10.2139/ssrn.2849303

Taylor, L. (2013), *Hacking a Path through the Personal Data Ecosystem*, https://linnettaylor.wordpress.com/2013/12/12/hacking-a-path-through-the-personal-dataecosystem/

World Economic Forum. (2020). A Roadmap for Cross-Border Data Flows: Future-Proofing Readiness and Cooperation in the New Data Economy.

https://www3.weforum.org/docs/WEF_A_Roadmap_for_Cross_Border_Data_Flows_2020.pdf