



---

# **LEGAL ROADMAP FOR NON-PERSONAL DATA GOVERNANCE IN INDIA**

**A STUDY WITH A SPECIAL FOCUS ON INTELLECTUAL PROPERTY,  
COMPETITION, REGULATORY AND THE DATA SECURITY DIMENSION**

**Authors: Dr Yogesh Pai and Kruti Kachhwaha**

**SPRIHA- DPIIT IPR Chair**

**&**

**Centre for Innovation, Intellectual Property and Competition (CIIPC)**

**National Law University Delhi**

## **TABLE OF CONTENTS**

|  |               |
|--|---------------|
| <b>ACKNOWLEDGEMENTS</b>  | <b>- 6 -</b>  |
| <b>LIST OF ABBREVIATIONS</b>   | <b>- 8 -</b>  |
| <b>CHAPTER I: NON-PERSONAL DATA-DEFINITION AND CONTEXT</b>                     | <b>- 10 -</b> |
| 1. The Concept of Data and Big Data  | - 10 -        |
| 2. Data: New Oil?  | - 12 -        |
| 2.1. Data: A Public Good?  | - 13 -        |
| 2.2. Data and Openness   | - 15 -        |
| 2.2.1. Open Data: Trade Offs   | - 16 -        |
| 2.2.2. Data and Information Privacy: Is Non-Personal Data Personal?            | - 18 -        |
| 2.3. Data and its Benefits to Economy  | - 20 -        |
| 3. Defining Personal and Non-Personal Data                                     | - 23 -        |
| 4. Development of Data Protection and Governance Laws in India                 | - 26 -        |
| 5. Global Perspectives on Non-Personal Data Governance                         | - 29 -        |
| <b>CHAPTER II: LEGAL PROTECTION FOR NON-PERSONAL DATA</b>                      | <b>- 33 -</b> |
| 1. Perspectives on Legal Scope and Definition of Data:                         | - 33 -        |
| 2. Data and Intellectual Property Protection                                   | - 36 -        |
| 2.1. Understanding the Intersection of Intellectual Property and Data          | - 36 -        |
| 2.2. Copyright Protection for Original Databases                               | - 37 -        |
| 2.3. The EU Database Directive   | - 39 -        |
| 2.3.1. EU's Unique Approach: Sui Generis Protection for Non-Original Databases | - 40 -        |
| 2.3.2. Scope and Impact of the Sui Generis right:                              | - 41 -        |
| 2.3.3. Problems and Challenges with the Sui Generis Right                      | - 42 -        |
| 2.3.4. The EU Database Directive vis-a-vis Big Data                            | - 45 -        |
| 2.4. The Data Producer's Right   | - 47 -        |

|  |               |
|--|---------------|
| 2.5. Challenges in Protecting Data as Intellectual Property  | - 49 -        |
| 3. Unfair Competition and Data   | - 50 -        |
| 3.1. Trade secret  | - 50 -        |
| 3.2. Big Data and Trade Secret Protection  | - 52 -        |
| 3.3. Big Data and Issues with Trade Secret Protection  | - 53 -        |
| 3.4. Data Misappropriation and Unjust Enrichment   | - 55 -        |
| 4. Contract Law and Data   | - 56 -        |
| 4.1. Contractual Issues in Data Transactions/Sharing   | - 58 -        |
| 4.2. Non-disclosure agreements and their significance  | - 61 -        |
| 4.3. Standard Rules for Data Sharing Contracts   | - 62 -        |
| <b>CHAPTER III: CYBER-SECURITY DIMENSION INVOLVING DATA BREACH AND DATA LEAK</b>                           | <b>- 66 -</b> |
| 1. Cyber Space and Cyber Security  | - 66 -        |
| 2. Data as Intellectual Capital and How Firms Protect it   | - 69 -        |
| 2.1. Cyber Security Threats  | - 71 -        |
| 2.2. Data Breach and Its Impact  | - 73 -        |
| 3. Overview Of Indian Data Protection Laws   | - 75 -        |
| 4. Remedies for Data Leak and Breach   | - 79 -        |
| 4.1. Risk Implementation and Breach Notification   | - 81 -        |
| 4.2. China's Approach to Data Security: Cross Border Data Flow Restrictions and Data Localization Measures | - 85 -        |
| <b>CHAPTER IV: COMPETITION LAW AND NON-PERSONAL DATA</b>   | <b>- 88 -</b> |
| 1. Data Markets and Competition Law  | - 88 -        |
| 1.1. Data as Competitive Intelligence  | - 89 -        |
| 1.2. A Rising Concern  | - 91 -        |
| 2. Anti-Competitive Practices Related to Data  | - 93 -        |

|  |                |
|--|----------------|
| 2.1. Exclusionary Conduct and Data as an Essential Resource  | - 95 -         |
| 2.1.1. Data as an Essential Resource   | - 97 -         |
| 2.1.2. Big Data: Pro-Competitive?  | - 100 -        |
| 2.2. Data and Abuse of Dominance   | - 101 -        |
| 2.3. Data Driven Mergers   | - 105 -        |
| 3. Conclusion  | - 108 -        |
| <b>CHAPTER V: REGULATORY APPROACHES TO DATA GOVERNANCE</b>   | <b>- 110 -</b> |
| 1. Approaches to Data Governance   | - 110 -        |
| 1.1. An Ex-Ante Approach to NPD?   | - 112 -        |
| 2. International Approaches to NPD Regulation  | - 114 -        |
| 2.1. European Union  | - 114 -        |
| 2.2. Germany   | - 118 -        |
| 2.3. United Kingdom  | - 119 -        |
| 2.4. Japan   | - 121 -        |
| 2.5. USA   | - 123 -        |
| 2.6. China   | - 125 -        |
| 2.7. Australia   | - 128 -        |
| 2.8. Other Ex Ante Initiatives   | - 129 -        |
| 3. Indian Approach to Ex-Ante Regulation of Digital Market   | - 130 -        |
| 4. Indian Approach to NPD Regulation   | - 136 -        |
| 4.1. Highlights from the Report  | - 136 -        |
| 4.2. Criticisms of the Report  | - 138 -        |
| <b>CHAPTER VI: REGULATORY AUTONOMY FOR NON-PERSONAL DATA:<br/>NAVIGATING THE PATCHWORK OF GLOBAL DATA GOVERNANCE</b> | <b>- 142 -</b> |
| 1. Introduction  | - 142 -        |

|   |         |
|---|---------|
| 2. WTO Framework and its Limitations  | - 143 - |
| 2.1. GATT and Digital Trade   | - 144 - |
| 2.2. GATS and Digital Trade   | - 145 - |
| 2.3. Lingering WTO Developments   | - 146 - |
| 2.4. WTO Exceptions and Data Governance                                     | - 147 - |
| 3. TRIPS and Non-Personal Data  | - 151 - |
| 3.1. Copyright and Big Data   | - 151 - |
| 3.2. Mandatory Data Sharing vis-a-vis Article 13 of the TRIPS Agreement     | - 152 - |
| 3.3. Trade Secret Protection and Big Data                                   | - 155 - |
| 4. FTAs and Non-Personal Data Governance                                    | - 156 - |
| 5. India's Data Localization Measures in Light of International Obligations | - 159 - |
| 6. Conclusion   | - 161 - |

## **CONCLUSION**

**164**

## **ACKNOWLEDGEMENTS**

### **About NFCG**

In 2003, the Ministry of Corporate Affairs (MCA) led a unique PPP model to set up the National Foundation for Corporate Governance in partnership with the Confederation of Indian Industry, the Institute of Company Secretaries of India, and the Institute of Chartered Accountants of India. Subsequently, the Institute of Cost Accountants of India, National Stock Exchange and the Indian Institute of Corporate Affairs also joined with an objective to promote good Corporate Governance practices both at the level of individual corporates and Industry as a whole. NFCG endeavors to create a business environment that promotes voluntary adoption of good corporate governance practices.

### **Vision**

Be the Key Facilitator and Reference Point for highest standards of Corporate Governance in India

### **Mission**

- To foster a culture of good Corporate Governance
- To create a framework of best practices, structure, processes and
- Ethics
- To reduce the existing gap between Corporate Governance framework & actual compliance by corporates
- To facilitate effective participation of different stakeholders
- To catalyse capacity building in emerging areas of Corporate Governance

**National Law University Delhi**

We would like to express our sincere gratitude to National Law University Delhi for providing the platform and resources necessary for the successful completion of this research report. Our heartfelt thanks go to Prof. (Dr) G.S. Bajpai, Vice Chancellor, and Prof. (Dr) Ruhi Paul, Registrar, for their unwavering support and encouragement throughout this project.

We are deeply grateful to Anupriya Dhonchak for her invaluable contribution in conducting the qualitative survey and for guiding the research team on some aspects of the report at the early stages. However, all infirmities in the report are entirely ours. Additionally, we extend our appreciation to the various industry experts who participated in the qualitative survey. Their expertise and willingness to share their knowledge have been crucial to the depth and quality of our research.

We would also like to express our gratitude to the team members of the Centre for Innovation, Intellectual Property and Competition (CIIPC) and the SPRIHA- DPIIT IPR Chair at NLU Delhi, for their continued support and dedication to this project. We thank the student fellows, associates and assistants at the CIIPC-IPR Chair for providing research assistance. We are truly thankful for the support and contributions of all those involved in the journey of bringing this report to fruition.

#### Funding and Copyright Notice:

This research project has been funded by the National Foundation for Corporate Governance. The authors declare no other conflict. As per the funding requirements “[T]he Copyright, Trademarks, and other Intellectual property rights on the research work/ study would be owned jointly by NFCG and the Institution.”

## LIST OF ABBREVIATIONS

|          |   |
|----------|---|
| 3VS      | Volume, Velocity and Variety  |
| AI       | Artificial Intelligence   |
| CERT     | Computer Emergency Response Team                                      |
| CERT-In  | Computer Emergency Response Team India                                |
| CCI      | Competition Commission of India                                       |
| CPTPP    | Comprehensive and Progressive Agreement for Trans-Pacific Partnership |
| DPDP Act | Digital Personal Data Protection Act                                  |
| FTA      | Free Trade Agreement  |
| FRAND    | Fair, Reasonable and Non-Discriminatory                               |
| GDPR     | General Data Protection Regulation                                    |
| GATS     | General Agreement on Trade in Services                                |
| GATT     | General Agreement on Trade and Tariffs                                |
| HVD      | High Value Datasets   |
| IDC      | International Data Corporation  |
| IIA      | International Investment Agreement                                    |
| IoT      | Internet of Things  |
| ISO      | International Standard Organisation                                   |
| IT       | Information Technology  |
| MCA      | Ministry of Corporate Affairs   |
| MeiTY    | Ministry of Electronics and Information Technology                    |
| MFN      | Most Favored Nation   |
| NDA      | Non-Disclosure Agreement  |
| NPD      | Non-Personal Data   |
| OECD     | Organisation for Economic Co-operation and Development                |

|          |   |
|----------|---|
| PDP Bill | Personal Data Protection Bill                   |
| PII      | Personally Identifiable Information             |
| QoD      | Quality of Data                                 |
| TDM      | Text and Data Mining                            |
| TFEU     | Treaty On the Functioning of the European Union |
| WTO      | World Trade Organisation                        |
| WIPO     | World Intellectual Property Organisation        |

## CHAPTER I: NON-PERSONAL DATA-DEFINITION AND CONTEXT

With the rise of Artificial Intelligence (AI), data has become an essential commodity input fueling Large Language Models and other emerging technologies. This has led to the rise in fortunes of AI tech companies and the Big-Data Industry, where there is significant convergence in terms of the business trajectory. Access to data and how the terms of access to such data are regulated remains one of the greatest industrial policy questions of our milieu. World over, regulations are being shaped and reshaped in this regard. While there is significant investment in AI technologies, there are concerns about such data being harvested without appropriate compensation to producers of the data. While the critical question of AI companies scraping copyrighted works through web-crawling on the internet is being litigated across various jurisdictions as of 2024, the question of what it means for non-copyrighted works, particularly non-personal data has not been studied comprehensively. This is pertinent in the context of the prevailing regulatory impulse to impose sharing requirements on owners of data. While these big questions have different dimensions which have been discussed in this report, an essential question of what data constitutes in the context of non-personal data is key to understanding why it is crucial to other aspects studied in this report.

### 1. The Concept of Data and Big Data

The word “data” comes from the latin word “datum” which literally means “*thing given*”.<sup>1</sup> According to the Merriam Webster dictionary, data means “*factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation*” or “*information in digital form that can be transmitted or processed*”.<sup>2</sup> The Cambridge dictionary defines data as “*information, especially facts or numbers, collected to be examined and considered and used to help decision-making, or information in an electronic form that can be stored and used by a computer*”.<sup>3</sup> The oldest census was conducted back in 2000-1000 BCE which indicates that data has been used since quite a long time. Throughout history, the use and collection of data has evolved, progressed and has become more systematic. The term ‘big data’ was first used in the

---

<sup>1</sup> Leland Wilkinson, *The Grammar of Graphics* (Springer 2005).

<sup>2</sup> Merriam-Webster, ‘Definition of Data’ <<https://www.merriam-webster.com/dictionary/data>> accessed 10 December 2024.

<sup>3</sup> Cambridge Dictionary, ‘Data’ <<https://dictionary.cambridge.org/dictionary/english/data>> accessed 10 December 2024.

mid-1990s in reference to huge or massive datasets. Since then, big data has been evolving constantly and quickly, so much so that there is no universal statement denoting its meaning exists.<sup>4</sup> Big data has been defined in several ways. One of the first and most popular definitions is by Doug Laney and it focuses on the characteristics of big data i.e. volume, velocity and variety (the three Vs).<sup>5</sup> According to this definition, Big Data has three traits that are volume (consisting of enormous quantities of data), velocity (created in real-time) and variety (being structured, semi-structured and unstructured). The velocity of data is also twofold; the speed at which the data flows and the pace at which it is collected, analyzed, and retrieved.<sup>6</sup> Big Data consists of enormous amounts of data that is created in realtime and can be structured, unstructured or semi-structured.<sup>7</sup> Some definitions of Big Data focus on its “*scalable architecture for efficient storage, manipulation, and analysis*”<sup>8</sup> and some discussions highlight that a data set is big when it “exceeds the processing capacity of conventional database systems” and requires the choice of “an alternative way to process it”.<sup>9</sup> The Department of Science and Technology, India also characterizes big data using the three Vs and defines Big Data as “*data whose scale, diversity, and complexity require new architecture, techniques, algorithms, and analytics to manage it and extract value and hidden knowledge from it.*”<sup>10</sup> Further, according to Microsoft, Big Data is about applying ‘*serious computing powers*’ to massive sets of information and Big Data Analytics refers to the “*methods, tools, and applications used to collect, process, and derive insights from varied, high-volume, high-velocity data sets.*”<sup>11</sup> Some definitions also focus on the social role of Big Data.

---

<sup>4</sup> Andrea De Mauro, Marco Greco and Michele Grimaldi, ‘What Is Big Data? A Consensual Definition and a Review of Key Research Topics’ (2015) AIP Conf Proc 97.

<sup>5</sup> Rob Kitchin and Gavin McArdle, ‘What Makes Big Data, Big Data? Exploring the Ontological Characteristics of 26 Datasets’ (2016) 3(1) Big Data & Society <<https://journals.sagepub.com/doi/full/10.1177/2053951716631130>> accessed 24 January 2025 .

<sup>6</sup> Harry E Pence, ‘What is Big Data and Why is it Important?’ (2014) 43(2) J Educ Technol Syst 159.

<sup>7</sup> Rob Kitchin and Gavin McArdle (n 5) .

<sup>8</sup> National Institute of Standards and Technology, ‘NIST Big Data Interoperability Framework: Volume 1, Definitions, Version 3’ (NIST 2019) <<https://doi.org/10.6028/NIST.SP.1500-1r2>> accessed 11 December 2024.

<sup>9</sup> Andrea De Mauro, Marco Greco and Michele Grimaldi, ‘What Is Big Data? A Consensual Definition and a Review of Key Research Topics’ (2015) 1644(1) AIP Conf Proc 97-104.

<sup>10</sup> Department of Science & Technology, ‘Big Data Initiative’ <<https://dst.gov.in/big-data-initiative-1>> accessed 10 December 2024.

<sup>11</sup> Microsoft Azure, ‘What Is Big Data Analytics?’ <<https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-big-data-analytics>> accessed 11 December 2024.

According to Boyd and Crawford, big data is “a cultural, technological, and scholarly phenomenon that rests on the interplay of technology, analysis and mythology.”<sup>12</sup>

The core concept of big data can therefore be said to comprise massive amounts of information (Volume, Velocity & Variety) that require specific technology & analytical methods to transform the information and extract economic values from its insights.<sup>13</sup> These attributes of big data create great potential for predictive analytics and decision making<sup>14</sup> and thus set it apart from traditional data or small datasets. Since big data holds voluminous amounts of information, it can be used for predictive analytics with the primary objective of predictive analytics is to analyze the available data to optimize decision making and use intelligent algorithms that can learn solutions from data.<sup>15</sup>

## **2. Data: New Oil?**

Data is commonly being termed as “the new oil” in the digital economy. It is an essential resource for driving innovation and thus a critical asset. However, there are substantial notable dissimilarities between data and oil. They differ in scarcity/availability, rivalry and their relationships with people.<sup>16</sup> Data is available in massive amounts, on the contrary, oil is a scarce resource and scarce resources are governed differently than those available abundantly.<sup>17</sup> Further, unlike oil, data is an intangible asset that is a non-rivalrous & a non-excludable public good. If access to data is available, it can be used simultaneously by many people or entities because unlike oil, data can be copied and processed multiple times without getting depleted. Furthermore, it should be noted that data is heavily connected to people, in fact, it comes from the people and reflects their behavior. It can also trace the actions of enterprises and companies while oil is just a

---

<sup>12</sup> Danah Boyd and Kate Crawford, ‘Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon’ (2012) 15(5) *Information, Communication & Society* 662.

<sup>13</sup> Andrea De Mauro, Marco Greco and Michele Grimaldi (n 9).

<sup>14</sup> Caryn Devins, Teppo Felin, Stuart Kauffman, and Roger Koppl, ‘The Law and Big Data’ (2017) 27(2) *Cornell J L & Public Policy* 357.

<sup>15</sup> *ibid.*

<sup>16</sup> Lauren Henry Scholz, ‘Big Data Is Not Big Oil: The Role of Analogy in the Law of New Technologies’ [ 2019] 86 *Tennessee Law Review* 863. See also A R Greene and Samuel J Gilbert, ‘More Data, More Power? Towards a Theory of Digital Legitimacy’ (2020) <<http://dx.doi.org/10.2139/ssrn.3773898>> accessed 11 December 2024.

<sup>17</sup> S A Aaronson, ‘Data is different, and that’s why the world needs a new approach to governing cross-border data flows’ (2019) *CIGI Papers* No 197 — November 2018 <[https://www.cigionline.org/static/documents/documents/paper%20no.197\\_0.pdf](https://www.cigionline.org/static/documents/documents/paper%20no.197_0.pdf)> accessed 20 January 2025 .

natural resource. It is due to these logical differences, that the analogy equating data with oil fails. Oil and data both have different characteristics, and these characteristics shape how both the sectors work. Oil, a scarce natural resource, requires expensive infrastructure and only a few players are involved in processing it. On the contrary, data is cheap to access and process. Therefore, it is argued that the structural entry barriers that are present in processing oil do not exist in processing data as every entity, at some level, is engaged in collecting some form of data.<sup>18</sup>

### **2.1. Data: A Public Good?**

Firms have access to large amounts of data and have the ability and sophisticated technology to transform and process this data to extract value from it. Data collection and transformation requires a lot of investment and it's due to this opportunity cost that firms try to fiercely protect data. With the increased data collection and processing by firms, the question of “who owns data” becomes critical.

Big data analysis can be very advantageous. Since data is non-rival and non-exhaustive, it can be used multiple times by multiple entities and its value won't be diminished and neither would its use by one entity exclude the possibility of use by others. Further data can be repurposed and reused. Using and analyzing data effectively can generate information that not only holds economic value but also social value.<sup>19</sup> Further, with the development of advanced technology and artificial intelligence, even more value can be generated from data to promote innovation and foster development of new products. Therefore, access to data is essential for competition and innovation, for both businesses and individuals. Data sharing and access can aid in identifying the future needs of the government and society and facilitate improving infrastructure and utilities. The Mckinsey Global Institute conducted a study and found that the re-use of public and private sector data in the areas of education, transportation, consumer products, electricity, oil and gas, health care and consumer finance can create value of about USD 3 trillion dollars globally.<sup>20</sup>

---

<sup>18</sup> Lauren Henry Scholz (n 16) .

<sup>19</sup> OECD, ‘Enhancing Access to and Sharing of Data : Reconciling Risks and Benefits for Data Re-use across Societies’ (OECD 2019) <<https://doi.org/10.1787/276aaca8-en>> accessed 11 December 2024.

<sup>20</sup> James Manyika and others, ‘Open Data: Unlocking Innovation and Performance with Liquid Information’ (McKinsey Global Institute 2013) <<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information>> accessed on 11 December 2024.

The OECD has categorized the economic and social benefits of data access and sharing into five categories.<sup>21</sup> These categories are: “i) *greater transparency, accountability and empowerment of users, for instance, when open data are used for cross-subsidising the production of public and social goods; ii) new business opportunities, including the creation of start-ups and in particular for data intermediaries and mobile app developers; iii) competition and co-operation within and across sectors and nations, including the integration of value chains; iv) crowdsourcing and user-driven innovation; and v) increased efficiency thanks to linkage and integration of data across multiple sources.*”

The non-rivalrous nature of the data signifies that it can be used by a wide range of people for a wide range of purposes. Due to this non-excludable and non-rivalrous nature, data can be termed as public good in economic terms. However, despite these characteristics, data in practice is not a public good since it is in *de facto* possession of the firms collecting it. Further, there are several legal and technological means through which data can be protected by the collecting entity to protect data against misappropriation. There are various legal means to protect unauthorized use of data or functionally exclude data from the public.<sup>22</sup> For example, copyright, trade secrets and contracts. Private entities tend to protect their private intent data and exclude others from using it since there aren't enough incentives to share their data or there is no safe and secure way to share data.<sup>23</sup> As it is possible to exclude others from using data or availing the benefits of data, which is essentially a non-rivalrous good, an issue arises regarding balancing the interests of private entities and the greater public interest. When data is limited to the private sphere, the public can get deprived of its benefits. This deprivation not only denies the public of the chance of growth but it can also put the public in a worse off situation. For illustration purposes, the 2010 Haiti earthquake can be looked at. In 2010, during the Haiti earthquake incident, the donations and the disaster relief could not reach the people efficiently since they were scattered. Census data could not have helped in ensuring that relief reaches everyone in need. However, by using mobile phone and location

---

<sup>21</sup> OECD (n 19) .

<sup>22</sup> Linnet Taylor, ‘The Ethics of Big Data as a Public Good: Which Public? Whose Good?’ [ 2016] 374 Philosophical Transactions: Mathematical, Physical and Engineering Sciences 1.

<sup>23</sup> Vivien Foster and others, World Development Report 2021: Data for Better Lives (World Bank Group 2021) <<https://wdr2021.worldbank.org/the-report/>> accessed 11 December 2024.

data, the exact real-time location of the people could have been determined.<sup>24</sup> Data was also an integral part of the COVID-19 pandemic for monitoring and risk management.<sup>25</sup> Therefore, combining public<sup>26</sup> and private interest data can offer significant inputs for the benefit of the public.

Data as a public good can be defined in two possible ways. Firstly, the notion that international organizations should have access to data so that they can promote social good. The second is the notion that data should formally be considered as a public good for the benefit of the society.<sup>27</sup> The public does not have fair and equitable access to data neither can they have the ability to analyze and understand the data. Similarly, not every country has the ability to collect large amounts of data and monetise it.<sup>28</sup> Due to this data divide, data can be both a commercial asset and a public good.<sup>29</sup>

## 2.2. Data and Openness

It is in this context that the discussions on open data are gaining traction. Considering the expense and resources required to be invested in creation of databases, its access is restricted in different ways and the open data movement seeks to reform this.<sup>30</sup> The Open Data Movement is driven by three principles; openness, participation and collaboration.<sup>31</sup> Open data means data which can be freely accessed, used, modified and shared for any purposes.<sup>32</sup> Open data can be beneficial for the greater public good in multiple ways and its uses can be categorized into three categories, (i)

---

<sup>24</sup> *ibid.*

<sup>25</sup> European Data Protection Supervisor, 'Data as a Public Good: Building a Healthier Digital Future' (EDPS Blog, 10 November 2020) <[https://www.edps.europa.eu/press-publications/press-news/blog/data-public-good-building-healthier-digital-future\\_en](https://www.edps.europa.eu/press-publications/press-news/blog/data-public-good-building-healthier-digital-future_en)> accessed 11 December 2024.

<sup>26</sup> *ibid.*

<sup>27</sup> Linnet Taylor (n 22) .

<sup>28</sup> UNCTAD, *Digital Economy Report 2021: Cross-border Data Flows and Development – For Whom the Data Flow* (UNCTAD 2021) <[https://unctad.org/system/files/official-document/der2021\\_en.pdf](https://unctad.org/system/files/official-document/der2021_en.pdf)> accessed 11 December 2024.

<sup>29</sup> Linnet Taylor (n 22) . See also Susan Ariel Aaronson, 'Why Are We Talking about Data as a Public Good?' 'Could a Global "Wicked Problems Agency" Incentivize Data Sharing?' (2023) Centre for International Governance Innovation <<http://www.jstor.org/stable/resrep48561.8>> accessed 13 December 2024.

<sup>30</sup> Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences* (SAGE Publications Ltd 2014).

<sup>31</sup> *ibid.*

<sup>32</sup> Open Data Handbook, 'What is Open Data?' (Open Data Handbook, 2024) <<https://opendatahandbook.org/guide/en/what-is-open-data/>> accessed 13 December 2024.

innovation & economic growth, (ii) political accountability & democratic participation and, (iii) public sector efficiency.<sup>33</sup> It can facilitate use and re-use of information and can push the creation of new businesses, new products, services and therefore promote innovation & entrepreneurship.<sup>34</sup> Since data can provide valuable insights into consumer behavior, it can be used to make efficient decisions regarding product and service development and informed consumption.<sup>35</sup> It can also help in increasing political accountability and dealing with corruption practices. According to the G8 Open Data Charter, open data on natural resources, land management and development spending can increase government accountability. Further, it can help in saving resources and improving public services like road and traffic management, monitoring energy consumption and infrastructure planning. More importantly, it can help in preventing duplication in information collection and thus saving resources & increasing investment.<sup>36</sup>

While open data can be beneficial, it can also pose several challenges. There are several factors that can affect a platform's decisions when it comes to open data. For instance, it can enhance already existing privacy concerns regarding sensitive personal information.<sup>37</sup> As discussed below, although these concerns may be acute in cases of open data, all data is subjected to the risk of breach of privacy.

### **2.2.1. Open Data: Trade Offs**

When it comes to open data or opening platforms to others, there are several factors that affect a platform owner's decision. Several trade-offs are made on a case-by-case basis. Open platforms generally indicate that the restrictions on access, use and development of platforms are eased, and

---

<sup>33</sup> Frederik Zuiderveen Borgesius, Jonathan Gray and Mireille van Eechoud, 'Open Data, Privacy and Fair Information Principles: Towards a Balancing Framework' (2015) 30(3) Berkeley Technology Law Journal 2073.

<sup>34</sup> Stott Andrew, 'Open Data for Economic Growth' (World Bank) <<https://www.worldbank.org/content/dam/Worldbank/document/Open-Data-for-Economic-Growth.pdf>> accessed 11 December 2024. See also G8, Open Data Charter: Principles (2013) principle 5 <<https://www.gov.uk/government/publications/open-data-charter>> accessed 11 December 2024.

<sup>35</sup> James Manyika and others, 'Open Data: Unlocking Innovation and Performance with Liquid Information' (McKinsey Global Institute 2013) <<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information>> accessed on 11 December 2024.

<sup>36</sup> *ibid.*

<sup>37</sup> Paul Ohm, 'The Underwhelming Benefits of Big Data' (2013) 161(1) University of Pennsylvania Law Review Online 339 <[https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=1113&context=penn\\_law\\_review\\_online](https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=1113&context=penn_law_review_online)> accessed 11 December 2024.

the platforms are made available and accessible to third parties to innovate.<sup>38</sup> There are two main approaches for opening platforms. The platform owners can either “relinquish control over their resources or selectively open up technologies through interfaces”. In addition, the openness of a platform can be twofold, first, platform-to-platform openness and second, platform-to-app openness. Both of these facilitate functionality, but they do differ from each other.<sup>39</sup> A platform is a “stand-alone structure” and an app is a “complementary product” which can add to a platform’s functionality. Therefore, platform-to-app openness will add functionality to the platform and platform-to-platform openness will create “standalone functionality”.

As per a study, there can be four levels of openness: “(1) on sponsor level (i.e. towards platform owners); (2) provider level (i.e. towards other platform providers); (3) technology level (i.e. interoperability between platforms); and (4) user-level (i.e. towards users from other platforms) but the definition of openness also depends on how interoperable two platforms are.”<sup>40</sup> Firms often relinquish control strategically, as a matter of a rational business practice to encourage growth and adoption of their platforms. They do not always seek to maintain full control over their platform. Rather they aim for a strategic forfeiture that can encourage user adoption and prevent lock-in concerns that can sustain long-term progress.<sup>41</sup>

Platform openness can have direct and indirect costs and benefits and while evaluating whether or not a platform should be open and accessible, various trade-offs have to be made. Studies have shown that openness decisions are made very selectively by platform owners depending on the business opportunities that it can advance for them. Openness can stimulate growth and innovation but at the same time if the openness is too high, it can induce a fear of competition. Further, “low quality complements can harm a platform’s integrity and reputation”. On one hand platforms can be open and attract users but at the same time closed platforms can maintain their exclusivity by

---

<sup>38</sup> Kevin Boudreau, ‘Open Platform Strategies and Innovation: Granting Access vs Devolving Control’ (2010) 56(10) *Management Science* 1849.

<sup>39</sup> Lars Mosterd and others, ‘Context Dependent Trade-Offs around Platform-to-Platform Openness: The Case of the Internet of Things’ [ 2021] 108 *Technovation*.

<sup>40</sup> Jan Ondrus, Avinash Gannamaneni and Kalle Lyytinen, ‘The Impact of Openness on the Market Potential of Multi-Sided Platforms: A Case Study of Mobile Payment Platforms’ (2015) 30(3) *Journal of Information Technology* 260, See also Lars Mosterd and others (n 39).

<sup>41</sup> Jonathan M. Barnett, ‘The Host’s Dilemma: Strategic Forfeiture in Platform Markets for Informational Goods’ (2011) 124 *Harvard Law Review* 1861.

retaining (or locking in) their customers and avoid imitation thereby, creating entry barriers in the market. Further, decisions regarding openness also vary depending on the nature of the platform. For example, a for-profit organization will consider the financial implications of platform openness. Whereas, a not-for-profit organization would be more willing to prioritize social good. On similar lines, a company that considers IOT as complementary to its business of hardware sales (a company producing and selling physical devices), would want to maximize openness and interoperability to increase their value. Therefore, openness of data and platforms is not only dependent on the nature & context of data but also other circumstances as discussed above.

Issues also arise when it comes to public disclosure of government data. Disclosure of data by government authorities can have both positive and negative effects. Positive effects include: social, political, economic, operational and technical benefits all of which can promote transparency, increase participation of the citizens, increase knowledge, stimulate growth, facilitate innovation, and etc. The negative effects of disclosure of government data include: breach of individual privacy, information overload, issues concerning undesirable surveillance, data misinterpretation & misuse. Further, releasing incorrect information or low quality information and information regarding wrong government decisions can decrease the faith of the citizens in the government. Moreover, it is also argued that open data might just empower the already empowered people. Since citizens do not have access to sophisticated technology and advanced skills through which they can interpret and process data, people who already possess this infrastructure will get an increased access to data that they will be able to use effectively.<sup>42</sup>

### **2.2.2. Data and Information Privacy: Is Non-Personal Data Personal?**

One of the pertinent definitional issues concerning NPD is whether the non-personal nature of the data remains its core feature. Even if data is anonymised before it is made public, it can continue to pose privacy risks. With technological developments in the digital economy, permanent anonymization of data is considered as an utopian concept. Data anonymization is a process through which identity can be preserved by using certain techniques to delete or remove identifiers that can link information to an individual. The International Organization for Standardization has

---

<sup>42</sup> Anneke Zuiderwijk and others, 'Towards Decision Support for Disclosing Data: Closed or Open Data?' [ 2015] 20 Information Polity 103.

defined “anonymisation” as the *"process by which personal data is irreversibly altered in such a way that a data subject can no longer be identified directly or indirectly, either by the data controller alone or in collaboration with any other party."*<sup>43</sup> Pseudonymisation of data is another way through which individual privacy can be protected. It is a technique that involves replacing direct identifiers with pseudonyms so that the identities of individuals or data principals are not disclosed or revealed.<sup>44</sup> The ISO has defined pseudonymisation as de-identification that *“removes the association with data subject and adds an association between a particular set of characteristics relating to the data subject and one or more pseudonyms."* Pseudonymisation allows the use of data sources without risking the identity of individuals. However, the utility of the data and the benefits it can accrue can get significantly affected by pseudonymisation at source as it can affect the likelihood of using other linkability methods for exploitation of dataset.<sup>45</sup>

The notion that anonymisation of datasets would pose no risks to privacy has been debunked as a myth.<sup>46</sup> It has been proven many times that anonymisation is not effective in protecting the privacy of individuals. For example, in 2006, Netflix released its data and offered a million-dollar prize for the improvement of Netflix’s movie recommendation algorithm. Researchers from the Texas University were able to identify individuals from Netflix’s publicly available anonymised data that contained its subscribers film ratings. This was done by using the Internet Movie Database as background knowledge and statistical methods. The de-anonymization as a result helped the researchers in revealing the political preferences and other sensitive information of individuals. The researchers thereafter explained that *“Once any piece of data has been linked to a person’s real identity, any association between this data and a virtual identity breaks the anonymity of the latter”*.<sup>47</sup> Therefore, the data was de-anonymised by cross referencing with other data found online.

---

<sup>43</sup> ISO/IEC, ISO/IEC 29100:2011 *Information technology — Security techniques — Privacy framework* (ISO 2011). Also see Sharp Cookie Advisors, ‘Anonymization and GDPR Compliance; an Overview’ (*GDPR Summary*, 21 July 2020) <<https://www.gdprsummary.com/anonymization-and-gdpr/>> accessed 11 December 2024.

<sup>44</sup> ISO/IEC, ISO/IEC 20889:2018, *Privacy Enhancing Data De-identification Techniques* (ISO 2018).

<sup>45</sup> Harvey Goldstein and Katie Harron, “Pseudonymisation at Source” *Undermines Accuracy of Record Linkage*’ (2018) 40(2) *J Public Health (Oxf)* 219.

<sup>46</sup> Ira Rubinstein and W Hartzog, ‘Anonymization and Risk’ (2016) 91(2) *Washington Law Review* 703.

<sup>47</sup> Big Data Privacy and Intellectual Property in a Comparative Setup. Arvind Narayanan and Vitaly Shmatikov, ‘How To Break Anonymity of the Netflix Prize Dataset’ [2006] *ArXiv* <<https://doi.org/10.48550/arXiv.cs/0610105>> accessed 11 December 2024.

This example shows that anonymisation of data sets by removing personal identifiers will not always be successful in protecting individuals from identification.

A question that arises in this context now is how data can be shared publicly without threatening privacy of the people and competition. Open data, even though a conceptually sound and an advantageous concept can thus pose critical and tricky privacy issues. Thus, there is a need to balance the benefits and burdens of data sharing. The issues that emerge with open data sharing do not call for abandoning the open data movement, rather they call for open data initiatives that are more mindful of these issues and how data is shared.

### **2.3. Data and its Benefits to Economy**

Vast volumes of data are produced every year, and the International Data Corporation (IDC), in its forecast, has projected that the volume of data will be increasing by 40 percent each year.<sup>48</sup> It has become a crucial input resource for different sectors and is bound to intertwine with the economy and trade intricately.<sup>49</sup> Data has high economic value and substantial investments are made in creating, collecting and processing data. A few uses of big data in various sectors and in various ways is discussed below to provide a picture of how beneficial big data and big data analysis can be.

→ Agricultural Sector and Smart Farming: Smart farming focuses on use of information and communication technology in managing farming cycles. The use of big data for predictive analysis helps in farming operations, real time operational decisions and developing new business models. Big data is currently at its early stage when it comes to farming but it is bound to change the farming sector and aid in solving global issues concerning food security, food safety and sustainability. Big data analytics will change farm management and operations for the better.<sup>50</sup> For example, The Climate Corporation, founded in 2006

---

<sup>48</sup> Carol Corrado and others, 'The Value of Data in Digital-Based Business Models: Measurement and Economic Policy Implications' (OECD, 2022) <<https://doi.org/10.1787/d960a10c-en>> accessed 11 December 2024.

<sup>49</sup> UNCTAD, '*The Value and Role of Data in Electronic Commerce and the Digital Economy and Its Implications for Inclusive Trade and Development*' (UNCTAD 2019) <[https://unctad.org/system/files/official-document/tdb\\_edc3d2\\_en.pdf](https://unctad.org/system/files/official-document/tdb_edc3d2_en.pdf)> accessed 11 December 2024.

<sup>50</sup> Jaak Wolfert and others, 'Big Data in Smart Farming – A Review' (2017) 153 *Agricultural Systems* 69.

takes weather data from various locations and along with data related to crop root structure and soil to locate weather patterns and events and helps farmers in maximizing their crop yields.<sup>51</sup>

- E-Commerce Sector: Big data analysis can provide deep insights regarding customer behavior and market trends and that can eventually help in improving customer experience. It also helps in detecting payment-related fraudulent activities, predict trends & demand, increase sales, etc. Big data analytics therefore go hand in hand with e-commerce.<sup>52</sup> For example, Amazon uses data collected from customers and uses it to improve their recommendation engine. Amazon's approach in collecting data is "the more it knows about you, the more likely it will be able to predict what you want to buy."<sup>53</sup>
- Banking Sector: Big data analysis is used in the banking sector for numerous purposes like collecting insights on the spending patterns, customer profiling, security and fraud management, etc.<sup>54</sup> It is also useful for studying the economy, studying past data and predicting future trends. It also helping banks in managing costs of compliance and risks of non-compliance.<sup>55</sup> However, it is said that the potential of big data analysis in the banking and finance industry is still unrealized.<sup>56</sup>
- Health Sector: The use of big data analytics in the healthcare sector has increased tremendously. According to the Study on Big Data in Public Health, Telemedicine and Healthcare of the European Commission, big data can help in: (i) increasing early diagnosis and improving quality of treatment by detecting early signals of disease, (ii) preventing diseases by identifying risks for disease, (iii) improve pharmacovigilance and patient safety

---

<sup>51</sup> Chunlei Tang, 'Data Services In Distinct Sectors', in *The Data Industry: The Business and Economics of Information and Big Data* (John Wiley & Sons 2016).

<sup>52</sup> Grace Bediako, 'The Application of Big Data Analytics in Improving eCommerce Processes: The Retail Sector User Experience' (Bachelor's thesis, Degree Programme in Computer Applications, Autumn 2023) <<https://www.theseus.fi/handle/10024/812302>> accessed 13 December 2024.

<sup>53</sup> Bernard Marr, *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results* (Wiley 2016).

<sup>54</sup> Utkarsh Srivastava, Santosh Gopalkrishnan, 'Impact of Big Data Analytics on Banking Sector: Learning for Indian Banks' (2015) 50 *Procedia Computer Science* 643.

<sup>55</sup> Rahul More and Yash Moily, 'Big Data Analysis in Banking Sector' (2021) 11(2) *International Journal of Engineering Research and Applications* 1.

<sup>56</sup> *ibid.*

through informed medical decisions and, (iv) predict outcomes.<sup>57</sup> Further, it can also help in public health surveillance and assessment.<sup>58</sup>

- Oil Industry: While data is being equated to oil, big data analytics have also been helping the oil industry. For instance, the Shell Company is using data analytics and comparing data collected from numerous drilling sites across the world. While comparing this data, the Company analyses if it resembles the profiles of drilling sites where abundant resources have been found previously. This helps in forecasting oil reserves.<sup>59</sup> Big data also helps oil companies in monitoring the performance and condition of their equipment, thereby lowering overhead costs by allowing effective routine maintenance.<sup>60</sup>
- Small Industries: Big data not only helps big companies but also small companies. A company called The Pendletons can be looked at for an example. The Pendletons used big data analytics to understand the behavior of local customers and gain insights to improve their business strategies.
- Sports and Big Data: Big data is also increasingly used for sports analytics. The US women's cycling team for the 2012 olympics followed the OAthlete (optimized athlete) approach wherein every aspect of an athlete's performance were monitored to amend their training programmes and optimize their performance sustainably.<sup>61</sup>
- Wildlife Conservation: The use of big data analytics is also helping in monitoring wildlife and tracking wildlife to fight against extinction. ZSL along with NASA and the European Commission's Joint Research Council was the first to highlight that remote sensing can help in understanding the impact of human activities on wildlife. It can also track

---

<sup>57</sup> Directorate-General for Health and Food Safety (European Commission) and others, *Study on Big Data in Public Health, Telemedicine and Healthcare: Executive Summary* (Publications Office of the European Union 2016) <<https://data.europa.eu/doi/10.2875/61543>> accessed 11 December 2024.

<sup>58</sup> Roberta Pastorino and others, 'Benefits and Challenges of Big Data in Healthcare: An Overview of the European Initiatives' (2019) 29 *European Journal of Public Health* 23.

<sup>59</sup> Bernard Marr, 'Big Data In Big Oil: How Shell Uses Analytics To Drive Business Success' (*Forbes*, 26 May 2015) <<https://www.forbes.com/sites/bernardmarr/2015/05/26/big-data-in-big-oil-how-shell-uses-analytics-to-drive-business-success/>> accessed 11 December 2024.

<sup>60</sup> Bernard Marr, *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results* (Wiley 2016).

<sup>61</sup> *ibid.*

movement and predict future movements of animals and detect whether there is a need for urgent human intervention to prevent loss to wildlife.<sup>62</sup>

Besides the few above-mentioned examples, big data analytics can also be useful in the manufacturing sector, shipping sector, real estate sector, etc.

### **3. Defining Personal and Non-Personal Data**

All the laws and regulations across different jurisdictions dealing with data, be it data protection or data sharing, are standing on the crossroads of personal and non-personal data. Hence, it is important to discern the meaning of both the terms.

Personal data, according to the General Data Protection Regulation (Regulation 2016/679), means “*any information relating to an identified or identifiable natural person*” and an identifiable natural person is one “*who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person*”.<sup>63</sup> According to this definition, personal data is data that can be identified to a natural person and can vary from objective like name, identification number, etc. to subjective information like opinions, evaluations, etc. The crucial element here is that personal information would describe something about a data subject. Data that does not have any meaning or is not related to any identifier is not personal data.

However, with the rapid development of technology it has become fairly easy to identify a person from non-personal data or re-identify a person from anonymised or pseudonymised data. The GDPR has addressed the concept of “*identifiability*” in Recital 26. It states that to determine identifiability of a person, all the means that are reasonably likely to be used to identify or single out an individual should be taken into account.<sup>64</sup> It further provides that in order to determine

---

<sup>62</sup> *ibid.*

<sup>63</sup> Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) [2016] OJ L119/1, art 4(1).

<sup>64</sup> Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) [2016] OJ L119/1, recital 26.

whether means are ‘reasonably likely’, objective factors like cost and amount of time required for re-identification and available technology and technological developments should be considered. The ISO has defined personally identifiable information (PII) as “*any information that can be used to establish a link between the information and the natural person to whom such information relates, or is or can be directly or indirectly linked to a natural person*”.<sup>65</sup>

The GDPR does not apply to anonymised data i.e. data where the data subject cannot be identified. The Directive also recognises that technological advancements should be given consideration while determining identifiability of data. This implies that the GDPR is taking account of the “technological context” and the same is crucial while defining and protecting personal data.<sup>66</sup>

The Free Flow of Non-Personal Data in the European Union (Regulation 2018/1807) has defined Non-Personal Data as “*data other than personal data as defined in point (1) of Article 4 of Regulation (EU) 2016/679*”. The Non-Personal Data Governance Framework has also defined non-personal data in a similar manner. It has defined non-personal data as data which is not personal data or data which does not have any personally identifiable information (PII).<sup>67</sup> By origin, non-personal data can either be non-identifiable to a person for example data related to weather, or data which was initially personal data but was made into non-personal data using techniques like anonymisation or pseudonymisation.<sup>68</sup> The definitions of personal data and non-personal data are mutually exclusive, interdependent and strongly connected.<sup>69</sup> Therefore, if pre-processing, a dataset is personal data, it can’t be said with certainty that it will maintain the same status after processing and this also applies to datasets that are non-personal data pre-processing.

---

<sup>65</sup> ISO/IEC 27701:2019, *Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines* (ISO 2019).

<sup>66</sup> Claudia Irti, 'Personal Data, Non-personal Data, Anonymised Data, Pseudonymised Data, De-identified Data' in *Privacy and Data Protection in Software Services* (Springer 2021) 49-57.

<sup>67</sup> Committee of Experts on Non-Personal Data Governance Framework, *Report by the Committee of Experts on Non-Personal Data Governance Framework* (Ministry of Electronics & Information Technology, 2020) [https://www.meity.gov.in/writereaddata/files/Non-Personal\\_Data\\_Committee\\_Report.pdf](https://www.meity.gov.in/writereaddata/files/Non-Personal_Data_Committee_Report.pdf) accessed 12 December 2024.

<sup>68</sup> European Commission, *Guidance on the Regulation on a Framework for the Free Flow of Non-Personal Data in the European Union* (Communication) COM(2019) 250 final.

<sup>69</sup> Emanuela Podda and Monica Palmirani, ‘Inferring the Meaning of Non-Personal, Anonymized, and Anonymous Data’ in Víctor Rodríguez-Doncel and others (eds), *AI Approaches to the Complexity of Legal Systems XI-XII* (Lecture Notes in Computer Science, vol 13048, Springer 2021) [https://doi.org/10.1007/978-3-030-89811-3\\_19](https://doi.org/10.1007/978-3-030-89811-3_19) accessed 12 December 2024.

There is lack of legal uncertainty regarding the definitions of both personal data and non-personal data and may be termed as obsolete definitions.<sup>70</sup>

The Expert Committee Report on Non-Personal Data Governance Framework has defined non-personal data as “*Any data which is not personal data (data pertaining to characteristics, traits or attributes of identity, which can be used to identify an individual)*” and also includes anonymised data. The Report further classifies non-personal data into three categories,: “(i) Public non-personal data, (ii) Community non-personal data, and, (iii) Private non-personal data.” The report has been criticized for the provided definition. The definition creates issues in distinguishing personal data from non-personal data as the PDP Bill (now DPDP Act) uses a context-based approach to define personal data wherein, depending on the context, data can be categorized as either personal or non-personal. Therefore, there is a lack of clarity as to what is non-personal data. Not only has the NPD Report adopted the definitions from EU that pose definitional problems, it has also failed to properly define the three categories of NPD.

The overlap between personal data and non-personal can also raise ethical and privacy concerns. The advancement of big data and the development of new technologies has made it possible to re-identify anonymous datasets. And the notion that once personally identifiable information is removed from a dataset, it becomes impossible to identify an individual has no scientific basis.<sup>71</sup> Anonymised datasets can often pose identification risks to data subjects, if not from the dataset itself then from the linking of the dataset with other sources of information available publicly and once datasets are linked, they can't be unlinked.<sup>72</sup> For example, in 2018, Strava, a GPS company released a heat map which indicated where people were exercising. The data released showed where people tracked and went for bike rides the most. This data made it possible to figure out that there are US army troops in Somalia, Afghanistan & Syria and Russian troops were revealed to be

---

<sup>70</sup> *ibid.*

<sup>71</sup> Henry Pearce, 'Big Data and the Reform of the European Data Protection Framework: An Overview of Potential Concerns Associated with Proposals for Risk Management-Based Approaches to the Concept of Personal Data' (2017) 26(3) *Information & Communications Technology Law* 312-335 <<https://doi.org/10.1080/13600834.2017.1375237>> accessed 12 January 2025.

<sup>72</sup> Mark Elliot, Kieron O'Hara, Charles Raab, Christine M. O'Keefe, Elaine Mackey, Chris Dibben, Heather Gowans, Kingsley Purdam, Karen McCullagh, 'Functional Anonymisation: Personal Data and the Data Environment' (2018) 34(2) *Computer Law & Security Review* 204-221 <<https://doi.org/10.1016/j.clsr.2018.02.001>> accessed 12 January 2025 .

present in Ukraine and Taiwan. The aggregated data revealed the patterns of and around a military base and the same posed security risks as the information revealed was of an extremely sensitive nature.<sup>73</sup> The risks posed in this case were more related to security and threat and less with privacy. However, this raises a concern regarding the notion of privacy and whether it extends to identification of communities/groups or is limited to identifications of individuals.

Mixed data sets raise further questions regarding parallel application of laws dealing with personal and non-personal data. Although it is possible to apply the two laws in parallel, in practice disentangling the two types of data in order to comply with both laws becomes difficult.<sup>74</sup> As per Article 2(b) of the Free Flow of Non-Personal Data Regulation provides that in cases of a data set comprising both personal and non-personal data, the Regulation applies to the non-personal data part of the data set and where personal and non-personal data in a data set are inextricably linked, the Regulation shall not prejudice the application of the GDPR (EU) 2016/679.<sup>75</sup> However, the Directive has not clarified what is meant by 'inextricably linked'. The term could imply that it is not possible to disentangle or differentiate personal from non-personal data or it could be a reference to the dynamic nature of data and the context based approach.<sup>76</sup> In addition to this, difficulty also arises in determining exactly when a dataset becomes a mixed dataset and requires compliance with both personal and non-personal data regulations. In other words, when the nature of a dataset changes from personal to non-personal or vice versa subsequently, it is not easy to determine when the switch happened.<sup>77</sup>

#### **4. Development of Data Protection and Governance Laws in India**

The rapid technological revolution and the increased digitisation of international trade acted as a push for the UNCITRAL to introduce the Model Law on Electronic Commerce in 1996. The Model

---

<sup>73</sup> Andrew Moseman, 'U.S. Troops Accidentally Reveal Secret Bases by Going Jogging' *Popular Mechanics* (29 January 2018) <<https://www.popularmechanics.com/technology/apps/a15912407/strava-app-military-bases-fitbit-jogging/>> accessed 31 January 2025.

<sup>74</sup> Inge Graef, Raphael Gellert and Martin Husovec, 'Towards a Holistic Regulatory Approach for the European Data Economy: Why the Illusive Notion of Non-Personal Data is Counterproductive to Data Innovation' (September 27, 2018). TILEC Discussion Paper No 2018-029 <<https://ssrn.com/abstract=3256189>> accessed 11 January 2025.

<sup>75</sup> Free Flow of Non-Personal Data Regulation, art 2(b).

<sup>76</sup> Inge Graef, Raphael Gellert and Martin Husovec (n 74).

<sup>77</sup> *ibid.*

Law was introduced with the objective of facilitating e-commerce transactions and for providing internationally acceptable rules on e-commerce and foster efficiency in international trade.<sup>78</sup> The Model Law operated on three principles i.e. non-discrimination, technological neutrality and functional equivalence. Many countries including India adopted the Model Law in order to fulfill its international obligations and commitments. The Information Technology Act was to provide “legal recognition for e-commerce. The Information Technology Act and the IT Rules provided for certain provisions for the protection of privacy and data protection in India. Section 43 of the IT Act enumerates and penalizes certain acts with respect to computer and computer resources like accessing computer resources without permission, extracting data, replicating information, contaminating computers or introducing viruses, destroying, deleting or altering information residing in the computer. Further, section 43A of the Act provides protection of sensitive personal information. It states that if a company possessing, dealing or handling sensitive personal data, causes wrongful loss or gain to any person due to its negligence in maintaining security measures to protect the data will be held liable under the Act.<sup>79</sup> The IT Act is supported by the Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules 2011 (Privacy Rules).<sup>80</sup> The Privacy Rules define sensitive personal data or information (SDPI) as personal information that contains information related to passwords, financial information (like bank account, credit card, debit card), physical, physiological and mental health condition, sexual orientation, medical records and history, Biometric information and any related detail. The IT Act and the Privacy Rules impose certain obligations on body corporates to ensure that data processing is done properly. It provides that reasonable security practices, procedures and standards must be implemented while handling sensitive personal data.<sup>81</sup> It also provides that a company can only collect sensitive personal data with prior consent of the data subject and only when it is essential and is done lawfully. Sensitive data must also not be stored or retained for longer than necessary and shall only use it for the purpose it was collected for.<sup>82</sup>

---

<sup>78</sup> “UNCITRAL Model Law on Electronic Commerce (1996) with additional article 5 bis as adopted in 1998” <[https://uncitral.un.org/en/texts/ecommerce/modellaw/electronic\\_commerce](https://uncitral.un.org/en/texts/ecommerce/modellaw/electronic_commerce)> accessed 30 January 2025.

<sup>79</sup> Information Technology Act 2000, s 43.

<sup>80</sup> Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules 2011 (Privacy Rules).

<sup>81</sup> *ibid*, rule 8.

<sup>82</sup> *ibid*, rule 5.

In 2017, the Puttaswamy judgment set a landmark precedent holding that privacy is intrinsic to freedom, liberty and dignity and is an inherent element of Article 21 of the Indian Constitution.<sup>83</sup> Following the Puttaswamy judgment, an expert committee was set up under the chairmanship of Justice B.N. Sri Krishna to address the issues pertaining to privacy and data protection which submitted the draft Data Protection Bill in 2018. The bill was tabled in Rajya Sabha and then sent to the Joint Committee of Parliament (JPC) for review in 2019. The review by the JPC was sent in the year 2021 and was heavily criticized being very state centric.<sup>84</sup> The criticisms on the bill surprisingly came from within the JPC. While the draft Personal Data Protection bill was under review, another expert committee was set up by the Ministry of Electronics and Information Technology under the chairmanship of Mr. Kris Gopalan to address the issues related to non-personal data (NPD). The Committee noted that NPD should be regulated to promote transparency and data sharing.<sup>85</sup> It further suggested that the Personal Data Protection Bill should be amended to include provisions dealing with NPD. While both the drafts on PD and NPD were under review, the MEITY introduced the draft India Data Accessibility and Usage Policy in February 2022 with the objective to maximize access, sharing and use of non-personal data and to “transform India’s ability to harness public sector data for catalyzing large scale social transformation”.<sup>86</sup> The policy also aims to enhance interoperability of data and improve the quality of data. It is one of the first policies in the country to focus on the commercial and economic value of public sector data and how it can enhance businesses.

The latest step towards data protection in India is the Digital Data Protection Act which was passed in August 2023 to provide for processing of personal data. The Act recognises an individual's right to protect their personal data and the need to process personal data for lawful purposes. The Act is the first law in the country dealing with protection of personal data across sectors. It defines

---

<sup>83</sup> *K S Puttaswamy v Union of India* (2017) 10 SCC 1.

<sup>84</sup> Shruti Dvivedi Sodhi , Bansari Samant and Tushar Sinha, ‘The Journey of India’s Data Protection Jurisprudence - Lexology’ (*Lexology*, 11 April 2022) <<https://www.lexology.com/library/detail.aspx?g=57720842-f709-4dd4-947b-44c3c6e4ed10>> accessed 12 December 2024.

<sup>85</sup> Summary of the Personal Data Protection Bill 2019’ (*PRS India*, 2024) <<https://prsindia.org/policy/report-summaries>> accessed 12 December 2024.

<sup>86</sup> Draft India Data Accessibility & Use Policy 2022 <<https://www.meity.gov.in/content/draft-india-data-accessibility-use-policy-2022>> accessed 12 December 2024.

personal data as “any data about an individual who is identifiable by or in relation to such data”.<sup>87</sup> The Act allows processing of personal data for lawful purposes i.e. if there is consent of the data principal or if it is for legitimate purposes.<sup>88</sup>

Since the 2000s, the government has made multiple attempts to regulate data usage and data sharing in India. However, it is unlikely that the question regarding the comprehensiveness and effectiveness of data protection laws in India can be answered in the positive.

## **5. Global Perspectives on Non-Personal Data Governance**

This section briefly evaluates how different jurisdictions have made attempts to deal with data governance and non-personal data. This would provide an overview of the global position on NPD which can provide useful insights in fostering a framework for NPD governance.

In 2022, Australia enacted the Data Availability and Transparency Act that authorized the sharing of public sector data. The Act has defined ‘information’ as “any information in a form capable of being communicated, analyzed or processed (whether by an individual or by computer or other automated means)”.<sup>89</sup> The objective of this Act was to enhance data availability, enable sharing of public sector data that would be in compliance with the Australian Privacy Act of 1988, enhance integrity and transparency in data sharing, confidence building in public sector data and to establish institutional mechanisms for data sharing.<sup>90</sup> The Act also establishes a DATA scheme which would be regulated by the National Data Commissioner. The Act also provides for purposes of data sharing i.e. delivery of government services, informing government policy and programs, and research and development.<sup>91</sup>

---

<sup>87</sup> Digital Personal Data Protection Act, 2023, s 2(t).

<sup>88</sup> *ibid*, s 4.

<sup>89</sup> 'Data Availability and Transparency Act and Commonwealth Records' (National Archives of Australia) <<https://www.naa.gov.au/information-management/information-management-legislation/data-availability-and-transparency-act-and-commonwealth-records>> accessed 12 December 2024.

<sup>90</sup> Home Page' (Office of the National Data Commissioner) <<https://www.datacommissioner.gov.au/>> accessed 12 December 2024.

<sup>91</sup> Digital Personal Data Protection Act, 2023, s 15.

In Japan, the Basic Act on the Advancement of Public and Private Sector Data Utilization was introduced with the purpose to effectively utilize and appropriate large amounts of data generated by the internet in order to achieve outcomes like enhancing safety and development of the society and the citizens. The Act defined public and private sector data as “information (excluding information that is likely to damage national security, hinder the maintenance of public order, or be an obstacle to the protection of public safety) recorded in an electronic or magnetic record”.<sup>92</sup> It lays down the principles for effective use and smooth circulation of public sector data.<sup>93</sup>

In 2014, the Canadian Government launched an Open Government Portal to facilitate access to information to enhance accountability and transparency.<sup>94</sup> The objective for the launch of the portal was to provide a one-stop access to information to government data and information to the public and businesses in order to increase transparency, accountability, citizen engagement, and socio-economic benefits. Subsequently in 2019, the Standards Council of Canada introduced the Canadian Governance Standardization Collaborative with the objective to accelerate industry wide data governance standardization strategies. The result of this collaboration was a standardization roadmap that made 35 recommendations for data governance, data collection, organization & grading, data access & sharing and data analytics & commercialisation.<sup>95</sup>

The UK in 2020 introduced its National Data Strategy with an overall objective of unlocking the power of data and to set out a framework and an approach for investing in data and strengthening the economy. The Strategy also focuses on ensuring data accessibility and states that it is important for the government to remove all the unnecessary access barriers to data. Furthermore, the UK also has a Data Ethics Framework which provides responsible use of data and which acts as a guidance for law and policymakers. It focuses on principles of transparency, accountability and fairness.<sup>96</sup>

---

<sup>92</sup> Basic Act on the Promotion of Utilization of Public and Private Sector Data (Japan, 2016) <[https://japan.kantei.go.jp/policy/it/data\\_basicact/data\\_basicact.htm](https://japan.kantei.go.jp/policy/it/data_basicact/data_basicact.htm)> accessed 12 December 2024.

<sup>93</sup> *ibid.*

<sup>94</sup> ‘Open Government Portal’ (Government of Canada) <<https://open.canada.ca/en>> accessed 12 December 2024.

<sup>95</sup> Public Services and Procurement Canada Government of Canada, ‘Canadian Data Governance Standardization Roadmap.: Iu81-3/18-2021E-PDF - Government of Canada Publications - Canada.Ca’ (2021) <<https://publications.gc.ca/site/eng/9.906188/publication.html>> accessed 12 December 2024.

<sup>96</sup> ‘Data Ethics Framework’ (GOV.UK, 16 September 2020) <<https://www.gov.uk/government/publications/data-ethics-framework>> accessed 12 December 2024.

When it comes to laws and regulations governing data, the European Union has been at the forefront. In 2018, the EU introduced a 'Framework for the Free Flow of Non-Personal Data' with the aim to ensure free flow of data (other than non-personal data) by laying down requirements like data portability and data localisation. The regulation aimed to address the issues regarding market distortion and remove any obstacles to trade in the market that distort competition and the internal functioning of the market.<sup>97</sup> The Regulation was also introduced with the objective to harmonize different national laws and regulations across the EU's jurisdiction. In addition, the EU Data Act, 2022 was also introduced to harmonize laws on data use and data access across the EU. The Act seeks to ensure fairness in the use of data generated by the Internet of Things devices. The objective of the Act is to make more data available through measures like increasing legal certainty and providing incentives to entities for investment in generation of data, develop model contract clauses for data sharing to prevent abuse of contractual imbalances, provide access to private sector data to public sector data for public interest purposes and make new rules for data portability.<sup>98</sup>

The exploration of big data and its regulatory landscape underscores a critical intersection of technology, law, and economics in the digital age. As big data continues to be a catalyst for innovation, its regulation poses unique challenges and opportunities for policymakers, businesses, and society at large.

Big data's transformative potential lies in its ability to drive innovation, improve decision-making, and create new business models. Its volume, velocity, and variety enable organizations to derive valuable insights and gain a competitive edge over their rivals. In the digital economy, big data is often termed as the "new oil" due to its ability to drive innovation. However, the analogy does not go beyond metaphors as it often overlooks the significant differences between data and oil. The same attributes that make big data valuable also introduce complexities in its regulation. Due to the benefits and advantages big data can offer, its access has become a debatable issue. The challenge is to strike a balance between fostering innovation and implementing robust regulatory

---

<sup>97</sup> European Commission, 'Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union' (Communication) COM(2019) 250 final.

<sup>98</sup> 'Data Act' (European Commission, Shaping Europe's Digital Future) <[https://ec.europa.eu/digital-strategy/our-policies/data-act\\_en](https://ec.europa.eu/digital-strategy/our-policies/data-act_en)> accessed 12 December 2024.

frameworks that protect data and information privacy, ensure fair competition, and also safeguard public interest.

## CHAPTER II: LEGAL PROTECTION FOR NON-PERSONAL DATA

### 1. Perspectives on Legal Scope and Definition of Data:

As discussed in the previous chapter, there are several definitions of the term “Big Data” and the most used formulation is provided by Laney which focuses on the three Vs of Big Data. The legal meaning of the term “Data” or “Big Data” has remained uncertain. Big Data is said to refer to a “*a phenomenon with a multitude of different implications in scientific disciplines, such as economics, technical disciplines, legal and social science*”. To understand the legal implications of Big Data, it is not crucial to have a precise legal definition. Rather, what is important is to understand the big data ecosystem.<sup>99</sup> Big data business models can essentially be categorized into three distinct categories, i.e., data collection/acquisition, data processing and data analytics. Each of these stages of big data interacts with law. The first stage is the process of data collection and data acquisition wherein issues concerning the ownership of data arise. Data acquisition also raises concerns regarding contractual relationships when data is being traded like a commodity. The second stage is the processing of data which can raise issues in relation to data protection laws if processing of data involves personal data or mixed datasets. The third stage is big data analysis, i.e. using the data, which has also raised several issues.<sup>100</sup> For example, access-related issues.

One of the most important issues that arises in the context of data is the ownership of data. The emergence of ownership rights can be traced back to different theories of ownership (Kant, Hegel, Bentham, Rawls). Different perceptions of ownership can also be seen in the context of data ownership. The dominant perspective in legal theory however has been that data cannot be owed since in many jurisdictions, right in facts and information has not been recognised.<sup>101</sup> Scholars have been critical of data ownership since there are significant differences between “data” and “regular” tangible property. Firstly, material tangible goods are inherently rivalrous in nature i.e.

---

<sup>99</sup> Nikolaus Forgó, Stefanie Hänold and Benjamin Schütze, ‘The Principle of Purpose Limitation and Big Data’, in Marcelo Corrales, Mark Fenwick, Nikolaus Forgó (eds), *New Technology, Big Data and the Law* (1st edn , Springer 2017).

<sup>100</sup> *ibid.*

<sup>101</sup> Patrik Hummel, Matthias Braun and Peter Dabrock, ‘Own Data? Ethical Reflections on Data Ownership’ (2020) 34 *Philosophy & Technology* 545 < <https://doi.org/10.1007/s13347-020-00404-9>> accessed 11 December 2024.

they cannot be used simultaneously by multiple parties at the same time. The rivalrous nature of tangible property makes it easy to claim exclusivity. On the other hand, data is an intangible and a non-rivalrous good that can be simultaneously used by many without losing its value or getting depleted.<sup>102</sup> This non-rivalrous nature of data makes it difficult to claim exclusivity and third parties cannot be excluded from using and accessing data, unless it is being kept a secret.<sup>103</sup> Non-rivalrous goods can be made excludable through legal or technical intervention like intellectual property rights.

Second, the proponents of ownership rights in data argue that in the current digital era of the internet of things (IoT), data holds tremendous economic value, and the existing legal regime of intellectual property and contract law fails to provide ideal protection to data. The existing legal regime does not create *erga omnes* rights in data and thereby there is always a risk of data misappropriation which can withhold data holders from confidently using and freely trading data.<sup>104</sup> This can create antitrust issues in the digital market. However, on the other hand, it is argued that the argument for data ownership to solve the incentive problem fails and there is no market failure that needs to be addressed. This is because the creation of data is generally a “by-product of profitable economic activities and does not require additional incentives”.<sup>105</sup>

The law regarding ownership in data isn't clear. Some legal instruments like GDPR and the EU Database Directive have created rights in data but only to some extent and therefore data holders protect their data by restricting access through various legal and technical means. The proponents of “virtual property” and ownership rights argue that despite the non-rivalrous nature of data there can be distinct legal mechanisms that can be used to protect ownership rights in data. For example, domain names, URL, email accounts, websites, etc do “mimic physical properties of tangible goods”. Furthermore, proponents also argue that the notion of property is open and has a functional

---

<sup>102</sup> Néstor Duch-Brown, Bertin Martens and Frank Mueller-Langer, ‘The Economics of Ownership, Access and Trade in Digital Data’ (2017) JRC Digital Economy Working Paper 2017-01 <<http://dx.doi.org/10.2139/ssrn.2914144>> accessed 14 December 2024.

<sup>103</sup> *ibid* 22.

<sup>104</sup> P Bernt Hugenholtz, ‘Data Property: Unwelcome Guest in the House of IP’ (Paper Presented at Trading Data in the Digital Economy: Legal Concepts and Tools, Münster, Germany, 2017) <[https://pure.uva.nl/ws/files/16856245/Data\\_property\\_Muenster.pdf](https://pure.uva.nl/ws/files/16856245/Data_property_Muenster.pdf)> accessed 14 December 2024.

<sup>105</sup> *ibid*.

role.<sup>106</sup> The functional role of property is that it is inherently connected to the property owner's freedom and even if data is not a material good, it can be protected due to its functional role.

Further, while examining the question of data or information ownership, Heverly argues that it must be analyzed how the extent of ownership right in data is comparable to existing tangible property rights. Heverly analyzes the extent of ownership in the context of right to use, right to exclude and right to transfer.<sup>107</sup> Furthermore, if data or information is protected, it is also crucial to note that it can be protected at different levels i.e. syntactic (code level), semantic (meaning of the data), structural (physical embodiment of information) and pragmatic level (effect, use & purpose of information) and there is a need to determine which level or levels the ownership right would govern.<sup>108</sup>

The non-rivalrous and non-excludable nature of data, in economic terms, makes it a public good. According to Zech, the possession of data should not be equated with ownership. Rather, for data and information, the right term is "access".<sup>109</sup> Van Alstyne has also made a distinction between "usage rights in data" and "ownership in data". Usage rights determine the "ability to access, create, standardize and modify data" and data ownership is the "right to determine these privileges for others".<sup>110</sup> Data holders enjoy *de facto* ownership of data and while there is a lack of clarity regarding data ownership. Data holders rely mainly on three barriers that prevent others from *accessing* their data. These barriers are legal, behavioral and technological.<sup>111</sup> Legal barriers to access are in the form of intellectual property protection (discussed at length in the subsequent sections). Behavioral barriers emerge in the context of contractual limitations. For example, exclusionary practices, unequal bargaining power between parties, restrictive clauses etc. Technological barriers include reliance on advanced technology and implementation of technical

---

<sup>106</sup> *ibid* 22.

<sup>107</sup> Robert A Heverly, 'The Information Semicommons' (2003) 18(4) Berkeley Technology Law Journal <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=450280](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=450280)> accessed 14 December 2024.

<sup>108</sup> *ibid* 22.

<sup>109</sup> *ibid* 22.

<sup>110</sup> Martin Fadler and Christine Legner, 'Who Owns Data in an Enterprise? Rethinking Data Ownership in Times of Big Data and Analytics (2020), Twenty-Eight European Conference on Information Systems (ECIS2020).

<sup>111</sup> Tomasso Fia, 'An Alternative to Data Ownership: Managing Access to Non-Personal Data through the Commons' (2020) Global Jurist <<https://www.degruyter.com/document/doi/10.1515/gj-2020-0034/html>> accessed 14 December 2024.

protection measures for access management. This *de facto* ownership raises two problems. First, unequal bargaining power between the negotiating parties and second, issues pertaining to innovation as data holders can retain data and data driven innovation for their own interest and have no duty or incentive to share.<sup>112</sup>

In the context of lack of legal clarity regarding ownership rights in data, data holders protect their interest through various means. The following sections of the chapter discuss elaborately on the same.

## **2. Data and Intellectual Property Protection**

### **2.1. Understanding the Intersection of Intellectual Property and Data**

Intellectual properties like inventions, literary and artistic works, designs, trademarks, etc, are creations of mind. Intellectual property grants creators an exclusive right and control over their creations. It provides creators with an economic incentive for creating their work and, therefore, solves the public good problem by providing benefits of the creation to the public.<sup>113</sup> IP rights also incentivize research and development by allowing creators to commercialize their work.

The ever-increasing value of data in the digital economy has raised questions regarding data ownership and whether data should be given intellectual property protection. The proponents for IP protection for data argue that the value of data is increasing significantly, and the current laws do not protect data adequately. Therefore, for an efficient data market, it is crucial to give IP protection to solve the incentive and the public good problem.<sup>114</sup> On the other hand, it is argued that there is no evidence to confirm an incentive problem in data markets. Protecting data would be complicated since it is difficult to determine the subject matter and the scope of data. Moreover,

---

<sup>112</sup> *ibid*, n 24.

<sup>113</sup> Stanley M Besen, 'An Introduction to the Law and Economics of Intellectual Property' (1991) 5(1) *The Journal of Economic Perspectives* 3-27 .

<sup>114</sup> Wolfgang Keber, 'A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis, (2016) *erwerblicher Rechtsschutz und Urheberrecht, Internationaler Teil (GRUR Int)* < [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2858171](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2858171)> accessed 11 December 2024.

an IP right over data could negatively impact the market and competition as it would affect the free movement of data in the economy.

Data protection in the IP regime is a complicated affair and is under debate. This chapter provides an overview of data protection under the current intellectual property regime. It focuses on copyright protection for databases, data producer's rights, unfair competition vis-a-vis trade secrets and protection of data under contract law.

## **2.2. Copyright Protection for Original Databases**

Copyright is an author's right over their original works. It grants authors a bundle of rights for fixed time duration. These rights include the right to reproduce, distribute and display work in public. The rationale behind granting copyrights to authors is to promote creativity and encourage the progress of science and arts.<sup>115</sup>

One of the most essential principles of Copyright law is the idea-expression dichotomy. This principle distinguishes ideas from expressions and asserts that copyright law protects expressions, not ideas. The principal aim is to promote innovation and competition by ensuring that ideas are freely available, and protection is granted only to the particularised expressions of these ideas.

For data, copyright protection can be granted to databases that are original in their selection and arrangement. Internationally, three documents can be referred to for the protection of collections and compilations. The Berne Convention for the Protection of Literary and Artistic Works protects original collections of literary and artistic works but does not explicitly provide for the protection of databases.<sup>116</sup> The TRIPS Agreement protects the compilation of data wherein protection is granted to the selection and arrangement of data and not the data itself.<sup>117</sup> Lastly, the WIPO Copyright Treaty, 1996 (WCT) also provides a provision similar to TRIPS wherein it grants

---

<sup>115</sup> WIPO, Copyright & related rights <[https://www.wipo.int/export/sites/www/sme/en/documents/pdf/ip\\_panorama\\_5\\_learning\\_points.pdf](https://www.wipo.int/export/sites/www/sme/en/documents/pdf/ip_panorama_5_learning_points.pdf)> accessed 19 December 2024.

<sup>116</sup> Berne Convention for the Protection of Literary and Artistic Works, 1886, .

<sup>117</sup> Agreement on Trade-Related Aspects of Intellectual Property Rights , 1994, Art 10.

protection to compilations of data which, because of their selection and arrangement, constitute intellectual creations.<sup>118</sup>

Copyright is granted to original works, and the concept of originality is based on multiple doctrines across jurisdictions. One of these doctrines is based on Locke's labor theory, wherein copyright protection is granted for the labor put into the creation of the work. Lockean theory suggests that an author should have the rights over the fruits of their intellectual labor.<sup>119</sup> This right originates from a person's right on their own body and thus, anything that their body creates, belongs to them and cannot be separated from them.<sup>120</sup> The author's effort, skill and time were considered sufficient to satisfy the originality requirement for copyright protection.<sup>121</sup> This is also known as the 'sweat of the brow' doctrine. Thus, copyright was granted to the author's independent creations and the work "originated" from the author must not be copied.<sup>122</sup> The question of originality is one of the most important concerns in copyright law and has been under judicial scrutiny time and time again. The Feist Publications case laid down the principle that a minimal level of creativity is a prerequisite for originality.<sup>123</sup> The Court noted, "*Compilations would be considered original if the author's selection and arrangement is done independently and entail minimal creativity.*"

In the Canadian case of CCH Canadian Ltd, the Supreme Court held that Originality is not to be interpreted as novelty or uniqueness. Rather, work should originate from the author themselves and it was further noted that the exercise of skill and labor should not be merely mechanical or negligible. It must involve the exercise of judgment and skill that involves intellectual effort i.e. "*use of one's knowledge, developed aptitude or practiced ability in producing the work.*"<sup>124</sup>

---

<sup>118</sup> WIPO Copyright Treaty 1996, Art 5.

<sup>119</sup> John Locke, Two Treatises of Government, Second Treatise (3rd edn, Cambridge University Press 1988) p 287- 288 "*Though the earth, and all inferior creatures be common to all men, yet every man has a property in his own person. This nobody has any right to but himself. The labour of his body and the work of his hands, we may say, are properly his. Whatsoever, then, he removes out of the state that nature hath provided and left it in, he hath mixed his labour with it and joined to it something that is his own, and thereby makes it his property...*"

<sup>120</sup> Edwin C Hettinger, 'Justifying Intellectual Property' (1989) 18(1) Philosophy & Public Affairs 31-52 <<https://www.jstor.org/stable/2265190>> accessed on 11 December 2024.

<sup>121</sup> Walter v Lane [1900] A C 539, Jeweler's Circular Pub Co v Keystone Pub Co, 281 F 83.

<sup>122</sup> University of London Press Ltd v University Tutorial Press Ltd [1916] 2 Ch 601.

<sup>123</sup> Feist Publications v Rural Telephone 499 U S 340 (1991).

<sup>124</sup> CCH Canadian Ltd v. Law Society of Upper Canada 2004 SCC 13.

India, in compliance with its international obligations under the TRIPS Agreement and the Berne Convention, protects databases that are original in the manner of their selection and arrangement. According to section 2(o) of the Copyright Act, literary work includes databases. The sweat of the brow doctrine was followed in India for a while. In *Burlington Home Shopping Pvt. Ltd. v Rajnish Chibber & Anr*, one of the first cases dealing with databases and copyright in India, the Court granted protection. While recognising the sweat of the brow doctrine, the Court noted that the author's hard work, labor and time invested in creating the work would be protectable as an original work.<sup>125</sup> Later, in the *Eastern Book Company v. D.B. Modak*, the Supreme Court rejected the sweat of the brow theory, and the Court's approach shifted to the modicum of creativity. The Court held that to get copyright protection over compilations and databases, some amount of creativity, exercise of skill and judgment must be involved.<sup>126</sup>

### **2.3. The EU Database Directive**

The originality threshold has developed through precedents over the years. This has led to divergence in the originality threshold required to be fulfilled across jurisdictions to get copyright protection. The European Union introduced the EU Database Directive, intending to achieve uniformity and harmonization across the EU concerning the protection of databases and to prevent market distortion due to variations in law. The Directive is a significant development for the protection databases and aims to deal with the non-uniformity in intellectual property protection for databases across the EU. The Recital of the directive notes that the protection for databases is uneven across the EU and the same is detrimental to the functioning of internal markets which can affect the free movement of goods and services in the EU.<sup>127</sup> The Directive recognises databases are crucial for development of information market and unauthorized use and/extraction of databases can have severe consequences.

Under the Directive, a database means “*a collection of independent works, data or other materials arranged systematically or methodically and individually accessible by electronic or other*

---

<sup>125</sup> *Burlington Home Shopping Pvt Ltd v Rajnish Chibber & Anr* 1995 PTC (15) 278.

<sup>126</sup> *Eastern Book Company v D B Modak*, 2002 SCC OnLine Del 1131, (2002) 25 PTC 641.

<sup>127</sup> Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases [1996] OJ L77/20.

means.”<sup>128</sup> The Directive grants copyright protection of 70 years to databases that are the author's own creation and are original in selecting and arranging contents.<sup>129</sup> It further clarifies that no other criteria than the one provided in the Directive shall be applied to determine whether protection should be granted to a database. The Directive also provides for exceptions to copyright. The limitations are applicable in cases where reproduction is for private purposes, for teaching or scientific research purposes to the extent that it is non-commercial, for purposes of public security and other traditional exceptions to copyright applicable in member states.<sup>130</sup>

Original databases are protected under the copyright law provided they are original in the selection and arrangement of their contents. Copyright cannot protect non-original databases or data as they do not fulfill the minimum level of creativity required, and copyright law does not protect facts.

### **2.3.1. EU’s Unique Approach: Sui Generis Protection for Non-Original Databases**

The EU Database Directive grants rights to the database maker where it is shown that “*substantial investment has been made in obtaining, verifying or presenting the contents of the database*”.<sup>131</sup> Thus, the Directive creates a *sui generis* right in the contents of databases. Data does not qualify for copyright protection in its raw form. However, the information that databases contain is a product of substantial investment, effort and time; thus, the need to protect such databases was recognised.<sup>132</sup> The objective behind the right is to grant database makers a right to prevent unauthorized extraction of the contents of the database in order to protect investment for a limited period of time. Thereby, providing an incentive to database makers to share the data instead of holding data in secrecy.<sup>133</sup>

---

<sup>128</sup> *ibid*, Art 1.

<sup>129</sup> *ibid*, Art 3.

<sup>130</sup> *ibid*, Art 6(2) .

<sup>131</sup> *ibid*, Art 7.

<sup>132</sup> Martin Zeitlin, Everything Counts in Large Amounts: Protection of Big Data under the Database Directive (Master’s Thesis, Department of Law, Spring Term 2018) <https://www.diva-portal.org/smash/get/diva2:1211909/FULLTEXT01.pdf> accessed 11 December 2024.

<sup>133</sup> Dev S Gangjee, 'The Data Producer's Right: An Instructive Obituary' (7 March 2022) in Ernest Lim and Phillip Morgan (eds), *The Cambridge Handbook of Private Law and Artificial Intelligence* (Cambridge University Press, 2022) <<https://ssrn.com/abstract=4051831> or <http://dx.doi.org/10.2139/ssrn.4051831>> accessed 11 December 2024.

The accrual of rights under the Directive does not require a database to be creative or original. The *sui generis* right is accrued on fulfilling the requirement of showing that substantial investment has been made in obtaining, verifying and presenting the contents of the database. As a result of this, databases which are not original are now granted protection.<sup>134</sup>

### **2.3.2. Scope and Impact of the Sui Generis right:**

The Directive grants two rights to the database maker. Firstly, the right to prevent extraction and the right to prevent re-utilization of the contents of the database. It further provides that extraction implies the “*permanent or temporary transfer of all or substantial parts of the contents of the database to another medium.*”<sup>135</sup> Further, re-utilization means “*making available to the public all or substantial parts of the contents of the database by the distribution of copies, by renting, by online or other forms of transmission.*”<sup>136</sup>

As per the Directive, lawful users of publicly available databases cannot be prevented from using non-substantial parts of the database. However, the Directive prohibits lawful users from performing acts that would disrupt the regular use of the database or unreasonably harm the legitimate interests of the database producer. Furthermore, the Directive also provides that lawful users of a database should not in any manner perform any acts that would come in conflict with the normal exploitation of the database or prejudice the legitimate interests of the database maker.<sup>137</sup>

Similar to the exceptions to the copyright protection, the Directive also provides exceptions for the *sui generis* right. The lawful users of the database may extract or reutilize substantial parts of the contents of the database for certain specific circumstances. Firstly, the users may extract data from non-electronic databases for private purposes. Secondly, the users may extract without authorization for educational or scientific research purposes, provided that it is for non-commercial

---

<sup>134</sup> P B Hugenholtz, 'Something Completely Different: Europe's Sui Generis Database Right' in S Frankel and D Gervais (eds), *The Internet and the Emerging Importance of New Forms of Intellectual Property* (Wolters Kluwer 2016) 205-222, Faculty of Law, Institute for Information Law (IViR).

<sup>135</sup> Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases [1996] OJ L77/20, Art 7(2a).

<sup>136</sup> *ibid*] Art 7(2b).

<sup>137</sup> *ibid* Art 8.

use. Lastly, extraction or re-utilization is permitted by public security, administrative or judicial procedures.<sup>138</sup>

The European Court noted that the database right protects the investment in obtaining the contents of the database. It was clarified that the “*investment made in obtaining the contents of the database*” implies the “*resources used to seek out existing materials and collect them in the database but does not cover the resources used for creating materials which make up the contents of a database.*” It does not protect the investments made in the creation or generation of new data or contents of the database.<sup>139</sup> It further clarified that the investment, qualitative or quantitative, in creating a database must be substantial. The objective behind the right is not to protect investment in the creation of data but to “promote the establishment of storage and processing systems for existing information.”<sup>140</sup>

### **2.3.3. Problems and Challenges with the Sui Generis Right**

The sui generis protection for non-original databases emerged as a legal framework to protect the investments and effort put into creating databases. The EU Directive has been criticized for posing several problems and challenges. Certain complexities surrounding the data producer’s right within the Directive have attracted backlash.

One of the key challenges relating to the sui generis right is access to databases. The Directive grants the database-maker the right to prevent extraction or re-utilization of the database contents. While this right aims to protect database makers, it also raises issues regarding equitable access to the database. It is also contended that the database producer’s right is a result of a fear of market failure and, in consequence, will cause multiple mini monopolies which would disrupt market efficiency.<sup>141</sup> The Directive has been criticized for not providing enough provisions to accommodate the needs of third parties who would want access to the database for the development of value added goods and services based on the databases. Compulsory licensing can be a potential

---

<sup>138</sup> *ibid* Art 9.

<sup>139</sup> *Fixtures Marketing Ltd v Oy Veikkaus Ab*, ECJ 9 November 2004, case C-46/02, ECR [2004] I-10396; Id. 23.

<sup>140</sup> *The British Horseracing Board Ltd and Others v William Hill Organization Ltd*. Case C-203/02.

<sup>141</sup> WIPO, *A Study on the Impact of Protection of Unoriginal Databases on Developing Countries: Indian Experience* (SCCR/7/5, 4 April 2002).

solution to this access problem. In fact, the proposal on the legal protection for databases did include a provision in this regard. Article 8 of the proposal provided for compulsory licensing on fair, reasonable and non-discriminatory terms.<sup>142</sup> Including a compulsory licensing provision in the final Directive would have promoted competition and solved access problems in cases where the *sui generis* right holder refuses to license the database. Similarly, Article 8(3) of the proposal also provided that member states should provide an arbitration mechanism for the purposes of disputes regarding such licenses.

Further, the *sui generis* right has raised concerns regarding its impact on research and development. It has been argued that the right would create an exclusive monopoly on data which can hamper access to knowledge and free flow of information. This exclusive control over data will give right holders an option to deny access to data or set unreasonably high prices for access. The pool of knowledge available to the public freely will be restricted since data holders would increasingly privatize data. However, on the other hand it has also been argued that creating a right in data could potentially benefit research and development. Since researchers and academicians are not only data consumers but also data producers, a right on data would help in improving their position in the commercial sector. Therefore, data producer's rights create both negative and positive effects on research and development.<sup>143</sup>

The Directive provides that the term for protection for non-original databases is fifteen years. However, these databases are updated continuously.<sup>144</sup> In such cases, the question that arises is would a new right be accrued every time substantial changes are made to the database? If the protection terms are renewable, it would raise concerns about perpetual protection. Furthermore, the *sui generis* right holder is the database maker. The Directive defines "Database maker" as "*the person who takes the initiative and the risk of investing.*" However, in certain situations, it will be

---

<sup>142</sup> European Commission, Proposal for a Council Directive on the Legal Protection of Databases COM(92) 24 final - SYN 393 (13 May 1992) <https://aei.pitt.edu/8653/1/8653.pdf> accessed 11 December 2024; Matthias Leistner, 'Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform' (7 September 2018) <https://ssrn.com/abstract=3245937> or <http://dx.doi.org/10.2139/ssrn.3245937> accessed 11 December 2024.

<sup>143</sup> G M Hunsucker, 'The European Database Directive: Regional Stepping Stone to an International Model?' (1997) 7 Fordham Intellectual Property, Media & Entertainment Law Journal 697.

<sup>144</sup> Peter K Yu, 'Data Producer's Right and the Protection of Machine-Generated Data' (2019) 93(4) Tulane Law Review 859.

challenging to identify the database maker or the person who invested in cases involving multiple participants. This lack of clarity regarding the right holder can cause legal uncertainties and thus, it is imperative to identify the database maker correctly.

Additionally, the exceptions to the database right have also attracted criticism. Critics have argued that the exceptions are too narrow compared to the copyright exceptions. For instance, the Directive does not provide for limitations for databases established by public authorities as provided for in national laws. As a result of which, the *sui generis* right covers copyright-exempted work created by public bodies. Therefore, these exceptions might need to be revised to align them with the copyright law to ensure harmonization. Limited and narrow exceptions can restrict the scope of new forms of research and development like text and data mining.<sup>145</sup> Text and data mining contribute to research and development as computational analysis of data can offer new knowledge and insights and promote innovation. Article 7 of the Directive restricts “*extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database.*” Extraction here includes both temporary and substantial transfer of parts of the contents of the database. Therefore, any extraction for the purposes of text data mining would infringe the *sui generis* right of the database maker. This restriction is concerning as it could potentially curtail development through text and data mining and the EU should re-evaluate the exceptions to the *sui generis* right.

The introduction of a *sui generis* right by the European Union is a significant and a bold step towards regulating data and the digital economy. As discussed in this section, the *sui generis* right is riddled with certain issues concerning its scope, identifying database maker, narrow exceptions, etc. The Directive was first evaluated in 2005. The purpose of the evaluation was to examine whether the Directive’s policy goals were achieved or not and whether the *sui generis* right has negatively affected competition.<sup>146</sup> In its analysis, the report noted that the scope of the legal protection granted by the Directive for non-original databases is not clear. Further, it notes that the

---

<sup>145</sup> European Commission, *Study in Support of the Evaluation of the Database Directive* (25 April 2018) <https://digital-strategy.ec.europa.eu/en/library/study-support-evaluation-database-directive> accessed 11 December 2024.

<sup>146</sup> European Commission, *First Evaluation of Directive 96/9/EC on the Legal Protection of Databases* (DG Internal Market and Services Working Paper, 12 December 2005) [https://openfuture.eu/wp-content/uploads/2021/12/2019EC-evaluation\\_report\\_legal\\_protection\\_databases\\_december\\_2005\\_en.pdf](https://openfuture.eu/wp-content/uploads/2021/12/2019EC-evaluation_report_legal_protection_databases_december_2005_en.pdf) accessed 11 December 2024.

economic impact of the *sui generis* right is unclear and unproven. In 2017, a public consultation was conducted with the aim to assess the functioning, application and impact of the Directive.<sup>147</sup> In the consultation, the majority of the participants were from the publishing sector, research, scientific and education sector, IT sector and the transport sector. With regard to the objective of balancing the interests of owners and users, participants from the research sector were of the opinion that the Directive does not achieve the objectives and the *sui generis* right is too broad with insufficient exceptions. The general view of the participants was that the Directive does not achieve its objectives. The view on the efficiency of the *sui generis* right is divided. If the EU does not abolish the right, it should consider dealing with the problems and challenges faced by the *sui generis* right. A clarification should be provided on the notion of the database maker and substantial investment. The exceptions to the right need to be revised again to ensure harmonization with public access policies.

#### **2.3.4. The EU Database Directive vis-a-vis Big Data**

It has been well established now that the definition of ‘database’ under the Database Directive is wide.<sup>148</sup> Therefore, the fact that the nature of big data is significantly different from traditional databases is inconsequential. Big data can fall under the scope of the Directive but ascertaining that would require a case-by-case assessment of databases. One of the requirements of ‘database’ under Article 1 is that the contents of the database should be a “collection of independent works, data or materials”.<sup>149</sup> The contents must have informative value or have “potentially informative content”. Since a majority of big data is machine generated and incomprehensible to the human eye, this requirement can be a worry. However, with the help of big data, machine generated data can be easily understood and information can be extracted. Thus, even if individual elements of big data hold no value, with big data analysis, can make the data comprehensible and the requirement under Article 1(2) can be fulfilled.

---

<sup>147</sup> European Commission, Summary Report of the Public Consultation on the Evaluation of Directive 96/9/EC on the Legal Protection of Databases <<https://digital-strategy.ec.europa.eu/en/library/summary-report-public-consultation-evaluation-directive-969ec-legal-protection-databases>> accessed 11 December 2024.

<sup>148</sup> *Ibid*, 33.

<sup>149</sup> Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases [1996] OJ L77/20, Art 1.

The second requirement under Article 1(2) is that data must be “arranged in a systematic or methodical way and individually accessible by electronic or other means.”<sup>150</sup> If the contents of the data is a mere accumulation of vast volumes of data, it cannot fall under the definition of ‘database’. However, to fulfill this requirement, it is sufficient to show that the material is “accessible by electronic or other means” or “searchable with a search engine”. In the case of big data, data can be arranged in a systematic way and made searchable through the application of sophisticated big data analytics technologies.<sup>151</sup> Therefore, whether or not big data falls within the scope of ‘database’ under the Database Directive is circumstantial.

The next issue in this context is whether big data can get copyright protection under the Directive. The Directive protects databases that are “*by reason of the selection or arrangement of their contents, constitute the author's own intellectual creation shall be protected as such by copyright*”<sup>152</sup> and must be the “author’s own intellectual creation”. This is the only criteria for copyright protection under the Directive. Therefore, originality must be reflecting the author’s personality or as emphasized by the CJEU; there should be author’s “personal touch” in the work.<sup>153</sup> Big data has mass volume and variety and is generally auto generated or machine generated and it is beyond human capacity to create or arrange big data.<sup>154</sup> However, the same can be arranged through big data analytics and that output can be arranged by a human being. Therefore, it is being argued that the output of big data processed with the help of big data analytics can get copyright protection to the extent that there is human involvement.<sup>155</sup>

The other issue concerns whether big data can get sui generis protection under the Directive. Unlike copyright protection, it is inconsequential for sui generis right that the content of big data is largely machine generated.<sup>156</sup> It has been clarified by the CJEU that the investment made in the

---

<sup>150</sup> *ibid* art 1.

<sup>151</sup> M Zeitlin, 'Everything Counts in Large Amounts: Protection of Big Data under the Database Directive' (2018) Law, Computer Science, Political Science.

<sup>152</sup> Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases [1996] OJ L77/20, Art 3(1).

<sup>153</sup> C-145/10 Painer, para 92; C-604/10 Football Dataco and Others, para 38.

<sup>154</sup> M Zeitlin (n 151).

<sup>155</sup> *ibid*.

<sup>156</sup> *ibid*.

database must be in obtaining, verifying or presenting the contents of the database and not in creation or making of the content.<sup>157</sup> In cases where the database maker has also created the contents of the database, as long as it is distinguishable from investments made in obtaining, verifying or presenting the contents, it can be protected. Vast volumes of raw data in big data databases is not within the scope of protection as investment in “obtaining” signifies that “material in the database must be pre-existing”.<sup>158</sup> Since the majority of big data databases involve investment in machine generated data which is essentially an investment in creation of data, many argue that it will not fall within the scope of the Directive.<sup>159</sup>

#### **2.4. The Data Producer’s Right**

In 2016, an impact assessment by the European Commission was published expressing that the legal uncertainty around the ownership of data is acting as a barrier for free flow of data and called for the creation of a new right on data. Later in 2017, to deal with this issue, the idea of a novel right for data producer’s was put forth. The new data producer’s right was a right on raw data and would protect machine generated data. The Proposed data producer’s right was a right in rem i.e. a right against the world and was introduced keeping 5 objectives in mind, “*to improve access to machine generated anonymised data, protect investment, incentivise data sharing, protect confidential data and to minimize lock-in effects*”. The Commission discussed issues related to machine generated data and the objective was to provide clarity regarding sharing of non-personal data across the EU, foster reliable protocols for identification and data sharing, foster the development of contract rules for B2B sharing, improve the functioning of the public sector, and etc.<sup>160</sup>

---

<sup>157</sup> Fixtures Marketing Ltd v Oy Veikkaus Ab, ECJ 9 November 2004, case C-46/02, ECR [2004] I-10396; Data Property: Unwelcome Guest in the House of IP P. Bernt Hugenholtz , See also Commission of the European Communities, 'First Evaluation of Directive 96/9/EC on the Legal Protection of Databases' (DG Internal Market and Services Working Paper, 12 December 2005).

<sup>158</sup> M Zeitlin , (n 151).

<sup>159</sup> Matthias Leistner, 'Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform' (5 October 2018) 20 <<https://ssrn.com/abstract=3253197>> accessed 17 January 2025..

<sup>160</sup> European Commission, 'Building a European Data Economy' COM(2017) 9 final (Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 10 January 2017).

It has been argued that the data producer's right would affect the intellectual property regime in the EU and would also come in conflict with the European Convention on Human Rights and the EU Charter. The EU Database Directive created two rights: copyright for protection of original databases and, a *sui generis* right for the protection of non-original databases. Both these rights did not extend to the contents of the database. The Database Directive either protects the originality and creativity of the database or the substantial investment that was put into the creation of a database. Therefore, the creation of a data producer's right would entail a protection that would encompass the protection granted by the intellectual property regime in the EU. This would eventually lead to an overlap between the two regimes and could give rise to multiplicity in claims. Further, if such a right were to be introduced, the exception for the same would have to align with the exceptions provided in the Database Directive.<sup>161</sup> Otherwise, the exceptions given in the EU database directive would be rendered useless. The right was also criticized for the lack of uncertainty it could potentially create regarding the scope of protection and the allocation of the right i.e. what all data would be protected and who will get the right.<sup>162</sup>

Furthermore, the European Convention of Human Rights provides for the right to freedom of expression and includes the right to receive information without any interference. The database rights would interfere with the citizen's right to free flow of information and would also restrict the access to data for scientific research. The ECHR in Article 16 enshrines the right to freedom of competition. Big data is an essential tool for innovation and restrictions on access to such data can create competition law issues as well. The data producer's right would create monopoly in the market and create barriers to competition.<sup>163</sup> Owing to these challenges, the data producer's rights failed and merely remained a thought experiment.<sup>164</sup>

---

<sup>161</sup> Francesco Banterle, 'Data Ownership in the Data Economy: A European Dilemma' in Tatiana-Eleni Synodinou and others (eds), *EU Internet Law in the Digital Era: Regulation and Enforcement* (Springer 2020).

<sup>162</sup> Bernt Hugenholtz, "Data Property: Unwelcome Guest in the House of IP" (Paper presented at Trading Data in the Digital Economy: Legal Concepts and Tools, Münster, Germany, 2017). <[https://pure.uva.nl/ws/files/16856245/Data\\_property\\_Muenster.pdf](https://pure.uva.nl/ws/files/16856245/Data_property_Muenster.pdf)> accessed 16 December 2024.

<sup>163</sup> *ibid.*

<sup>164</sup> Dev S Gangjee, 'The Data Producer's Right: An Instructive Obituary' in Ernest Lim and Phillip Morgan (eds), *The Cambridge Handbook of Private Law and Artificial Intelligence* (CUP 2022).

## 2.5. Challenges in Protecting Data as Intellectual Property

Data is a valuable asset for individuals, businesses and the state. The ever-increasing value of data in today's digital economy has raised issues regarding protecting data and the rights of data holders. However, the intellectual property protection of data raises complex issues and challenges. Copyright protects databases that are original in their arrangement and selection. The protection also does not extend to the contents of the database since copyright does not protect readily available and non-rivalrous data. Therefore, possible risks of data misappropriation remain and threaten development and efficient trade. It has also been held in *Football Dataco and others* that copyright protection cannot be granted to work generated by machines without any human intervention. Thus, copyright requires work to be of human authorship. Due to these concerns, raw data is preempted from copyright protection. In fact, Geoffrey Bowker has said that "*Raw data is both an oxymoron and a bad idea; to the contrary, data should be cooked with care.*" This is because the term "raw" refers to something untouched and "cooked" implies "result of cognitive process" Since data is a product of cognitive process in the sense that it needs to be determined what to collect.<sup>165</sup>

The concept of data and "raw data" is complex and difficult to define. Ascertaining the right holders and the extent of their rights over the data would also be a challenging task. The question regarding extending intellectual property protection to data would involve considerations on the impact on innovation and development. On one hand it is contended that the lack of protection for databases would act as a disincentive for creation of value-added databases. While on the other hand, is the dilemma that providing protection to data would hamper access to information and that information should remain part of the public domain. The potential consequences of non-access to data would be more severe for developing countries like India as it could obstruct scientific and academic research.

Zech has argued for four justifications for granting intellectual property protection to machine generated raw data. Firstly, an IP right in data would help in optimal data allocation in market since it can help in disclosing data. Secondly, it could help in determining who should get benefits

---

<sup>165</sup> Nick Barrowman, 'Why Data Is Never Raw' [2018] 56 *The New Atlantis* 129.

arising out of data and thirdly, help in understanding data ownership. Lastly, data is an important asset for companies and companies enter into contracts dealing with data and therefore, data should be treated like a property.

Further, when it comes to creating a novel data producer's right to protect investment to promote data production and analysis, economists have argued that there is insufficient evidence supporting the need to create an economic incentive for database creators.<sup>166</sup> Meaning thereby that there are enough incentives already available in the market.

Additionally, the data producer's right could raise concerns regarding the producer's control over the dissemination of information. Consequently, exclusive control over information and data would contravene freedom of expression and free flow of information and curtail fair competition in the market. In our expert interviews, it was noted that data sharing would give different people opportunities to access the same non-rivalrous data, and more people will innovate. Therefore, sharing of data is being considered as a catalyst for innovation.<sup>167</sup> Further, when asked whether an intellectual property right should be granted over data, Kris Gopalan responded in his interview that raw data is not proprietary and cannot be protected. Furthermore, data can be used in multiple ways, it can be combined with other data and new data can be derived from it. Using data in this manner can boost innovation and therefore free flow of information should not be hampered with.

### **3. Unfair Competition and Data**

#### **3.1. Trade secret**

A trade secret is confidential information that has commercial value by virtue of being kept a secret. The TRIPS Agreement under Article 39 provides protection for undisclosed information that is secret in the sense that it is not readily accessible to the public, has commercial value, and reasonable efforts have been made by the holder of the information to keep it secret.<sup>168</sup> Secrecy is

---

<sup>166</sup> W. Kerber, 'A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis, [2016] *GRUR Int.* 989, 997.

<sup>167</sup> Interview Responses.

<sup>168</sup> TRIPS 1994, Art 39.

not an absolute criterion for protection but a relative one. Therefore, the secrecy status of the information can be maintained even when it is shared with third parties by the holder.

In 2016, the European Union adopted the Trade Secret Directive (TSD) to harmonize the principles regarding trade secrets across the union and to promote better sharing of knowledge in business to business transactions. It is an important tool for protection of confidential information in the digital economy. The Directive defines trade secrets under Article 2(1) as information that can fulfill threefold requirements that are (1) it is a secret and is not generally known or readily available to persons normally dealing with the kind of information in question, (2) information has commercial value that is owed to its secrecy and (3) the person in lawful control of the information has taken reasonable steps to maintain the secrecy.<sup>169</sup> A reading of Recital 14 of the Directive suggests that the Commission's aim was to ensure that the definition of trade secrets is homogenous that doesn't restrict the subject matter. A broad definition of trade secrets encourages that commercially valuable and confidential private information can be protected.

Trade secret protection is sometimes preferred over other intellectual property protection. There are mainly two reasons for this. Firstly, it provides for an unlimited duration of protection as long as the secrecy is maintained, and secondly, it requires a lower threshold for protection as compared to other IPs. However, a trade secret does not grant an exclusive right to use. Recital 16 of the EU Trade Secret Directive provides that no exclusive right would be created in trade secrets with the intention to foster innovation and competition.<sup>170</sup> It further notes that reverse engineering of a "*lawfully acquired product will be considered a lawful means to acquire information.*" Therefore, businesses can rely on trade secret protection for their business information. However, they cannot rely on it since there will always be a risk of reverse engineering by competitors.

India does not have a specific law governing trade secrets. However, confidential information is protected through contracts and law of equity. India is a signatory of the TRIPS Agreement and the requirements provided under Article 39 are to be fulfilled to get trade secret protection in India as well. Information which is confidential and reasonable steps have been taken by the holder of

---

<sup>169</sup> Trade Secret Directive (EU), Art 2(1).

<sup>170</sup> *ibid*, rule 16..

the information to maintain its secrecy and if disclosed would be detrimental to the interests holder of the information is considered as trade secret in India.<sup>171</sup> This information can include agreements, client lists, information about clients, customer lists, business information, market strategies, etc. can qualify as trade secrets.<sup>172</sup>

### 3.2. Big Data and Trade Secret Protection

Trade secrets allow the information holder to protect business know-how and business information. It provides safeguards and protects business information without the requirement of originality. Therefore, businesses often use trade secret protection to protect information that can otherwise not be protected by other kinds of intellectual property rights.<sup>173</sup> Therefore, due to this broad understanding of 'information', trade secrets are another way data can be protected if it satisfies the three requirements mentioned above. Therefore, non-personal data that is a secret and has potential commercial value can be protected as a trade secret. However, in the context of big data as trade secrets, the issue arises whether "individual data" or "*datum*" can fall within the scope of trade secret protection.<sup>174</sup> The European Commission has also expressed doubts about "individual data generated by interconnected machines and devices could be regarded as 'trade secret'" since it cannot have commercial value.<sup>175</sup> However, the Commission does acknowledge that a combination of data can fall within the scope of the TSD if all requirements are met.<sup>176</sup> Further, Recital 14 of the TSD also states that trade secrets cannot protect trivial information. However, in the context of big data and the digital economy, trivial information, when compiled and analyzed using big data analytics can have immense commercial value. Therefore, raw data has value (albeit low) that hasn't yet been utilized. It is in this context that Zech argues that since there is no minimum threshold for requirement of commercial value, one of the requirements of

---

<sup>171</sup> *Diljeet Titus Vs Alfred A Adebare and Ors* (2006), *Hi-Tech Systems and Services Ltd Vs Suprabhat Ray and Ors* (2015) and *Burlington Home Shopping Pvt Ltd Vs Rajnish Chibber*.

<sup>172</sup> *Ambience India Pvt Ltd v Shri Naveen Jain; Diljeet Titus v Alfred A Adebare & Ors*.

<sup>173</sup> Inp pan and Paweł Podrecki 'Protection of Trade Secrets and Know-How in the European Union' (WIPO Regional Seminar on Trade Secrets).

<sup>174</sup> Tanya Aplin and others, *Trading Data in the Digital Economy: Trade Secrets Perspective* (Nomos, Baden-Baden 2017).

<sup>175</sup> European Commission, 'Staff Working Document on the Free Flow of Data and Emerging Issues of the European Data Economy' SWD(2017) 2 final (10 January 2017).

<sup>176</sup> *ibid*.

trade secrets are fulfilled.<sup>177</sup> The other requirements for trade secret are secrecy and taking steps to maintain secrecy and both of these would require a subjective assessment. If data sets are part of the open data movement, they cannot be secret but if datasets are not openly accessible or are “not generally accessible to people” and protective measures are adopted to maintain confidentiality, it fulfills all the requirements.<sup>178</sup>

However, a key issue concerning trade secret protection for big data is how “information” is interpreted, i.e. “*trade secret as a body of information distinct from its individual components.*” While analyzing trade secret protection for big data, it is important to create a distinction between data on a semantic basis and data on a syntactic basis.<sup>179</sup> In this context, an Italian ruling has noted that the “commercial value of the secret information derived not from the information itself, but from the way it might be processed by dynamic technologies.” The Tribunal further noted that the potentiality of the information to qualify as a trade secret has to be assessed not by judging the value of individual information, rather by judging the combination of data as a whole. Therefore, it was held that the commercial value lies in the technological methods applied on the dataset and not the individual elements in isolation.<sup>180</sup>

### **3.3. Big Data and Issues with Trade Secret Protection**

Trade secrets protect a broad range of information. Therefore, it is very flexible when it comes to the subject matter. This wide ambit of protection has raised specific over-protection problems. The broad nature of protection allows businesses to protect any sort of information, which, as a consequence, can obstruct public access to critical information. For instance, big data companies

---

<sup>177</sup> Herbert Zech, 'A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data' (2016) 11 *Journal of Intellectual Property Law & Practice* 460-470 <<https://doi.org/10.1093/jiplp/jpw146>> accessed 17 January 2025.

<sup>178</sup> Tanya Aplin, 'Trading Data in the Digital Economy: Trade Secrets Perspective' (2017) 59 <<https://doi.org/10.5771/9783845288185-59>> accessed 18 January 2025.

<sup>179</sup> European Union Intellectual Property Office, *Trade Secrets Litigation Trends in the EU* (EUIPO 2023) <[https://euipo.europa.eu/tunnel-web/secure/webdav/guest/document\\_library/observatory/documents/reports/2023\\_Trade\\_Secrets\\_Litigation\\_Trends\\_in\\_the\\_EU/2023\\_Trade\\_Secrets\\_Litigation\\_Trends\\_Study\\_FullR\\_en.pdf](https://euipo.europa.eu/tunnel-web/secure/webdav/guest/document_library/observatory/documents/reports/2023_Trade_Secrets_Litigation_Trends_in_the_EU/2023_Trade_Secrets_Litigation_Trends_Study_FullR_en.pdf)> accessed 16 December 2024.

<sup>180</sup> Leonardo Assicurazioni s r l v Pro Insurance s r l & A&A Insurance Broker s r l & Unknown (14 May 2018)

like Google and Uber hold vast amounts of vital information that can be used for public good and infrastructural development.<sup>181</sup>

Concerns about restricting access to small businesses with lower bargaining powers have also been raised. Businesses can simply deny sharing information to gain a competitive market advantage. This can result in worsening the positions of smaller market players. Further, the velocity of the data is crucial when it comes to the usability of data. The value of data decreases with time, and by the time small companies gather data or get access to data, the value of the said data is depleted. To illustrate this issue, the US case of *Lyft Inc v City of Seattle*<sup>182</sup> can be examined. In 2014, Lyft and Raiser were asked to share data relating to the number of rides, ride percentage for zip codes and pick-up and drop-off zip codes in the form of a standardized quarterly report with the municipality. *Lyft* and *Raiser* argued that the information cannot be shared as it is confidential in nature. However, the Washington Supreme Court held that the report is a trade secret but is not exempted from disclosure under the Public Records Act. The records of Lyft are public record and need to be shared for public interest. It was noted that trade secrecy hampered the municipality's access to important data that could be used for public interest. Considering that the nature of non-personal data held by the firms was time-sensitive, delayed access to the data further delayed urban planning. Herein, the use of trade secret protection for proprietary data is not enough to keep aside public interests.<sup>183</sup>

On these lines, our interviewee, Kris Gopalan, also supported that raw data cannot qualify as a trade secret and should be made available for the public good. Data relating to infrastructure and phenomena like traffic patterns, climate, etc., can be helpful for the public and should be made available for the public good.<sup>184</sup>

---

<sup>181</sup> Interview Responses.

<sup>182</sup> *Lyft Inc v City of Seattle* 94026-6 (WA 2018); Tommaso Fia, 'Resisting IP Overexpansion: The Case of Trade Secret Protection of Non-Personal Data' (2022) 53 IIC - International Review of Intellectual Property and Competition Law 917.

<sup>183</sup> Interview Responses.

<sup>184</sup> Interview Responses.

Article 5 of the EU Trade Secret Directive provides for certain exceptions to third party use or acquisition of trade secrets. These exceptions include “*exercising right to freedom of expression, revealing misconduct, wrongdoing or illegal activity for public interest, disclosures by workers in accordance with law and, for the purpose of protecting a legitimate interest recognised by Union or national law.*”<sup>185</sup> Recital 21 of the directive further provides that while protecting trade secrets, the principle of proportionality should be kept in mind for smooth functioning of the market for research and innovation. These measures can be used to protect parties not having access to information.

Another point that needs to be considered is that trade secrets do not grant an exclusive right in the data or the database. It merely grants protection against misappropriation of confidential information. For instance, if there is a breach of confidentiality, the data holder can take an action against the breaching party but cannot restrict third parties from accessing or using the information thus leaked.<sup>186</sup> Therefore, trade secret protection is not sufficient to protect data.

### **3.4. Data Misappropriation and Unjust Enrichment**

The doctrine of Misappropriation emerged in the United States and is defined as “*the unauthorized, improper, or unlawful use of funds or other property for purposes other than that for which intended.*”<sup>187</sup> The doctrine grants a “*quasi-property*” right to people who invest substantial time, skill and labor to create an intangible asset. The US Supreme Court in *International News Service v Associated Press*, wherein International News Service (INS) started lifting news stories from Associated Press (AP), noted that a quasi-property right lies in the facts gathered by AP. However, this right is not a right in rem because facts are unprotect-able, but is a right against competitors. The rationale for protecting AP’s news stories was threefold. Firstly, when substantial time, effort and resources are invested by someone, they should be able to reap its benefits. Secondly, news should be protected as it is valuable like any other property. Thirdly, to provide an economic incentive to gather news.<sup>188</sup> Misappropriation is also understood as “*taking use of another’s*

---

<sup>185</sup> EU TSD Article 5.

<sup>186</sup> Kerber (n 166).

<sup>187</sup> ‘Misappropriation’, Wex Definitions <<https://www.law.cornell.edu/wex/misappropriation>> accessed 16 December 2024

<sup>188</sup> *ibid.*

*property solely for the purposes of benefiting unfairly from it.*<sup>189</sup> Under IP jurisprudence, facts cannot be protected. However, if substantial investment has been made to gather and acquire these facts and someone tries to take advantage of others' efforts, it can be protected through misappropriation.

One of the main motives why businesses rely on trade secrets to protect their confidential and commercially viable data is to prevent third party misappropriation of their data. According to a study, the European Trade Secret Directive has not conclusively proved to promote disclosure of information or data sharing. Rather, it has encouraged them to resort to taking actions that would preserve the secrecy of their data.<sup>190</sup> Further, it is very difficult to determine exactly when information has been misappropriated and by whom.

#### **4. Contract Law and Data**

The EU Regulations (EU) 2016/679 and 2018/1807 established the principle of free flow of data and laid the foundation of considering data as an intangible asset and a tradable commodity. This led to the objectification of data as something valuable that can be shared and licensed.<sup>191</sup> With the increase in the perception of data as a property, contractual transactions treating data as a commodity also increased.

Contract law plays a vital role in protecting data and facilitating data sharing. Intellectual property protection for data and non-original databases is not uniform across jurisdictions. In this context, Contract law is relied on for protecting an enterprise's data and databases in the context of data sharing and maintaining data secrecy. Therefore, stakeholders rely on contracts for clearly determining the ambit of their rights in the data they hold. Furthermore, it allows the data holder to share data based on its terms and conditions. With the freedom to determine the terms and conditions of transactional contracts, big companies can retain complete control and ownership over the shared data. There is a freedom to give data access on preferential terms to preferred

---

<sup>189</sup> *Pocket Books, Inc v Dell Publ'g Co* 267 N.Y.S.2d 269, 272 (App Div 1966).

<sup>190</sup> Tanya Aplin and others, 'The Role of EU Trade Secrets Law in the Data Economy: An Empirical Analysis' (2023) 54 IIC - International Review of Intellectual Property and Competition Law 826.

<sup>191</sup> Carolina Perlingieri, 'Data as the Object of a Contract and Contract Epistemology' (2019) 5(2) Italian LJ 613.

parties and simultaneously restrict use and access. Therefore, contract law protects the rights and interests of data holders and creators in an economy where the law around data is unclear and developing, for example, data transfer agreements and data pools.<sup>192</sup>

When data is perceived as a commodity, two important aspects come up. Firstly, if data is governed by a law, then there are certain people who would be entitled to trade such data. For example, personal data can be traded by the person who owns the personal data. Secondly, does tradability of data encompass the prospects of reselling the data beyond primary markets.<sup>193</sup>

There can be multiple ways to commercialize data. Data can be sold as a commodity, access to data can be offered as a service and services can be offered based on the data that the data holder has access to, wherein the data holder can collect and produce data and analyze it to offer services based on it. For example, Google collects data and analyzes it and offers a service like Google Maps. Offering services based on data does not require the data holder to disclose or share the information they hold. Another way to commercialize data is to offer access to collected and produced raw data. When data is being transacted through contracts, a distinction can be drawn out between data being traded as a commodity and offering services based on data.

In cases where a corporeal object is being sold as a commodity, there is a transfer of possession. However, data is a non-rival and non-excludable good that can be simultaneously used by many people. Therefore, when data is being commercialized like a commodity, its possession does not get transferred, rather its access is granted. Thus, data is only traded like a commodity when the data holder, while granting access to data, also refrains from using the said data. Otherwise, if the data holder does not grant exclusive use to the other party, access to data is being offered as a service and there is no ‘transfer’ of data per se. When it comes to creating a distinction between goods and services, the aspect of time also comes into consideration. In a seller-purchase contract, there is a final delivery of goods and these goods can be kept for however long the purchaser wants

---

<sup>192</sup> Tommaso Fia, ‘An Alternative to Data Ownership: Managing Access to Non-Personal Data through the Commons’ (2021) 21 *Global Jurist* 181.

<sup>193</sup> Herbert Zech, ‘Data as a Tradeable Commodity’ in Alberto De Franceschi (ed), *European Contract Law and the Digital Single Market: The Implications of the Digital Revolution* (Intersentia 2016).

to. On the other hand, services are consumed when they are rendered.<sup>194</sup> Zech explains this using an example: downloading music or buying CDs/Cassettes is a commodity paradigm and streaming music on platforms like Apple Music, Spotify and YouTube is a part of the service paradigm. These differences in the service and commodity paradigm are important for contract law as it affects the kind of contract and contractual obligations parties enter into

#### **4.1. Contractual Issues in Data Transactions/Sharing**

A data sharing agreement is an agreement between two parties that covers how data between them is to be shared. It covers what data will be shared, how much data will be shared, how it will be shared and how the data will be used and analyzed. Data sharing agreements help parties by giving them clarity about their roles and scope of contract, helps in determining the purpose of data sharing, it covers and provides clarity about every stage of data sharing and, it sets standards for data sharing.<sup>195</sup> Use of data enhances productivity, fosters innovation and generates revenue for businesses. Data is being increasingly commercialized by businesses. Most of this commercialisation takes place through contractual transactions for trading and licensing data. However, despite its benefits, it has been noted that data is not being commercialized at its optimum level. In a research survey of data and analytic businesses, it was found that only one-third of businesses were commercializing their data.<sup>196</sup>

Transacting data through agreements grants autonomy to the contracting parties, which is their right to enter into a contract freely. The parties have the ability to mold the contractual relationship the way they want to. However, since rights and obligations are determined by the contracting parties themselves, there are chances that these rights and obligations may not be allocated correctly. Further, in data-sharing contracts, the data controller holds a higher bargaining power and is an influential party. The *de facto* possession of data more than often translates into

---

<sup>194</sup> *ibid.*

<sup>195</sup> Information Commissioner's Office, 'Data Sharing Agreements' (19 November 2024) <<https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/data-sharing-a-code-of-practice/data-sharing-agreements/>> accessed 16 December 2024.

<sup>196</sup> IBM, 'Forrester Study: Unlock the Power of Data to Transform Your Business' (August 2019) <https://adtmag.com/whitepapers/2019/08/ibm-forrester-study-unlock-the-power-of-data-to-transform-your-business-iuk.aspx#:~:text=In%20September%202018%2C%20IBM%20commissioned%20Forrester%20Consulting%20to,artificial%20intelligence%20to%20unlock%20the%20power%20of%20data> accessed 31 January 2025.

ownership and causes unequal negotiating powers between the contracting parties.<sup>197</sup> Parties are free to mold the terms of the data-sharing contracts as they want to and data-requesting party's lack of bargaining power raises issues concerning their party autonomy.<sup>198</sup> Therefore, due to the unequal bargaining power, data sharing through contracts can sometimes be unfair.<sup>199</sup> For instance, in the agricultural sector, it has been noted that farmers hesitate in entering into contracts with international technology providers because they feel vulnerable and less powerful.<sup>200</sup>

The construction of a data sharing agreement is another challenge since a wide range of factors are covered in such agreements. For example: IP protection, pricing models, legal compliances, data rights, quality of data (QoD), commercial/non-commercial use of data, liability, etc. Data sharing agreements are unique as “*data are heterogeneous and contract terms are contextual*” i.e. the laws on contract vary across jurisdictions.<sup>201</sup>

There are several other issues that haunt contractual sharing of data. There is heavy economic uncertainty concerning how the data sharing parties are going to share profits. The lack of clarity on how the profits accrued from data sharing are to be shared discourages firms from sharing data or investing into the same. These “economic barriers” create hurdles in creating “complete” and effective data sharing contracts. Incompleteness of contract can be harmful, and studies show that many data contracts do not cover crucial aspects concerning data. For example, most DaaS contracts do not cover quality of data (QoD) which is an essential component of DaaS contracts.<sup>202</sup>

---

<sup>197</sup> Fia n 192

<sup>198</sup> Michiel Rhoen, ‘Beyond Consent: Improving Data Protection through Consumer Protection Law’ (2016) 5 Internet Policy Review <<https://policyreview.info/articles/analysis/beyond-consent-improving-data-protection-through-consumer-protection-law>> accessed 16 December 2024.

<sup>199</sup> OECD, *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-Use across Societies* (Organisation for Economic Co-operation and Development 2019).

<sup>200</sup> Leanne Wiseman and others, ‘Farmers and Their Data: An Examination of Farmers’ Reluctance to Share Their Data through the Lens of the Laws Impacting Smart Farming’ (2019) 90–91 NJAS: Wageningen Journal of Life Sciences 1.

<sup>201</sup> Llewellyn D W Thomas and Aija Leiponen, ‘Big Data Commercialization’ (2016) 44 IEEE Engineering Management Review 74.

<sup>202</sup> Hong-Linh Truong and others, ‘Data Contracts for Cloud-Based Data Marketplaces’ (2012) 7(4) International Journal of Computational Science and Engineering 280.

Additionally, creating legal certainty by including all relevant factors in a contract is not practical. For example, there are many unforeseen costs that are traditionally the burden of the owner of the service/goods. However, since there is no legal clarity on who owns data, there is no clarity on who will pay these unforeseen costs.<sup>203</sup>

Further, another issue that arises with data sharing through contracts is that the information shared between parties can be used by the receiving party for purposes not specified in the contract. A situation like this can be detrimental to the economic interests of the data sharer. To deal with such situations, a contract with a strict confidentiality clause and end use restrictions can be entered into. However, this does not curtail the risk of information being misappropriated by the information receiver because it is difficult to trace whether the information/data shared by the data sharer was in fact used by the receiving party. It cannot be easily determined where the information goes once it is disclosed to the recipient.<sup>204</sup>

Furthermore, one way to enable text and data mining activities is through licensing contracts since unauthorized TDM activities can infringe on the rights of data holders. However, licensing contracts cannot be solely relied upon to facilitate TDM activities as it can pose certain problems. Firstly, only selective databases would be available for licensing and secondly, it is highly likely that the license would be offered on unreasonable terms.

Thus, even though contract law acts as a cushion for stakeholders to fall back on, it does pose other issues like transaction costs, unequal bargaining power for determining the terms of the contract, potential risk of data misappropriation etc. Some scholars present that an effective data sharing regime should include: “(1) data right affirmation to ensure that economic benefits are allocated properly, (2) clear determination of the nature of transaction (this is relevant for all contracts but considering the peculiarity of data, this is crucial for data sharing agreements) and, (3) supporting

---

<sup>203</sup> Nadine Stüdlein, ‘Data as a Common Good: Essays on Data Portability and B2B Industrial Data Sharing’ (Doctoral thesis, Universität Passau 2022) <<https://opus4.kobv.de/opus4-uni-passau/frontdoor/index/index/docId/1120>> accessed 16 December 2024.

<sup>204</sup> Interview responses

system for data transaction (personal data protection, data security, cross border flows).<sup>205</sup> However, the European Commission has noted that data sharing through contracts can contribute to data lock-ins that can disrupt free flow of data<sup>206</sup> and these lock-in issues cannot be overcome by contract law.<sup>207</sup> Therefore, contract law in the paradigm of data markets is relevant for data sharing but there are some loopholes that need attention.

#### **4.2. Non-disclosure agreements and their significance**

Non-disclosure agreements are legal agreements between parties to a contract to maintain confidentiality and secrecy. An NDA binds the contracting parties to not disclose the information shared by any party. NDAs provide confidence and flexibility to the contracting parties in sharing information and data along with non-disclosure obligations.

Non-disclosure Agreements play a vital role in protecting sensitive information during business transactions, especially data sharing. NDA ensures that confidentiality between the contracting parties is maintained by allowing data holders to restrict the other contracting party from sharing information with third parties, giving them a competitive advantage in the market. It also allows contracting parties to seek compensation and damages if confidentiality breaches.

In data sharing agreements, one of the major issues is having no control over who can access the data once it is shared. This bolsters the fear of data holders that their confidential data is at risk of being misappropriated and once a firm shares their data, it won't be able to control or anticipate how their data will be utilized by their competitors. The lack of standards governing data sharing results in firms having a fear of sharing information.<sup>208</sup>

---

<sup>205</sup> Hongbin Yu, 'The Nature and Rule Construction of Data Transaction Activities' (2023) 20(1) US-China L Rev 24.

<sup>206</sup> Maria Jose Schmidt-Kessen, 'The Impact of Data Ownership Rights on Competition in Big Data Markets: Reflections in the Context of the EU and Global Data Economy' (2020) 2(2) VIT L Rev 56.

<sup>207</sup> Josef Drexl, *Data Access and Control in the Era of Connected Devices Study on Behalf of the European Consumer Organisation BEUC* (BEUC 2018).

<sup>208</sup> Nadine Stüdlein, 'Data as a Common Good: Essays on Data Portability and B2B Industrial Data Sharing' (Doctoral thesis, Universität Passau 2022) <<https://opus4.kobv.de/opus4-uni-passau/frontdoor/index/index/docId/1120>> accessed 16 December 2024.

Herein, the digital economy where data is an asset for innovation and development, NDAs become crucial for data transactions. It allows businesses to share information and data while maintaining confidentiality and secrecy of the information or data so shared. This gives them a competitive edge in the market. Contracts can establish clear obligations of the parties, such as the scope of information shared, the purpose of use or limits to analyzing the data shared, limits on subsequent sharing, confidentiality, etc., which are set out in data sharing agreements and NDAs.<sup>209</sup> Thus, NDA can be a crucial tool for business in the digital economy. It can help nurture trust between the contracting parties as they assure the parties that their interest would be protected. However, NDAs are only enforceable against contracting parties and any case of unwanted data disclosure puts the data-sharing party at risk and sometimes contractual instruments cannot be sufficient to free use and share data.

#### **4.3. Standard Rules for Data Sharing Contracts**

In this context, creating standard or default rules for contracts dealing with data transactions may be considered to prevent abuse of party autonomy. These standard rules will act as guidance for parties, minimize unfair contract terms and reduce transaction costs.<sup>210</sup> Competition law deals with cases of abuse of dominance. However, the threshold that needs to be met to prove abuse of dominance is too high. Therefore, providing standard contract rules for data transaction agreements may be beneficial in ensuring that the benefits are allocated fairly to all the contracting parties. By using standard contract rules, the legislature can also exercise control and ensure that there is a balanced distribution of rights and obligations.<sup>211</sup> Further, it has been argued that having default or standard rules would be rooted in antitrust principles since it would promote free and fair trade and promote economic growth.<sup>212</sup> In furtherance of this, some governments have taken initiatives to encourage data sharing. Majority of these initiatives focused on enhancing access to

---

<sup>209</sup> Dev S Gangjee, 'The Data Producer's Right: An Instructive Obituary' in Ernest Lim and Phillip Morgan (eds), *The Cambridge Handbook of Private Law and Artificial Intelligence* (CUP 2022).

<sup>210</sup> Matthias Leistner, 'Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform' in S Lohsse, R Shulze and D Staudenmyer (eds.), *Trading data in the digital economy: legal concepts and tools* (Nomos, Baden-Baden 2017) <t: <https://ssrn.com/abstract=3245937>> accessed 14 December 2024.

<sup>211</sup> S Lohsse, R Shulze and D Staudenmyer (eds), *Trading Data in the Digital Economy: Legal Concepts and Tools* (Nomos, Baden-Baden 2017).

<sup>212</sup> Friedrich Graf von Westphalen, 'Contracts with Big Data: The End of the Traditional Contract Concept?' in S Lohsse, R Shulze and D Staudenmyer (eds), *Trading data in the digital economy: legal concepts and tools* (Nomos, Baden-Baden 2017).

public sector data. For example, Australia introduced the Data Sharing and Transparency Act. The objectives of the Act were to boost public sector data sharing, ensure integrity and transparency in data sharing, comply with privacy norms whilst data sharing, and to build confidence in the use of public data.<sup>213</sup>

Some initiatives have also been taken to regulate private sector data sharing. These initiatives are generally of two kinds, i.e. contract guidelines and public private collaborations.<sup>214</sup> Contractual guidelines provide defined guiding principles for the parties to negotiate agreements. These guidelines are not mandatory for the parties to follow, and parties can freely deviate from it. Japan has taken an initiative on these lines to provide guidelines for contracts relating to use of AI and Data. These guidelines by the Ministry of Economy, trade and Industry provide a direction to the contracting parties for negotiating the terms of the agreement. The contract guidelines categorize data sharing contracts into three categories. First, data provision contracts where the data is owned by one party and is shared with the other party. Second, a data creation contract where both/all the parties to the contract create data and negotiate the terms of use of the said data. Third, data sharing agreements where data is shared through a platform that collects, processes and analyzes the data.<sup>215</sup> The Ministry of Agriculture, Forestry and Fisheries (MAFF) also introduced a similar Guideline on Contracts Regarding Utilization of AI and Data in the Agricultural Sector to promote smart agriculture in Japan. The guidelines aim to instill confidence in farmers regarding their data and provide guidelines to protect trade secrets.<sup>216</sup>

But it should be kept in mind that determined contract standard rules would be helpful in avoiding issues like lock in and unequal bargaining power, balance must be struck within the value chain

---

<sup>213</sup> Data Availability and Transparency Act 2022 <<https://www.legislation.gov.au/C2022A00011/latest/text>> accessed 17 December 2024.

<sup>214</sup> OECD, *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-Use across Societies* (Organisation for Economic Co-operation and Development 2019).

<sup>215</sup> Akira Matsuda, Ryohei Kudo and Taiki Matsuda, ‘AI, Machine Learning & Big Data Laws and Regulations 2024 – Japan’ (*Global Legal Insights*, 2024) <<https://www.globallegalinsights.com/practice-areas/ai-machine-learning-and-big-data-laws-and-regulations/japan>> accessed 17 December 2024.

<sup>216</sup> Maria Jose Schmidt-Kessen (n 206) .

and investments must also be protected. A fair balance of allocation of rights and obligation among all the parties in the value chain would be appreciable.<sup>217</sup>

## **5. Conclusion**

Since data creation, collection, processing and storage entails substantial opportunity costs, firms tend to fiercely protect it from their competitors. This is done primarily through four ways. Intellectual property (IP) law plays a pivotal role in protecting big data. However, copyright law limits the protection to original databases in terms of their sequence, structure and organisation and data can per se not be protected through copyright law under the Modak decision of the Supreme Court as data does not satisfy the originality standard of minimum creativity. The EU's Database Directive, with its dual protection of original and non-original databases, exemplifies how IP law can incentivize data creation and sharing. The sui generis database right, in particular, aims to protect substantial investments made in obtaining, verifying or presenting the contents of the database, thereby encouraging innovation. The EU also introduced a novel Data Producer's right on raw and machine generated data with the objective to improve access to data and to promote data sharing. However, owing to certain challenges and massive criticism, the thought remained a mere unimplemented experiment. Intellectual property protection for data raises complex issues that call for a balance between creating a property right in data and ensuring optimal allocation of data along with information dissemination for the greater social and economic good.

Datasets can also be protected through common law remedies for protection of undisclosed information in the form of breach of confidence for misappropriation of undisclosed information. However, regarding trade secret protection for big data, an issue arises whether protection can be extended to individual datum. The limitation of these protections to datasets with commercial value raises questions about the scope and effectiveness of existing legal frameworks. As data becomes increasingly central to business strategies, there is a need to reassess and possibly expand the legal definitions and protections for data. Contract law is also a fundamental tool for governing data transactions, with non-disclosure and confidentiality agreements being standard practices. However, contractual protection has its pitfalls, including issues related to data ownership, party

---

<sup>217</sup> C Makridis and B Dean, 'Measuring the economic effects of data breaches on firm outcomes: Challenges and opportunities' (2018) 43(1) Journal Economic and Social Measurement 59.

autonomy, and unequal bargaining power. These challenges necessitate a careful consideration of the terms and conditions in data sharing agreements to prevent unfair practices and ensure equitable access to data. In furtherance of this, standard rules for regulating data sharing contracts that can help in providing foundation for effective and fair contractual sharing of data can be a great tool for instilling confidence in data sharing without the risks of data breaches.

## CHAPTER III: CYBER-SECURITY DIMENSION INVOLVING DATA BREACH AND DATA LEAK

### 1. Cyber Space and Cyber Security

The increased access to information and the expansion of the internet in the digital economy has made it challenging to protect valuable information. Information technology has been contributing to the growth of the economy everywhere and has aided in promoting global integration and innovation and yet information and cyber security lags.<sup>218</sup> In this context, it becomes crucial to understand what is meant by the terms “cyberspace” and “cyber security”.

The term cyberspace originates from the Greek word “Kyber” which literally means “to navigate” or “navigable space”. The term was coined in 1984 by an American science fiction author named William Gibson in his novel titled “Neuromancer”. Gibson’s definition of cyberspace referred to “*a navigable, digital space of networked computers accessible from computer consoles; a visual, colorful, electronic, Cartesian datascape known as ‘The Matrix’ where companies and individuals interact with, and trade in, information.*”<sup>219</sup> Since then, the term has been reappropriated and defined by many. Therefore, there is no uniform definition of “cyberspace” that has been established at an international level yet.<sup>220</sup> NATO has defined cyberspace as “*the global domain consisting of all interconnected communication, information technology and other electronic systems, networks and their data, including all those which are separated or independent, which process, store or transmit data.*”<sup>221</sup> Cyberspace is like a glue that binds many elements of technology together.<sup>222</sup> It is an interactive domain of digital networks that store, modify and communicate information.<sup>223</sup>

---

<sup>218</sup> C Makridis and B Dean (n 217) .

<sup>219</sup> Martin Dodge and Rob Kitchin, *Mapping Cyberspace* (Routledge 2000).

<sup>220</sup> Katarzyna Chałubińska-Jentkiewicz, Filip Radoniewicz and Tadeusz Zieliński (eds), *Cybersecurity in Poland: Legal Aspects* (Springer 2022).

<sup>221</sup> NATO Term ‘Cyberspace’ <<https://nso.nato.int/natoterm>> accessed 18 December 2024.

<sup>222</sup> Jason Whittaker, *The Cyberspace Handbook* (Routledge 2003).

<sup>223</sup> UK Cyber Security Strategy; Oğuz Kaan Pehlivan, *Confronting Cyberespionage Under International Law* (Routledge 2018).

The definition of cyberspace can range from technical to conceptual aspects and it has been claimed as the “fourth domain” after land, sea and air.<sup>224</sup> It is a combination of various expanding cyberspaces that enable digital interaction and communication.<sup>225</sup> According to Kuehl’s definition, cyberspace is “*a global domain within the information environment whose distinctive and unique character is framed by the use of electronics and the electromagnetic spectrum to create, store, modify, exchange, and exploit information via interdependent and interconnected networks using information communication technologies.*”<sup>226</sup> Further, According to Z. Trejnis and P. Trejnis, “*cyberspace encompasses all information and communication means in a collection of networks, techniques, users and digital space, which in turn is assigned three layers: material, logical and informational.*” This definition distinguishes between different layers of cyberspace i.e. the physical network or the hardware, the logical network or the software and the human network.<sup>227</sup> Some common features can be traced across all the definitions of cyberspace. These common features include its non-material nature, unambiguously identified boundaries, decentralization, lack of supervision, universal accessibility, highly accurate real time information processing & calculations and interactive & virtual nature.<sup>228</sup>

In cyberspace, cyber security is a prominent concern due to computer crimes. Cyber security issues are critical not only for firms storing valuable data but also for national security and privacy issues. There is a lack of universally accepted definition for computer crime or cyber-crime. It was defined by the OECD as “*any illegal, unethical, or unauthorized behavior involving the automatic data processing and/or transmission of data*”<sup>229</sup> Interpol has defined computer crime as “*criminal activities in the scope of computer technologies*”. These criminal activities can be categorized as breach of resource access rights, fraud with use of computers, modification of computer resources,

---

<sup>224</sup> Jennifer L Bayuk, *Cyber Security Policy Guidebook* (Wiley 2012).

<sup>225</sup> Martin Dodge and Rob Kitchin, *Mapping Cyberspace* (Routledge 2000).

<sup>226</sup> Oğuz Kaan Pehlivan, *Confronting Cyberespionage Under International Law* (Routledge 2018).

<sup>227</sup> Tomasz Zdzikot, ‘*Cyberspace and Cybersecurity*’ in Katarzyna Chałubińska-Jentkiewicz, Filip Radoniewicz and Tadeusz Zieliński (eds), *Cybersecurity in Poland: Legal Aspects* (Springer International Publishing 2022).

<sup>228</sup> *ibid.*

<sup>229</sup> OECD, ‘*OECD Economic Outlook*’, 1983(1) (Organisation for Economic Co-operation and Development 1983) <[https://www.oecd-ilibrary.org/economics/oecd-economic-outlook-volume-1983-issue-1\\_eco\\_outlook-v1983-1-en](https://www.oecd-ilibrary.org/economics/oecd-economic-outlook-volume-1983-issue-1_eco_outlook-v1983-1-en)> accessed 17 December 2024; Filip Radoniewicz, ‘*Cyberspace, Cybercrime, Cyberterrorism*’ in Katarzyna Chałubińska-Jentkiewicz, Filip Radoniewicz and Tadeusz Zieliński (eds), *Cybersecurity in Poland: Legal Aspects* (Springer 2022) <[https://doi.org/10.1007/978-3-030-78551-2\\_4](https://doi.org/10.1007/978-3-030-78551-2_4)> accessed 17 December 2024.

software reproduction, crime on the internet, hardware & software destruction, etc.<sup>230</sup> As per the UN, cyber-crimes are crimes that are perpetrated using information and communication technology (ICT) and are different from traditional crimes as the geographical boundaries are blurred. Europol has categorized cyber-crime into two groups, i.e. *cyber-enabled* crimes and *cyber dependent* crimes.<sup>231</sup> Cyber enabled crimes involve traditional crimes like fraud that are facilitated by ICT. Cyber dependent crimes are acts where ICT is the target of attack and the confidentiality, integrity and availability of computer data and computer systems are affected.<sup>232</sup> This classification between new and traditional crimes has also been adopted in the Convention of Cybercrime.<sup>233</sup>

Cyber security is therefore a crucial part of smooth functioning of an entity. Cyber security refers to the “*ability to control access to networked systems and the information they contain*”. Cyber security measures can make cyberspace “*reliable, resilient and trustworthy*”.<sup>234</sup> For the private sector, the biggest cybersecurity threat has been the issue of data breach.<sup>235</sup> These threats can be of various types like Viruses, Worms, Trojan Horse Software, Harmful Spyware Attack, Impersonation Attacks, Man-in-the-Middle attack and Denial of Service Attack (DoS).<sup>236</sup> Therefore, a large number of a firm’s cybersecurity measures are focused on preventing data breaches and maintaining confidentiality, integrity and accessibility of data. This is not only driven by the obligation to protect the company from breach of personal data but also driven by self-interest to protect the company’s trade secrets and valuable information. Therefore, data security is a crucial part of a company’s cybersecurity measures<sup>237</sup> as it can enable them to fulfill their goals, take appropriate steps to avoid cyber threats, create discipline, deliver secure services, ensure security of resources and reduce instances of cyber-attacks.

---

<sup>230</sup> Ibid.

<sup>231</sup> Europol, ‘Home’ (*Europol*) <<https://www.europol.europa.eu/home>> accessed 17 December 2024; UNODC, ‘Cybercrime Module 1 Key Issue: Cybercrime in Brief’ (*Sherloc UNODC*, May 2019) <<https://sherloc.unodc.org/cld/en/education/tertiary/cybercrime/module-1/index.html>> accessed 17 December 2024.

<sup>232</sup> *ibid*, UNODC.

<sup>233</sup> Convention on Cybercrime (Adopted 23 November 2001) ETS No 185.

<sup>234</sup> Jennifer L Bayuk, *Cyber Security Policy Guidebook* (Wiley 2012).

<sup>235</sup> Jeff Kosseff, ‘*Positive Cybersecurity Law: Creating a Consistent and Incentive-Based System*’ (2016) 19(2) *Chap L Rev* 401.

<sup>236</sup> Alok Mishra and others, ‘*Cybersecurity Enterprises Policies: A Comparative Study*’ (2022) 22(2) *Sensors* (Basel) 538.

<sup>237</sup> *ibid*.

## 2. Data as Intellectual Capital and How Firms Protect it

At every stage, be it data collection, data processing or data storage, there is a need to take measures to protect data.<sup>238</sup> Since data breach incidents can cause significant economic and monetary losses to corporations, the protection of data is not only driven by the motivation to protect the privacy of the stakeholders but also by self-driven interests in the market. A majority of the cyber-crimes committed against firms are driven by an intention to commit industrial espionage and extract financial value from valuable sensitive data.<sup>239</sup>

The sum of all the information or data held by a company and all the knowledge of a company that provides them with competitive advantage in the market is considered as Intellectual Capital. Intellectual Capital includes “*intellectual material, knowledge, expertise, intellectual property and information*” and can help in value creation.<sup>240</sup> The use of big data provides a substantial competitive edge to companies and big data analysis can create value and “support new intangible assets”. Research by the McKinsey Global Institute has shown that big data is a “*driver for innovation, productivity and growth*”.<sup>241</sup> Considering how valuable big data is to companies, it can be considered an Intellectual Capital. Now, one of the key challenges that companies face is protecting their valuable data and information from cyber-attacks.<sup>242</sup> The majority of incidents of cyber threats and cyber-attacks are intended as industrial espionage.<sup>243</sup> Stealing of data and buying hacked data has become an easy task and is essentially less risky than traditional physical theft.<sup>244</sup> The ease with which data can be stolen has thus created a “*marketplace for data on the dark web*” and has resulted in the creation of a “*hidden data economy*”.

---

<sup>238</sup> Perera C and others, ‘Big Data Privacy in the Internet of Things Era’ (2015) 17(3) IT Professional 32.

<sup>239</sup> Matteo La Torre, John Dumay and Michele Antonio Rea, ‘Breaching Intellectual Capital: Critical Reflections on Big Data Security’ (2018) 26(3) Meditari Accountancy Research 463.

<sup>240</sup> John Dumay, ‘A critical reflection on the future of intellectual capital: from reporting to disclosure’ (2016) 17(1) Journal of Intellectual Capital <<https://www.emerald.com/insight/content/doi/10.1108/jic-08-2015-0072/full/html>> accessed 18 December 2024 (168-184.). See also *ibid*.

<sup>241</sup> McKinsey Global Institute, ‘Big Data: The Next Frontier for Innovation, Competition, and Productivity’ (May 2011).

<sup>242</sup> La Torre (n 240).

<sup>243</sup> Verizon, ‘2017 Data Breach Investigations Report’ (Verizon, 2017) <[www.verizonenterprise.com/verizon-insights-lab/dbir/2017/](http://www.verizonenterprise.com/verizon-insights-lab/dbir/2017/)> accessed 18 December 2024.

<sup>244</sup> *ibid*.

The security of big data is related to three key qualities of data i.e. confidentiality, integrity and availability. Any incident that can affect any of these three qualities will affect the value of big data. For instance, unauthorized access to data can lead to theft, modification, alteration and deletion of data and the same significantly affects the integrity and reliability of big data. Such incidents can impact the value of the information and knowledge that could have been extracted from the big data prior to the attack. The ISO has chalked out aims of “*information security management system*” and it is to preserve the “*confidentiality, integrity and availability of information by applying a risk management process and give confidence to interested parties that risks are adequately managed*”.<sup>245</sup>

Firms deploy various ways and processes to understand and analyze their competitors, suppliers and customers. This is their competitive intelligence and it is used to increase the competitive advantage of the company and reduce the competitive advantage for other competitors.<sup>246</sup> One of the biggest issues firms face while extracting value from their datasets is the issue of data theft. Competitive data theft can be divided into two categories. Firstly, proprietary technology data that is related to creation of goods and services. Secondly, tactical data which is data that supports companies in making decisions and creating strategies. Theft of the former category of data can aid competitors in replicating the products and services of the attacked company and theft of the latter can provide information about the company's commercial decisions.<sup>247</sup> A cyber security incident can have multiple potential harms. This can include loss of sales where due to the illegally obtained information competitors can improve the quality of their information, reduce product costs, poach sales, etc. Therefore, scholars suggest deploying various policies like “*privacy policy, website security policy, cloud computing security policy, email security policy*”, etc. in order to ensure data integrity, confidentiality and accessibility.<sup>248</sup>

---

<sup>245</sup> ISO/IEC, *ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection — Information security management systems — Requirements* (ISO 2022).

<sup>246</sup> <https://media.proquest.com/media/hms/OBJ/JLVSV?s=yV1hfKvWGrZf0ZHdD48ITKcXkys%3D>

<sup>247</sup> Allan A Friedman, Austen Mack-Crane, and Ross A Hammond, ‘Cyber-enabled Competitive Data Theft: A Framework for Modeling Long-Run Cybersecurity Consequences’ (*Brookings Economics studies*, 6 December 2013) <<https://www.brookings.edu/articles/cyber-enabled-competitive-data-theft-a-framework-for-modeling-long-run-cybersecurity-consequences/>> accessed 17 December 2024.

<sup>248</sup> Alok Mishra and others, ‘Cybersecurity Enterprises Policies: A Comparative Study’ (2022) 22 *Sensors* (Basel) 538.

## 2.1. Cyber Security Threats

Information theft or espionage has been around since the 440 AD but in the era of the internet, it has taken an entirely new form. Cyber espionage has been termed as the greatest transfer of wealth in history.<sup>249</sup> The three layers of cyberspace i.e. the physical layer, logical layer and the social layer plays an important role for cyber espionage. The physical layer can facilitate cyber espionage in situations where a code is inserted in hardwares and that can enable remote access of the system. The logical layer of cyberspace can be infected with malware and that can provide unauthorized access. Lastly, the social layer can facilitate espionage by the use of techniques like phishing and can give access to credentials which can eventually give authorized access.

There are various ways to extract information through the use of ICT. Instances have been reported where “internet buffers” were used to watch live and stored data by attaching intercept probes to fiber-optic cables. In 2013, it was reported that data flows across fiber-optic cables transferring information to data centers can also be copied. Further, backdoors can be secretly installed in computer equipment by entering unauthorized codes that permit unauthorized access to decrypt information passing through the computer.<sup>250</sup>

With the development of advanced ICT, malware has also become advanced and stronger. Malware is malicious software designed to do malicious actions in computer networks and systems. There are various types of malware, some examples include viruses, worms, spyware, trojans, etc.<sup>251</sup> Malware attacks can cause damages like information leak, financial loss, privacy loss, facilities destruction, etc. Once information is acquired illegally, it is marketed anonymously on the deep web.<sup>252</sup> Data theft has created a marketplace on the dark web or a “hidden data economy” where hackers can buy and sell stolen information.<sup>253</sup>

---

<sup>249</sup> Richard Rivera and others, ‘An Analysis of Cyber Espionage Process’ in Álvaro Rocha, Carlos Hernan Fajardo-Toro and José María Riola Rodríguez (eds), *Developments and Advances in Defense and Security* (Springer 2022).

<sup>250</sup> Emily Price, ‘Juniper Networks security flaw may have exposed US government data | Hacking’ (*The Guardian*, 22 December 2024); Evan Perez and Shimon Prokupez, ‘First on CNN: Newly discovered hack has U.S. fearing foreign infiltration’ (*CNN Politics*, 19 December 2015).

<sup>251</sup> Rivera (n 250).

<sup>252</sup> *ibid.*

<sup>253</sup> Matteo La Torre, John Dumay and Michele Antonio Rea, ‘Breaching Intellectual Capital: Critical Reflections on Big Data Security’ (2018) *Meditari Accountancy Research* <https://doi.org/10.1108/MEDAR-02-2018-0267>.

The process of a malware attack has been categorized by Richard Rivera et al. in a few phases.<sup>254</sup> The first phase is Reconnaissance where the attacker studies the victim to collect useful information and identifies its vulnerabilities. The second phase is the Preparation phase where depending on the objective of the attack, different techniques are used. The third phase is the Attack phase where the attacker attacks and gets unauthorized access to the victim's system. The fourth phase is Information Gathering and the attacker needs to know the type of information he is looking for. The fifth phase is the Maintenance phase where the attacker analyses if there is a need to adapt the current attack or perform more attacks. The next phase is the phase of Information Leakage which can happen simultaneously with or after the fifth phase. The attacker can transmit the information through proxy networks or the deep web. In the seventh phase, the stolen information or data is sold at a lower price. The last phase is the phase where the attacker finishes information gathering and escapes the system. During the escape, intruders also leave a window open so that their next attack in the future can be easy.<sup>255</sup> For example, the Duqu malware attack was deployed through targeted emails and it spread in the system using MS Word Document. Once the system was infected with this malware, it enabled the intruders to not only extract sensitive information but also extract information that could help them in performing future attacks. The Duqu malware attack was primarily deployed for espionage purposes across 20 countries around the world.<sup>256</sup> The success of malware attacks like Duqu was attributed to the lack of efficiency in defense mechanisms. Developing a strong defense mechanism against attacks necessarily requires joint industry wide research and special expertise.<sup>257</sup> Individual firms need to cooperate and share information about attacks to encourage the development of industry tools and techniques to prevent malware attacks. Government to government cooperation is also necessary to tackle cross border cyber espionage attacks.<sup>258</sup> Sun Tzu notes that Japan has created a harmonious corporate culture which can help in tackling security issues. Unlike Japan, the US has taken an individualistic

---

<sup>254</sup> Rivera (n 250).

<sup>255</sup> *ibid.*

<sup>256</sup> Wangen G, 'The Role of Malware in Reported Cyber Espionage: A Review of the Impact and Mechanism' (2015) 6 *Information* 183; B Bencsáth, G Pék, L Buttyán and M Félegyházi, 'Duqu: Analysis, Detection, and Lessons Learned' (Proceedings of the ACM European Workshop on System Security (EuroSec), Bern, 10 April 2012) <<https://crysys.hu/publications/files/BencsathPBF12eurosec.pdf>> accessed 17 December 2024.

<sup>257</sup> Bencsáth (n 257).

<sup>258</sup> Peter R J Trim, 'Counteracting Industrial Espionage through Counterintelligence: The Case for a Corporate Intelligence Unit and Collaboration with Government Agencies' (2002) 15 *Security Journal* 7.

approach that has made it difficult to manage corporate goals and industrial espionage. The Japanese approach should be followed to protect competitive and economic intelligence.<sup>259</sup>

## 2.2. Data Breach and Its Impact

A data breach is described as “*any security incident where any party gains unauthorized access to sensitive data or confidential information*”.<sup>260</sup> It has also been defined as “*compromise of security that leads to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to protected data transmitted, stored or otherwise processed.*”<sup>261</sup> The loss of industrial information through cyber espionage has been said to be the “greatest transfer of wealth in history”.<sup>262</sup> According to IBM, data breaches may be caused by innocent mistakes, malicious insiders or hackers. Further, data breaches can be motivated by various objectives but most data breaches involving malicious intent are motivated by financial gains. As per the latest Cost of Data Report 2023, the global average cost for data breach was USD 4.45 million, 2.3% increase from the year 2020. Since 2020, the incidents of data breach cost has increased by 53.3%. Furthermore, according to the Cost of Data Report 2020, data breaches follow a pattern i.e. finding a weak target, identify the method of breach, launch the attack and then exfiltrate, sell, lock or destroy the data acquired unlawfully.<sup>263</sup>

In cyber security breaches, the attacker gets unauthorized access to valuable information including employee and customer’s personal information, company accounts, etc.<sup>264</sup> Multiple studies have been conducted to analyze the impact of data breaches on firms. However, there is a lack of consensus regarding the adequacy and reliability of the methods used to calculate the exact cost of data breach incidents.<sup>265</sup> There is also a lack of consensus on the cost factors to be considered to measure the costs as there can be multiple types of harms like physical harm, monetary and

---

<sup>259</sup> *ibid.*

<sup>260</sup> Matthew Kosinski, ‘What is a data breach?’(IBM, 24 May 2024) <<https://www.ibm.com/topics/data-breach>> accessed 17 December 2024.

<sup>261</sup> ISO/IEC, *ISO/IEC 27040:2024 Information technology — Security techniques — Storage security* (ISO 2024).

<sup>262</sup> Makridis C and Dean B, ‘Measuring the Economic Effects of Data Breaches on Firm Outcomes: Challenges and Opportunities’ (2018) 43 *Journal of Economic and Social Measurement* 59.

<sup>263</sup> IBM, *Cost of a Data Breach Report 2022* (IBM, 2022).

<sup>264</sup> P Wang, H D’Cruze, D Wood, ‘Economic Costs and Impacts of Business Data Breaches’ (2019) 20 *Issues In Information Systems* 162.

<sup>265</sup> *ibid.*

economic harm, social harm as well as direct and indirect costs like loss of corporate reputation and customers.<sup>266</sup> Furthermore, the hidden costs are not easy to detect and can negatively affect business, competition, reputation and recovering from these costs can be expensive and time consuming.<sup>267</sup> The very first study on the cost of data breach was done by Anderson et al. in 2012 wherein the cost of data was categorized into direct loss, indirect loss, defense cost and the cost to society. The direct loss is the economic and monetary loss to the attacked party and includes loss of time and effort put into recovering from the attack. The indirect loss is the loss of consumer trust and loss of business opportunities. The defense cost includes costs of developing and deploying security prevention measures and costs of law enforcement. Lastly, the cost to the society is direct, indirect and defense costs combined. This study has also suggested that often the indirect costs of cybercrimes are relatively higher than the direct costs incurred by traditional crimes.<sup>268</sup>

In another study, the cost of data breach was categorized into two categories i.e. direct and indirect financial costs, reputational costs and associated costs. The study also states that the direct and indirect financial costs can be measured but measuring the reputation cost is challenging.<sup>269</sup> Data breaches can significantly impact the reputation of the attacked party and its financial performance. Reputation is the “*overall opinion about a firm by customers, investors, employees and the general public*” and can act as a competitive advantage for a company.<sup>270</sup> A breach incident can entail a crisis for businesses and interrupt daily business and influence the consumer's perception of the company.<sup>271</sup> Reputational damage can make consumers switch to other competitors and can make it difficult for businesses to attract new customers.<sup>272</sup> In addition, another study has shown that data privacy breach can negatively impact the market value of the firm. It has shown that the

---

<sup>266</sup> *ibid.*

<sup>267</sup> IBM Report (n 264); *ibid.*

<sup>268</sup> Wang (n 266).

<sup>269</sup> Krausz & Walker, 2013.

<sup>270</sup> Griselda Sinanaj and Humayun Zafar, ‘Who Wins In A Data Breach? - A Comparative Study On The Intangible Costs Of Data Breach Incidents’ (Proceedings of the Pacific Asia Conference on Information Systems (PACIS), Chiayi City, Taiwan, 27 June–1 July 2016) <<https://aisel.aisnet.org/pacis2016/60>> accessed 17 December 2024.

<sup>271</sup> *ibid.*

<sup>272</sup> Cavusoglu H, Mishra B and Raghunathan S, ‘The Effect of Internet Security Breach Announcements on Market Value: Capital Market Reactions for Breached Firms and Internet Security Developers’ (2004) 9 International Journal of Electronic Commerce 70.

market value of smaller firms is more affected by data breach incidents as compared to larger firms.<sup>273</sup> However, on the contrary, it has also been shown that the value of the firms increases abnormally due to the perception that as a result of the breach, the attacked firms would invest more in protection and prevention measures.<sup>274</sup>

Data breach affects the integrity and availability of a dataset and this eventually affects the quality of data. Data gets compromised if the intruder steals, modifies, alters or deletes the data. For factual accuracy in big data analysis and results, it's important to ensure that big data remains uncompromised and clean. Malicious attacks compromising the integrity of data (data sabotage) is a great cause of concern and can cause financial losses to the data holder because it can result in wrong decision making. Data sabotaging or maliciously compromising the accuracy of data can cause graver damage than data theft or destruction as data sabotages can remain undetected.<sup>275</sup>

### **3. Overview of Indian Data Protection Laws**

With the increase in the value of data for firms, theft of corporate data has quickly become an issue. Firms hold various types of data ranging from personal information of customers & employees to confidential corporate information. There is no guarantee that confidential data will remain protected at all times. Considering that the risk of data breach is very plausible, laws must be in place to protect against data theft and provide remedies for the same.

The Indian Penal Code (IPC) defines “theft” under section 378 as “*whoever, intending to dishonestly take any moveable property out of the possession of any person without that person’s consent, moves that property in order to such taking, is said to commit theft.*”<sup>276</sup> Further, “moveable property” has been defined to include corporeal object or property.<sup>277</sup> The question that arises here is whether theft of data/information can be covered under the traditional definition of

---

<sup>273</sup> Tripathi M and Mukhopadhyay A, ‘Financial Loss Due to a Data Privacy Breach: An Empirical Analysis’ (2020) 30 Journal of Organizational Computing and Electronic Commerce 381.

<sup>274</sup> *ibid.*

<sup>275</sup> Matteo La Torre, John Dumay and Michele Antonio Rea, ‘Breaching Intellectual Capital: Critical Reflections on Big Data Security’ [ 2018] 26 Meditari Accountancy Research 463.

<sup>276</sup> Indian Penal Code 1860, s 378.

<sup>277</sup> Section 378, Indian Penal Code, 1860 (Act 45 of 1860).

theft as provided under IPC. The IPC was introduced back in 1860 and it was not made keeping in mind theft of information or data. It is therefore argued that data theft can't be covered under IPC's definition of theft for two reasons. Firstly, data is not a moveable property as defined under IPC as it is not a corporeal property. Secondly, even if data were to be considered as moveable property, data theft does not necessarily deprive the owner of the possession of data since data is transferred or copied.<sup>278</sup> The Calcutta High Court in *Adventz Investments and Holdings Limited & Ors. vs. Birla Corporation & Anr.* noted that information is not a moveable property according to section 22 of the IPC as it includes corporeal property. Further, the Court also noted that to fall under the ambit of section 378 of the IPC, the property “must be capable of being taken out of the possession of the person who has the right to retain it”.<sup>279</sup>

Despite the fact that data theft is not recognised under IPC, the provisions of the Indian Penal Code may come into play to protect data. Section 405 of the IPC provides for criminal breach of trust and penalizes dishonest misappropriation of property. In *Jagjeet Singh v. State of Punjab & Anr.*, the Court was dealing with a case of data theft by employees.<sup>280</sup> The Court observed that the instances of hacking and data theft are offenses not only limited to the IT Act but also fall within the ambit of the Indian Penal Code. Therefore, the application of the IT Act does not imply that acts of data theft fall outside the scope of the IPC. In the cited case, since there was a criminal breach of trust and theft of data by an employee, section 405 and 381 of the IPC and section 66 of the IT Act was applied. Section 405 of the IPC deals with criminal breach of trust. It provides that “*whoever, being in any manner entrusted with property, or with any dominion over property, dishonestly misappropriated or converts to his own use that property, or dishonestly uses or disposes of that property in violation of any direction of law prescribing the mode in which such trust is to be discharged, or of any legal contract, express or implied, which he has made touching the discharge of such trust, or willfully suffers any other person so to do, commits “criminal breach of trust”*”. The offense of criminal breach of trust is a non bailable and a cognizable offense.

---

<sup>278</sup> Bharat Chug, ‘Data Theft And The Indian Criminal Law’ (*Live Law*, 26 August 2018) <<https://www.livelaw.in/data-theft-and-the-indian-criminal-law/>> accessed 17 December 2024.

<sup>279</sup> *Adventz Investments and Holdings Limited & Ors. vs. Birla Corporation & Anr*, AIR 2019 SC 2390.

<sup>280</sup> *Jagjeet Singh v. State of Punjab & Anr* [Special Leave Petition (Criminal) No. 3583 of 2021].

The Information Technology Act, 2000 and the IT Rules provide for certain provisions for the protection of data privacy in India. The IT Act is primarily based on the United Nations Model Law on Electronic Commerce. It was enacted to give legal recognition to electronic transactions, facilitate electronic filing of documents with government agencies and to penalize cyber offenses. The Act defines cyber security as “means of protecting computer devices & resources and information stored therein from unauthorized access, use, disclosure, disruption, modification or destruction”.<sup>281</sup> This protected information includes “data, images, codes, computer programmes, software and databases,” etc.<sup>282</sup> The Act also defines data as “*representation of information, knowledge, facts, concepts or instructions which are being prepared or have been prepared in a formalized manner, and is intended to be processed, is being processed or has been processed in a computer system or computer network.*”<sup>283</sup> This data may be stored in any form. If any person, without the permission of the person in-charge of a computer system and resources: accesses computer system, downloads/extracts data from the computer system, introduces virus to the system, damages the system and its resources, disrupts or causes disruption to the system, denies authorized person to access the system, facilitates unauthorized access to the system for another person, destroys, alters or steal information stored on a computer system.<sup>284</sup> The Act also defines access “*with its grammatical variations and cognate expressions*” to mean “gaining entry into, instructing or communicating with the logical, arithmetical, or memory function resources of a computer, computer system or computer network.” Any person who does any of the above mentioned prohibited acts fraudulently or dishonestly, will be punished with imprisonment up to three years or with fine up to five lakh rupees or with both. Thus, the IT Act has enumerated and penalizes certain acts with respect to computer and computer resources. Further, section 43A of the Act provides protection of sensitive personal information.<sup>285</sup> It states that if a body corporate possessing, dealing or handling sensitive personal data, causes wrongful loss or gain to any person due to its negligence in maintaining security measures to protect the data will be held liable under the Act.

---

<sup>281</sup> Information Technology Act, s 2(nb).

<sup>282</sup> *ibid* s 2(v).

<sup>283</sup> *ibid* s 2(o).

<sup>284</sup> *ibid* s 43.

<sup>285</sup> *ibid*.

Section 70B of the Act establishes an agency called Computer Emergency Response Team (CERT) that serves as the national agency that responds to incidents related to cyber security.<sup>286</sup> The CERT aims to provide a central point of contact for reporting cyber issues and publishes information on cyber threats. It provides 24/7 assistance in preventing and managing cyber security incidents. CERT analyzes cyber incident trends and patterns to develop preventive strategies.<sup>287</sup> CERT responded to 13,91,457 cyber security incidents ranging from website intrusion, malware attacks, phishing attacks, ransomware attacks, data breach, etc in the year 2022 alone.<sup>288</sup> The Information Technology (Computer Emergency Response Team and Manner of Performing Functions and Duties) Rules, 2013 define cyber security breach as “*unauthorized acquisition or unauthorized use by a person as well as an entity of data or information that compromises the confidentiality, integrity or availability of information maintained in a computer resource.*”<sup>289</sup> The CERT-IN is responsible for addressing all types of cyber security incidents that occur in the country but the level of support it can provide can vary depending on the nature, type and severity of the incident. The Rules provide a list of priority incidents (in a decreasing order) and resources are assigned based on this priority list. These incidents are: (i) threats to physical safety of human beings, (ii) incidents of severe nature like denial of service, intrusion, contamination, etc. on the public information infrastructure, (iii) large scale incidents like identity theft or intrusion into computer resource, (iv) compromise of individual user accounts on multi-user systems, (v) any other type of incident depending on its severity and extent.<sup>290</sup>

Therefore, even if corporate data theft cannot be covered under the traditional definition of theft provided by the IPC, such instances of cyber security are covered by the IT Act and Rules. With the recent enactment of the Digital Personal Data Protection (DPDP) Act, India has taken the first

---

<sup>286</sup> *ibid* s 70B.

<sup>287</sup> CERT-in, ‘Indian - Computer Emergency Response Team’ (*CERT-in*, 2022) <<https://www.cert-in.org.in/>> accessed 17 December 2024.

<sup>288</sup> CERT-in, ‘CERT IN Annual Report 2022’ (CERT-in 2022).

<sup>289</sup> The Information Technology (Computer Emergency Response Team and Manner of Performing Functions and Duties) Rules, 2013, r 2(i).

<sup>290</sup> The Information Technology (Computer Emergency Response Team and Manner of Performing Functions and Duties) Rules, 2013, r 11.

step towards protecting data.<sup>291</sup> The DPDP Act provides that data fiduciaries must protect personal data under their possession by taking “reasonable security safeguards” to prevent breach.<sup>292</sup> Any contravention of the Act will attract penalties under section 33 of the Act. By enacting a law specifically designed to protect personal data, India has moved towards improving the data protection regime. However, the possibility of data breaches can not be negated. Consequently, there is still a need to create measures to minimize the impact of data breaches and to create robust mechanisms that can minimize the possibility of cyber security threats/attacks.

#### **4. Remedies for Data Leak and Breach**

Criminalizing an act requires a justification. John Mill’s harm principle is often looked at by criminal theorists in context of questions of criminalization. According to the harm principle, “*the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others.*” Criminalizing data theft can primarily be justified using the harm principle. With the rapid development of digital technology, computers and data have become crucial for society and it has become necessary to protect information. With the increased accessibility and portability of information, it has become easier to gain unauthorized access to computers or data stored therein and the same can cause tremendous amounts of damage since a large volume of information is stored and shared. Criminal provisions and sanctions can act as a deterrent for offenders and can provide a sense of security to the users.

The European Convention on Cyber Crime, a binding multinational instrument was introduced to address the issue of cybercrime. It is the only international legally binding document that deals with cybercrime. It provides that every contracting party shall take measures to establish criminal offenses when a computer is accessed in an unauthorized manner intentionally.<sup>293</sup> Article 4 of the Convention mandates parties to criminalize intentional damage, deletion, deterioration, alteration or suppression of computer data without right. The Convention was adopted to address the urgent need to “*deter action directed against the confidentiality, integrity and availability of computer*

---

<sup>291</sup> Anirudh Burman, ‘Understanding India’s New Data Protection Law’ (2023) Carnegie Endowment for International Peace <<https://policycommons.net/artifacts/4941168/understanding-indias-new-data-protection-law/5770466/>> accessed 17 December 2024.

<sup>292</sup> Digital Personal Data Protection Act, 2023, s 8(5).

<sup>293</sup> Convention on Cybercrime (Adopted 23 November 2001) ETS No 185, art 2.

*systems, networks and computer data*” as well as their misuse”<sup>294</sup> It is notable that the Convention recognises that confidentiality and integrity of data needs to be protected by criminal law.

However, in the context of digital crimes, the question arises regarding the interpretation of the term ‘access’ in cases of unauthorized access to computer or computer data. This has been interpreted differently in different jurisdictions.<sup>295</sup> In the US, the Computer Fraud and Abuse Act of 1986 (CFAA) makes it an offense to access computers without authorization. The Defendants are charged under the Act for unauthorized access only when “the defendant was not authorized to access the protected computer under any circumstances by any person or entity with the authority to grant such authorization; the defendant knew of the facts that made the defendant’s access without authorization; and prosecution would serve the Department’s goals for CFAA enforcement.”<sup>296</sup> The US has adopted a narrow approach in *State v Allen*<sup>297</sup> where it interpreted ‘access’ as “the ability to make use of the computer in the sense of obtaining the use of programs or data.” The Court further noted that the intention and objective of the law is to criminalize misuse of computer systems and not “merely approaching a computer”. In *State v Riley*<sup>298</sup>, ‘access’ was interpreted as approaching or otherwise making use of any resources of a computer, directly or by electronic means.

The UK and Australia have given the term ‘access’ a broader interpretation. Both the countries have criminalized unauthorized access to computer data and not the computer. As per the Criminal Code Act 1995, Australia has made it an offense “for a person to cause any unauthorized access to data held in a computer, knowing the access is unauthorized, and by that access intending to commit, or facilitate the commission of a serious offense”.<sup>299</sup> The offense can attract imprisonment of 5 years or longer. The Act has also defined data to include “any information in any form”.<sup>300</sup>

---

<sup>294</sup> *ibid*, art 2.

<sup>295</sup> Jonathan Clough, ‘Data Theft? Cybercrime and the Increasing Criminalization of Access to Data’ (2011) 22 Criminal Law Forum 145.

<sup>296</sup> US Department of Justice, ‘9-48.000 – Computer Fraud and Abuse Act’ (*Justice Manual, US Department of Justice*, May 2022) <https://www.justice.gov/jm/jm-9-48000-computer-fraud-and-abuse-act> accessed 17 December 2024.

<sup>297</sup> 260 Kan 107 (1996).

<sup>298</sup> 846 P 2d 1365 (Wash 1993).

<sup>299</sup> Criminal Code 2002, s 415.

<sup>300</sup> *ibid*, s 412.

Notably, it is also an offense to access restricted data without authorization. In the UK, the Computer Misuse Act 1990 imposes punishment on a person if “(a) *he or she causes a computer to perform any function with intent to secure access to any program or data held in any computer or to enable any such access to be secured; (b) the access he or she intends to secure or to enable to be secured is unauthorized; and (c) he or she knows at the time when he or she causes the computer to perform the function that that is the case*”.<sup>301</sup> The term “causes a computer to perform a function” keeps up with the rapid technological development. Since it is fairly easy for hackers to get unauthorized access to a computer or data without physically accessing it, the term used in the provision makes the law appropriate to the current fast paced development. Section 2 of the Act criminalizes the offense committed under section 1 with the intent to commit or facilitate access of other or further offenses.<sup>302</sup> Furthermore, if a person gains unauthorized access with the intent to impair the operation of a computer system, hinder access to program or data of the computer, impair the operation of program or data of the computer or enables these acts to be done would be guilty under the Act.<sup>303</sup>

#### **4.1. Risk Implementation and Breach Notification**

The law on data security and the prevention of data leak and breaches is focused on risk implementation of safeguards responses once data gets leaked or compromised. The underlying agenda is to decrease the likelihood of harm in cases of data breach since it is apparent that there can be no perfect data protection.<sup>304</sup> This approach has been adopted across many jurisdictions. For example, in Australia, when an entity becomes aware of any material information that could potentially affect the value of its securities, under the Australian Stock Exchange Rules, it must inform the ASX.<sup>305</sup> Similarly, the Critical Infrastructures Act puts an obligation on entities when they become aware of a cyber-security breach or an incident that has occurred or is about to occur which has affected or is likely to affect critical infrastructure, it must inform the appropriate

---

<sup>301</sup> *ibid*, s 1.

<sup>302</sup> *ibid*, s 2.

<sup>303</sup> *ibid*, s 3.

<sup>304</sup> Ira Rubinstein and Woodrow Hartzog, 'Anonymization and Risk' (2015) 91 *Washington Law Review* 703, NYU School of Law Public Law Research Paper No 15-36 <https://ssrn.com/abstract=2646185> accessed 30 January 2025.

<sup>305</sup> Corporations Act 2001, s 674 read with ASX Listing r 3.1 and 3.1A.

authorities immediately or within 72 hours of becoming aware.<sup>306</sup> Further, as per the “Telecommunications Sector Security Reforms (TSSR) Administrative Guidelines”, the Home Affairs Ministry can direct Carrier, carriage service provider (C/CSP) to take action if there is a risk of security breach that could impact confidentiality, availability and the integrity of the information. The Minister is also entitled to initiate proceedings in case of non-compliance, followed by civil remedies.<sup>307</sup> Another example could include China’s Non-personal Data Security Breach Notification Requirements that provide that a summary report must be submitted when a cyber-security breach or incident has taken place.<sup>308</sup> This applies to key institutions and operation institutions. The timeline for reporting breaches depends on the seriousness of the incident. The incident must be reported within five days of discovery if it is an ordinary security incident, 12 hours if it is a major security incident and 2 hours if it is an extremely serious incident. In Germany, an entity is required to notify the Federal Office for Information Security if it comes to the knowledge of any incident that compromises the security of telecom networks. The requirement covers incidents affecting the availability, integrity, authenticity and confidentiality of IT systems and that could adversely affect the operability of a critical structure.<sup>309</sup>

Similar risk control and management approach can be seen for protection of personal data. For example, in the EU, if a company or an organization suffers from a breach of confidentiality or integrity of data it holds and the same is likely to breach an individual’s right to privacy, the supervisory authorities have to be notified without delay, of no later than 72 hours.<sup>310</sup> If the data breach is likely to pose high risk to individuals, then the concerned individuals have to be informed

---

<sup>306</sup> Security of Critical Infrastructure Act 2018, s 24.

<sup>307</sup> Cyber and Infrastructure Security Centre, ‘Telecommunications Sector Security Reforms (TSSR) Administrative Guidelines’ (2020) Australian Government Department of Home Affairs <[https://www.cisc.gov.au/resources-subsite/Documents/tss\\_administrative\\_guidelines.pdf](https://www.cisc.gov.au/resources-subsite/Documents/tss_administrative_guidelines.pdf)> accessed 17 December 2024.

<sup>308</sup> Baker McKenzie, ‘Breach Notification Requirements’ (Baker *McKenzie Resource* Hub, 1 January 2024) <<https://resourcehub.bakermckenzie.com/en/resources/global-data-privacy-and-cybersecurity-handbook/asia-pacific/china/topics/breach-notification-requirements>> accessed 18 December 2024.

<sup>309</sup> Deutsche Telekom, ‘Reporting of data breaches’ (Deutsche Telekom, February 2022) <<https://www.telekom.com/resource/blob/594868/f752115910ce527a54869d4206b4a668/dl-meldung-datenschutzvorfaelle-en-data.pdf>> accessed 17 December 2024.

<sup>310</sup> The General Data Protection Regulation (GDPR), art 33.

too.<sup>311</sup> Any person who suffers damage has the right to receive compensation from the data controller.<sup>312</sup>

India has also taken an approach focusing on risk assessment and providing safeguards to minimize harms followed by data breach incidents and has imposed obligations on entities to notify about data breach and the Information Technology (The Indian Computer Emergency Response Team and Manner of Performing Functions and Duties) Rules, 2013 provide that CERT, an agency providing 24 hour services<sup>313</sup> will provide services like responding to cyber security incidents, predict, prevent and analyze cyber security incidents, train or upgrade technical know-how, scan cyber space for any vulnerabilities, breach and malicious activities, etc.<sup>314</sup> In 2022, CERT-In issued directions to service providers. These directions apply to intermediaries, data centers, body corporate and government organizations and they need to inform and report to CERT-In about cyber security incidents within 6 hours. The informant may be required by CERT-In to take actions to respond to cyber security incidents or assist CERT-In.<sup>315</sup> The Information Technology (Guidelines for Intermediaries and Digital Media Ethics Code Rules, 2021) also provides that intermediaries are required to implement cyber security measures to protect security and report any incidents of breach of cyber security to CERT-In.<sup>316</sup> The time limit within which cyber security incidents must be reported to CERT-In is 6 hours. Previously, the requirement was to only report within “a reasonable time period”. The six hour time limit is significantly stricter than requirements adopted around the world and it was both criticized and appreciated. The Ministry of Electronics and Information Technology (MeitY) pointed out that the nature of cyber-crimes has become more complex over the years and therefore, there is a need to ensure that all cyber security incidents are

---

<sup>311</sup> *ibid*, art 34.

<sup>312</sup> *ibid*, art 82.

<sup>313</sup> Information Technology (The Indian Computer Emergency Response Team and Manner of Performing Functions and Duties) Rules, 2013, r 5 <<https://www.cert.in.org.in/Downloader?pageid=5&type=2&fileName=CIPS-2018-0113.pdf>> accessed 17 December 2024.

<sup>314</sup> *ibid*, r 9.

<sup>315</sup> Government of India, Ministry of Electronics and Information Technology (MeitY), Indian Computer Emergency Response Team (CERT-In), ‘Directions under Sub-Section (6) of Section 70B of the Information Technology Act, 2000 Relating to Information Security Practices, Procedure, Prevention, Response and Reporting of Cyber Incidents for Safe & Trusted Internet’ (28 April 2022) No. 20(3)/2022-CERT-In <[https://www.cert-in.org.in/PDF/CERT-In\\_Directions\\_70B\\_28.04.2022.pdf](https://www.cert-in.org.in/PDF/CERT-In_Directions_70B_28.04.2022.pdf)> accessed 18 December 2024.

<sup>316</sup> The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021, s 3(1)(l)

reported at the earliest so prompt action can be taken by CERT-In.<sup>317</sup> Stakeholders opposing the 6 hour deadline state that it is not in line with global best practices. For example, the GDPR provides for a three day window to report incidents. Further, the shorter deadline was also criticized as this change in the deadline will create compliance costs for firms as they will have to adapt their internal procedures to comply with it.<sup>318</sup>

The blanket requirement of reporting all types of cyber security incidents within 6 hours is said to be unreasonable. It is instead suggested by experts that the nature and extent of the cyber-attack along with the size of the entity attacked should be considered while deciding on the timeframe for the reporting requirement.<sup>319</sup> It should also be noted that the CERT-In Directions do not provide for follow up requirements. For example, the GDPR provides that in cases of cyber security incidents where it is not possible to provide information within the stipulated time frame, it may be provided in phases without unnecessary delay.<sup>320</sup> This is absent from the CERT-In Guidelines but CERT-In FAQs clarify that additional information may be reported later. It is very plausible that relevant information can come up regarding the breach post the 6 hour window, therefore, CERT-In Directions should also include follow-up/additional reporting requirements.

The CERT-In Directions have listed the types of incidents that need to be mandatorily reported. The list has been expanded to include more incidents as compared to the old 2014 CERT-In Directions. However, the list was very broad and the Directions failed to define these incidents. For example, terms like data breach, data leak, etc. have not been defined and it is unclear as to why data breach and data leak have been listed separately. The Ministry of Electronics and Information Technology (“MeitY”) then released a list of FAQs to resolve some of the recurring

---

<sup>317</sup> Vinod Joseph, Aryan Mohindroo and Anushkaa Shekhar, ‘Cert-in’s Six Hour Reporting Rule for Cyber Security Incidents - Statutory Interpretation and Analysis’ (*Argus Partners*, 29 September 2024) <<https://www.argus-p.com/papers-publications/thought-paper/cert-ins-six-hour-reporting-rule-for-cyber-security-incidents-statutory-interpretation-and-analysis/>> accessed 18 December 2024.

<sup>318</sup> Novojuris Legal, ‘6 Hours To Report Cyber Incidents To CERT-In’ (*Mondaq*, 13 January 2023) <<https://www.mondaq.com/india/it-and-internet/1270332/6-hours-to-report-cyber-incidents-to-cert-in>> accessed 18 December 2024.

<sup>319</sup> Tejasi Panjiar, Anushka Jain and Prateek Waghre ‘CERT-In Directions on Cybersecurity: An Explainer’ (Internet Freedom Foundation, 5 May 2022) <<https://internetfreedom.in/cert-in-guidelines-on-cybersecurity-an-explainer/>> accessed 18 December 2024.

<sup>320</sup> The General Data Protection Regulation (GDPR) art 33(4).

concerns and issues regarding the Directions.<sup>321</sup> The FAQs, among other things, have elaborated on the type of incidents that need to be mandatorily reported. However, the stakeholders have expressed the need to have some thresholds that would determine whether an incident falls under the listed instances.<sup>322</sup>

The new Directions, unlike the old Directions, penalize non-compliance<sup>323</sup> and the same is a step towards better enforcement of cyber security standards in the Indian cyberspace.

#### **4.2. China's Approach to Data Security: Cross Border Data Flow Restrictions and Data Localization Measures**

China's approach to data security stands on stringent principles set by three laws including the Cyber Security Law (2017), Data Security Law (2018) and, Personal Information Protection Law (2021). These laws are supported by various regulations including the Outbound Data Transfer Security Assessment Measures, Measures for Standard Contract for Personal Information CrossBorder Transfer, Regulations on Promoting and Regulating the Cross-Border Data Flows, etc. These laws aim to strengthen national security, protect personal data and promote the Chinese digital economy. However, due to the stringency of these laws, they have been subject to a lot of criticism internationally, particularly concerning the data localisation requirements, restrictions on cross-border flow of information and its impact on international trade. This is discussed in detail in Chapter 6.

China's cyber security laws have heavily emphasized on data localization requirements and cross border data flow restrictions. The Cyber Security Law (CSL) requires the classification of business data into different categories based on the level of its importance, based on which, restrictions are

---

<sup>321</sup> Indian Computer Emergency Response Team, 'Frequently Asked Questions (FAQs) on Cyber Security Directions of 28.04.2022' (2022) <[https://www.cert-in.org.in/PDF/FAQs\\_on\\_CyberSecurityDirections\\_May2022.pdf](https://www.cert-in.org.in/PDF/FAQs_on_CyberSecurityDirections_May2022.pdf)> accessed 18 December 2024.

<sup>322</sup> Aniruddha Majumdar, Aparna Gaur, Milind PM and Gowree Gokhale, 'CERT-IN releases FAQs explaining the Direction on Cybersecurity' (*Nishith Desai Associates*, 25 May 2022) <<https://www.nishithdesai.com/generateHTML/6139/3>> accessed 18 December 2024.

<sup>323</sup> CERT-In (n 316).

imposed on cross border-transfer.<sup>324</sup> CSL mandates that critical information operators have “important data” and that must be stored within the Chinese territory. Further, it also provides that companies that want to transfer “important data” outside the Chinese territory are mandated to perform an internal security review, apply for security assessment and get the approval of the Cyberspace Administration of China (CAC).<sup>325</sup> Article 31 refers to “critical information infrastructure” as those sectors whose information if destroyed would severely harm national security, public welfare, or economic stability. Furthermore, there is no precise definition for the term “important data” as of yet. The draft Measures for Security Assessment of Cross-Border Transfer of Personal Information and Important Data, only state that important data refers to data that is closely tied to “national security, economic development, or social public interests”.

These requirements are further bolstered by the Personal Information Protection Law that requires personal information processor’s to pass the security assessment by the national cyberspace department, obtain a certificate from a relevant specialised institution and conclude a contract of transfer in accordance with the prescribed standard contract by the cyberspace department.

The Data Security Law aims to regulate data processing and ensure data security. It outlines the extent of export control over data that concern national security and interests and international obligations.<sup>326</sup> It requires processors of important data to conduct risk assessments of their processing on a regular basis and subsequently submit a report to the relevant authorities.<sup>327</sup> All the provisions of the Cyber Security Law also apply to the outbound security management of important data as per Article 31.<sup>328</sup>

China’s cyber security laws also focus on risk prevention. The CSL requires network operators to implement multi-level protection system (MLPS). They are also required to create internal security management systems and adopt measures to prevent computer viruses/attacks that could endanger

---

<sup>324</sup> Cyber Security Law 2017, art 21.

<sup>325</sup> Cyber Security Law 2017, art 21; *ibid* Article 37.

<sup>326</sup> *ibid*, art 25.

<sup>327</sup> *ibid*, art 30.

<sup>328</sup> *ibid*, art 31.

cybersecurity. They must also monitor and record network activity logs for at least 6 months.<sup>329</sup> This provision aims to enhance accountability and enable traceability of security breaches.

While China's strict regulations concerning movement of data and cyber security measures aim to protect national security and economic development, several criticisms towards its approach have been made. It can be argued that the focus of the laws is more on control and surveillance instead of providing a robust framework for cybersecurity incident response. Practices like data localization and restriction of data flows across borders have been termed as overly restrictive and can potentially harm global flow of information for international businesses, particularly in sectors where transfer of data is imperative. It also creates additional compliance costs for international businesses. Furthermore, the security assessments required to be fulfilled as a pre-requisite to outbound data transfer lacks transparency. The Chinese authorities have broad discretion to restrict transfer of data.

## **5. Conclusion**

Technological measures play a crucial role in protecting data from the ever-evolving nature of cyber threats. Cyber-attacks can lead to unauthorised access to sensitive and important business information. It allows attackers to steal, manipulate and destroy data. Cyber threats like hacking, phishing, malware, etc. compromise the integrity of data and consequently cause far reaching damage that can extend to financial and reputational harms. This financial and reputational harm can have both direct and indirect costs. Therefore, cyber-attacks pose a critical risk to organizations leading to a loss of intellectual capital. In the current digital economy, the importance of data has increased exponentially and so have the chances of it becoming the prime target of cyber-attacks.

The recognition of the growing importance of cyber-security and data protection measures is also reflected in legal and regulatory frameworks. In India, the Information Technology Act, 2000 and the IT Rules provide the foundation for data protection. The Act recognises a wide range of issues related to cyber-security including hacking, data theft, unauthorized access, etc. In addition to this

---

<sup>329</sup> *ibid*, art 21.

the IT rules mandate organizations to implement measures like encryption, access control, network security, etc. to safeguard data and prevent unauthorised access to data. A critical element of cyber security regulations reporting requirements for cyber security incidents to the CERT-In. This enables authorities to respond swiftly and mitigate the potential harm. A failure to respond to cyber-attacks timely can exacerbate the damage. This also aligns with practices across the globe focusing on reporting requirements and risk implementation. A well-structured legal and technological framework coupled with transparent remedial actions can act as a deterrent to cyber-attacks and can help build resilience.

Some new strategies to strengthen cyber security include data localisation measures and restrictions on cross border flow of data. Recently, China's stringent approach to cyber security measures have attracted a lot of attention and criticism. Its practices have been termed as overly restrictive with the potential for hindering global flow of information (issues concerning data localisation measures and restrictions on cross border flow of data have been elaborated in Chapter VI of the Report). Along with these issues, the issue of cross border data theft has also been highlighted in the context of challenges in implementing cyber security laws and enforcement measures across borders vis-a-vis regulatory autonomy. Thus the need for an international multilateral framework has become more urgent than ever. An international approach can facilitate cooperation and standardize legal framework that can help in creating a safe digital ecosystem in an interconnected world.

## **CHAPTER IV: COMPETITION LAW AND NON-PERSONAL DATA**

### **1. Data Markets and Competition Law**

The present-day markets are increasingly experiencing digitisation which has contributed to the increasing importance of the value of data as a critical asset for businesses across various sectors. Prior to this, data was merely a by-product of business activities but now it is a part of a firm's competitive intelligence providing valuable insights into consumer behaviors, consumer demands, market trends and competitor strategies. Collection and analysis of big data has the potential to provide relevant insights and improve a firm's service quality and innovation abilities. As a consequence of this increased asset value of data, firms are increasingly leveraging data in their

decision-making process to drive innovation and gain a competitive edge in the market. Amidst this data-driven transformation of the market, a pressing concern emerges in the market regarding the competitive landscape. Firms with rich data resources have access to vast amounts of data that can provide them valuable insights for data driven innovations and decision making which gives them a powerful market position. While some firms have rich data sources, other firms without access to rich and comprehensive data resources can possibly find themselves at a competitive disadvantage. This discrepancy in access to data thus raises questions concerning fairness, market competition and the potential of anti-competitive conduct by data-driven firms.

This chapter delves into the complexities surrounding the use of big data as competitive intelligence too in digital markets and analyzes anti-competitive issues that can potentially arise from use and access to data. The chapter also sheds light on challenges that firms may face due to lack of access to data in data driven markets.

### **1.1. Data as Competitive Intelligence**

Data can throw light on consumer behavior and shifting demands which can allow firms to personalize their content, and services i.e. enable them to “delivery of “the right content to the right person at the right time to maximize immediate and future business opportunities”.<sup>330</sup> This can create user satisfaction and customer loyalty and increase switching costs for consumers. Therefore, exclusive access to big data can act as a competitive advantage for firms.<sup>331</sup> Big data has hence been defined as *"high-volume, high-velocity, and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight, decision making and process automation."*<sup>332</sup> Accumulation of such data in the hands of one business provides them with the opportunity to create more products and gain more market share and users which would eventually lead to increased revenue. The accumulation of data in the hands of one

---

<sup>330</sup> Tam and Ho, “Understanding the Impact of Web Personalization on User Information Processing and Decision Outcomes” (2006) 30 MIS Quarterly 865.

<sup>331</sup> Fast V, Schnurr D and Wohlfarth M, “Regulation of Data-Driven Market Power in the Digital Economy: Business Value Creation and Competitive Advantages from Big Data” (2023) 38 Journal of Information Technology 202.

<sup>332</sup> “Definition of Big Data - Gartner Information Technology Glossary” (Gartner) <<https://www.gartner.com/en/information-technology/glossary/big-data>> accessed January 24, 2025.

business can therefore act as an entry barrier for other businesses as the data holder can leverage their position and capture the market.

Along with access to data, the ability to utilize its actual worth and value also contributes to a firm's potential to gain a competitive edge in the market. The value of data depends on the quality of the data, how the data transfers into actual learnings, how the data interacts with other data sources and the uniqueness of the data.<sup>333</sup> Due to its voluminous nature, big data cannot be analyzed by traditional methods but requires unique platforms that can handle huge amounts of data.<sup>334</sup> The quality of data held by firms also affects the value that can be extracted from it. When it comes to data quality, four factors play an important role i.e. usability, accuracy, relevancy and the time dependency of data. If the data is not accurate or relevant, it cannot provide any beneficial insights.<sup>335</sup> Further, the relevancy of data in certain markets can decrease over time and businesses can't make much use of stale data. For example, Facebook found out that their learning and performance were significantly affected when they used old data.<sup>336</sup> According to a study conducted in 2019, the relevancy of data decreases by 3% every month.<sup>337</sup> Updated, new and fresh data can be beneficial than voluminous but outdated data. Similarly, if the data is not unique and can be easily imitated by competitors, it cannot provide a competition advantage. Netflix, an online streaming service has been collecting data since its inception, but it did not help them avoid competition from other platforms like HBO Max, Amazon Prime, Hulu, Disney, etc. It is evident that mere access to data does not provide firms with a competitive advantage. The nature and quality of the data and a firm's capacity to utilize this data efficiently impacts whether a firm will get a competitive advantage over its competitors.<sup>338</sup>

The value of data is rooted in its analysis and the information and insights derived from it. It is the information and insights derived from the data that is valuable to firms and plays an important role

---

<sup>333</sup> Iansiti M, "The Value of Data and Its Impact on Competition" [2021] SSRN Electronic Journal.

<sup>334</sup> Rubinfeld DL and Gal MS, "Access Barriers to Big Data" [2016] SSRN Electronic Journal.

<sup>335</sup> *ibid.*

<sup>336</sup> *ibid.*

<sup>337</sup> Writer AS, "Control Data Decay To Enhance Your Sales Efficiency" (*AiThority*, September 4, 2019) <<https://aithority.com/ait-featured-posts/control-data-decay-to-enhance-your-sales-efficiency/>> accessed January 24, 2025.

<sup>338</sup> *ibid*

in enabling them to make better decisions for innovation.<sup>339</sup> Data in itself does not have any value and firms need the ability to analyze large volumes of data to be able to extract value from it. Therefore, firms with the ability to analyze and process big data can make predictions from it and learn from the past. This creates a “positive feedback loop”.<sup>340</sup> The more data a firm collects from its users, the more accurately they can target their customers.<sup>341</sup> These self-enforcing feedback loops can be of two types. First is the user feedback loop wherein, more users generate more data and create more business value for firms and simultaneously, firms can produce more value for its customers. Second is the monetization feedback loop wherein, with access to large amounts of user data, firms can optimize targeted advertising and invest more in improving quality of service and thereby generating more revenue.<sup>342</sup> These feedback loops become competitive advantages and can therefore act as entry barriers for new market entrants.

## **1.2. A Rising Concern**

Alongside the immense potential for innovations and growth driven by data, there are also the rising concerns over the budding anti-competitive behavior in digital markets. Data acts as both a catalyst for innovations and a barrier to entry for competitors.

Access and use of data can create barriers to entry in the data driven economy. In the context of search engine markets, positive feedback loops created by data are a fundamental problem faced by new entrants. New entrants in the market have little data to rely on to improve the relevance and responsiveness of their search results. With low quality search results, it is difficult for new entrants to develop a user base as users are less likely to switch to a lower quality search engine after being accustomed to a different search engine. New entrants will therefore have a small user base which would reduce the opportunities to attract advertisers as advertisers would want to reach a larger part of their target audience. This eventually creates a feedback loop wherein newer search engines would find it difficult to grow their user-base and revenue due to lack of data resulting in

---

<sup>339</sup> Daniel L Rubinfeld and Michael S Gal, “Access Barriers to Big Data” (2017) 59(2) Arizona Law Review 339-381 <<https://journals.librarypublishing.arizona.edu/arizlrev/article/id/6959/>> accessed 22 January 2025.

<sup>340</sup> *ibid.*

<sup>341</sup> *ibid.*

<sup>342</sup> Fast V, Schnurr D and Wohlfarth M, “Regulation of Data-Driven Market Power in the Digital Economy: Business Value Creation and Competitive Advantages from Big Data” (2023) 38 (2) Journal of Information Technology 202.

lower quality results. At the same time without a substantial user base and revenue, the search engine would be unable to improve its algorithms and expand its services. This loop can perpetuate the problem of limited growth and competitiveness. This chicken and egg problem emphasizes the value of data driven network effects wherein dominant players can reinforce their dominance by refining their algorithms and increasing their user base.<sup>343</sup> This cycle creates significant barriers to entry for new entrants in the market.

These data driven network effects become even more intense when platforms have two tiered structures for data collection. For instance, Facebook employed a two tier data collection structure wherein it collected data both “On Facebook” and “Off Facebook”. This two tiered structure for data collection enabled Facebook to highly enhance its personalized services. The activity was said to create entry barriers in the social networking market as the platform was availing the benefits of direct network effects wherein the more users it attracted, the more appealing it became to newer users. In other words, Facebook was using the feedback loop to reinforce its dominance and making it difficult for new entrants in the market to effectively compete.<sup>344</sup>

Furthermore, another concern in digital markets is that firms with the ability to analyze big data can detect their potential future competitors or “nascent firms” that have not fully grown to their full potential yet or their “potency as a competitor is as yet not fully developed and hence unproven.”<sup>345</sup> By identifying their competitors, firms can monitor their position in the market and trace competitive threats and take actions to eliminate them.<sup>346</sup> This practice has been termed as “acquire-copy-or-kill” or “killer acquisition”<sup>347</sup> where dominant firms eliminate their competition

---

<sup>343</sup> Vikas Kathuria, *Greed for Data and Exclusionary Conduct in Data-driven Markets* (4 December 2018) <https://ssrn.com/abstract=3295436> accessed 28 January 2025. An updated version appears in *Computer Law & Security Review* (2019) 35(1) 89-102.

<sup>344</sup> Wiedemann K, “A Matter of Choice: The German Federal Supreme Court’s Interim Decision in the Abuse-of-Dominance Proceedings *Bundeskartellamt v. Facebook* (Case KVR 69/19)” (2020) 51 IIC - International Review of Intellectual Property and Competition Law 1168.

<sup>345</sup> CS Hemphill and Tim Wu “Nascent Competitors” (2020) 168(7) *Uni of Penn Law Review* <[https://scholarship.law.upenn.edu/penn\\_law\\_review/vol168/iss7/1/](https://scholarship.law.upenn.edu/penn_law_review/vol168/iss7/1/)> accessed 24 January 2025.

<sup>346</sup> Case M, “Google, Big Data, & Antitrust” <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3917218](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3917218)> accessed January 24, 2025.

<sup>347</sup> Dai K and Deng J, “Big Data and Antitrust Risks In Close-up: From the Perspective of Real Cases” (*California Lawyers Association*, January 22, 2021) <<https://calawyers.org/publications/antitrust-unfair-competition-law/competition-fall-2020-vol-30-no-2-big-data-and-antitrust-risks-in-close-up-from-the-perspective-of-real-cases/>> accessed January 24, 2025.

by acquiring them, copying them or by using tactics to kill the smaller firms before they achieve their potential. Using this tactic, Facebook tried to acquire Snapchat in 2013. Snapchat refused the offer and consequently, Facebook copied Snapchat's services by introducing a similar feature on its platform Instagram. Eventually, within a small-time frame, Facebook gained more users than Snapchat and succeeded in maintaining its dominance in the social media market.<sup>348</sup> Similarly, Facebook also acquired Whatsapp for \$22 billion, as it feared that Whatsapp might “morph into Facebook overtime”.<sup>349</sup>

In this context, control and access to big data can raise competition law concerns and challenges with respect to market entry barriers. There are four kinds of restraints on the collection of big data that can act as barriers i.e. technological, legal, behavioral and storage barriers.<sup>350</sup> Technological barriers include issues like the inability to imitate data or temporal issues i.e. “the point in time the competitor starts collecting data”. Legal barriers concerning big data collection require complying with laws. For example, complying with laws protecting privacy. It also raises concerns regarding data ownership. Behavioral issues, for example, include issues related to the price that the data owner sets while transacting data.

## **2. Anti-Competitive Practices Related to Data**

It has been noted that the competitive significance of data can be difficult to ascertain due to the characteristics of big data. Data is a “non-rivalrous” resource and collection of data by one entity does not exclude other entities from collecting it.<sup>351</sup> At every given moment today, data is being extensively generated and collected. Nowadays, data brokers are also available as an alternative source to purchase comprehensive datasets from.<sup>352</sup> Since the relevancy of big data and big analysis also depends on the freshness of data (the newer the data the more is its relevance), some

---

<sup>348</sup> Shinal J, “Mark Zuckerberg Couldn't Buy Snapchat Years Ago, and Now He's Close to Destroying the Company” *CNBC* (July 12, 2017) <<https://www.cnn.com/2017/07/12/how-mark-zuckerberg-has-used-instagram-to-crush-evan-spiegels-snap.html>> accessed January 24, 2025

<sup>349</sup> Charlie Warzel and Ryan Mac, 'These Confidential Charts Show Why Facebook Bought WhatsApp' *BuzzFeed News* <https://www.buzzfeednews.com/article/charliwarzel/why-facebook-bought-whatsapp> accessed 31 January 2025.

<sup>350</sup> Rubinfeld DL and Gal MS, “Access Barriers to Big Data” [2016] SSRN Electronic Journal.

<sup>351</sup> Case M, “Google, Big Data, & Antitrust” <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3917218](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3917218)> accessed January 24, 2025.

<sup>352</sup> Comerford DDS and RE, “Antitrust and Regulating Big Data” (2016) 23 UF Law Faculty Publications.

argue that the issue of anti-competitive behavior concerning exclusive data control resulting in an entry barrier does not arise.

On the other hand, it has also been noted that despite data's "non-rivalrous" nature, data which is actually relevant and valuable remains difficult to access and use in reality. Firms with data of "competitive significance" have no incentives to disclose or share the data as it gives them an advantage in the markets as such data can be difficult and costly to replicate.<sup>353</sup> Moreover, even though some data can be purchased from data brokers, a lot of valuable data is under exclusive control of some dominant entities and can still create a barrier to entry and count as anticompetitive. Two sided or multi-sided platforms that are already dominant in the market have massive amounts of users on both sides of the platform that they can collect their data from.<sup>354</sup> The wide user base of entities gives them a lot of opportunities including the opportunity to generate revenues.<sup>355</sup> With larger quantities of data, platforms can improve their algorithms using machine learning. This bolsters their dominant positions. Thus, restrictive use of data can raise anti-competitive concerns. This competitive significance of data has also been recognised by the EU when it was found that the 2010 Microsoft and Yahoo Deal would improve the algorithm of the search engine through which the relevance of search results would become more accurate and relevant.<sup>356</sup> This characteristic of data makes it a non-fungible or a non substitutable resource in digital markets.

Therefore, new entrants would face a competitive disadvantage while trying to enter the markets without access to data like the data the already existing entities in the market have. Competition law aims to protect consumers, ensure redistribution of wealth and protect competitors. It attempts to achieve these aims by putting a check on anti-competitive agreements, abusive behavior and

---

<sup>353</sup> McSweeney CT, "Data, Innovation, and Potential Competition in Digital Markets – Looking Beyond Short-Term Price Effects in Merger Analysis".

<sup>354</sup> Vikas Kathuria, *Greed for Data and Exclusionary Conduct in Data-driven Markets* (4 December 2018) <https://ssrn.com/abstract=3295436> accessed 28 January 2025. An updated version appears in *Computer Law & Security Review* (2019) 35(1) 89-102.

<sup>355</sup> *ibid.*

<sup>356</sup> SUNNYVALE Calif. and REDMOND, Wash, "Yahoo! And Microsoft to Implement Search Alliance" (*Stories*, February 18, 2010) <<https://news.microsoft.com/2010/02/18/yahoo-and-microsoft-to-implement-search-alliance/>> accessed January 24, 2025.

mergers & combinations that could restrict competition.<sup>357</sup> The following sections will throw a light on the interaction between data and competition law in digital markets.

### **2.1. Exclusionary Conduct and Data as an Essential Resource**

Dominant firms often engage in exclusionary practices that can restrict competition in various ways like cartelisation, price fixing, allocating market shares, restricting output or an essential resource, etc. Dominant entities in digital markets can restrict access to data or leverage their datasets to reinforce their dominance. This can create entry barriers and hinder a new competitor's ability to compete in the market effectively or on a level playing field.

Barriers to entry play a significant role in ascertaining whether a market is competitive or anti-competitive. The inability of a market player to access essential resources to enter a market is a classic example of how a barrier to entry can prevent competition. The essential facilities doctrine imposes antitrust liability on firms holding a dominant position refusing to provide access to an essential facility for competition to other competitor firms. It is very likely that conflicts concerning access to data are bound to arise. Access to data that provides significant insights and inputs to firms provides a significant competitive edge to firms. Generally, firms have the freedom to deal with or to not deal with other parties in the way they please. However, in certain exceptional circumstances, a refusal to deal can be anti-competitive. In this context, since refusal to share data can impede a firm's ability to innovate and grow, a question arises regarding the applicability of the essential facilities doctrine to get access to data.

The essential facilities doctrine originated in the United States through application of section 1 and 2 of the Sherman Act, 1890 in multiple cases. The Terminal Railroad case<sup>358</sup> was one of the first cases where access to the terminal facility was held to be essential by the Court. The Court held that it was anti-competitive to deny access to an essential facility to the competitors and the same had to be provided on reasonable terms. In the US, there are four elements necessary to invoke the essential facility doctrine; (1) the control of the essential facility by a monopolist, (2) a competitor's

---

<sup>357</sup> Richard Whish and David Bailey, *Competition Law* (Oxford University Press 2021).

<sup>358</sup> *United States v. Terminal Railroad Association*, [1992] 224 U.S. 383.

inability practically or reasonably to duplicate the essential facility, (3) the denial of the use of the facility to a competitor and, (4) the feasibility of providing the facility.<sup>359</sup>

Similarly, in the EU, the doctrine has been applied to tackle abuse of dominant position by undertakings. Article 102 of the Treaty Establishing the European Community<sup>360</sup> provides that any abuse by an undertaking holding a dominant position is prohibited. Abuse of dominant position here can be imposing unfair prices or unfair trading conditions, discriminating other parties in similar transactions by applying dissimilar conditions to put them at a disadvantage, limiting production, market or technical development at the prejudice of the consumers, making the conclusion of contracts subject to acceptance by the other parties of supplementary obligations.<sup>361</sup>

This provision was used in the EU in the context of essential facilities doctrine in the case of *Commercial Solvents* where the Court noted that if an undertaking holds a dominant position in the market (here production of raw materials and the subsequent control over the manufacturers of derivatives) it cannot refuse to supply the raw materials to the manufacturers of derivatives in order to eliminate the competition in the derivative market. This would constitute abuse of dominant position under Article 86. However, the term “essential facility” was used for the first time in *Sea Containers v. Stena Sealink*<sup>362</sup> where the Commission held that if an entity holds a dominant position and deals with an essential facility, without the access of which competitors cannot operate in the market, cannot deny the access to the essential facility without any reasonable justification. If access to essential facilities is denied on unreasonable grounds or is being offered at unreasonable terms, it would be violative of Article 86 (of the then EC Treaty).

The jurisprudence on essential facilities has been in an expansionary phase.<sup>363</sup> It has developed through case laws across jurisdictions. In India, the doctrine was applied in *Arshiya Rail Infrastructure Limited v. Ministry of Railways & Ors.*<sup>364</sup> wherein the Commission observed that the doctrine can be invoked only in certain circumstances like “existence of technical feasibility

---

<sup>359</sup> *MCI Communications Corp. v. AT&T*. [1983] (708 F.2d 1081, 1132 (7th Cir.), cert. denied, 464 U.S. 891.

<sup>360</sup> Consolidated version of Treaty Establishing the European Economic Community [2002] OJ C 325.

<sup>361</sup> “Competition Policy” (Competition Policy) <[https://ec.europa.eu/competition/legislation/treaties/ec/art82\\_en.html](https://ec.europa.eu/competition/legislation/treaties/ec/art82_en.html)> accessed January 24, 2025.

<sup>362</sup> (OJ L 15/8 (1993).

<sup>363</sup> Areeda P, “ESSENTIAL FACILITIES: AN EPITHET IN NEED OF LIMITING PRINCIPLES” (1989) 58 Antitrust Law Journal 841.

<sup>364</sup> *Arshiya Rail Infrastructure Limited v. Ministry of Railways & Ors.*, [2010] Case No. 64 & [2011] Case No. 12.

to provide access, possibility of replicating the facility in a reasonable period of time, distinct possibility of lack of effective competition if such access is denied and possibility of providing access on reasonable terms.” Further, the doctrine was also dealt with in the case of *Shri Shamsheer Kataria v. Honda Siel Cars India Ltd. & Ors.* wherein the Director General in his investigation report pointed out that the points to be considered while determining essentiality are “(a) control of the essential facility by the monopolist; (b) the inability to duplicate the facility; (c) the denial of the use of the facility, and (d) the feasibility of providing the facility.” The Commission concurred with the report and it was held that the denial to access spare parts to independent repairers would constitute an abuse of dominance.

### 2.1.1. Data as an Essential Resource

If essential facilities doctrine were to be applied to refusal to share data by dominant firms, it would be necessary to establish that the data is essential for competition and that the entity in control of the data has refused to give access despite having the means to do so. Further, the refusal will curtail competition as the data cannot be imitated.<sup>365</sup>

As compared to other facilities, big data is significantly different. As mentioned earlier, big data is intangible and its value keeps decreasing with time. Data in itself is not useful as it has no intrinsic value and value is derived through analyzing it. Data by nature is non-rivalrous and non-excludable i.e. it can be used simultaneously by many people without the exclusion of others. Due to big data’s non-rivalrous nature, the same data can be used by multiple users and the same data will hold different value for different users.<sup>366</sup> However, this does not translate into practice as access to data is restricted through means like exclusive dealing and protecting trade secrets. In order to establish that data is essential, it would be necessary to also establish that it is indispensable and non-substitutable.<sup>367</sup>

---

<sup>365</sup> Essential Data Author(s): ZACHARY ABRAHAMSON [https://www-jstor-org.nludelhi.remotexs.in/stable/pdf/43617042.pdf?refreqid=fastly-default%3Ae8f3fb8601e500a388b61a0bfa20a703&ab\\_segments=0%2Fbasic\\_search\\_gsv2%2Fcontrol&origin=&initiator=&acceptTC=1](https://www-jstor-org.nludelhi.remotexs.in/stable/pdf/43617042.pdf?refreqid=fastly-default%3Ae8f3fb8601e500a388b61a0bfa20a703&ab_segments=0%2Fbasic_search_gsv2%2Fcontrol&origin=&initiator=&acceptTC=1)

<sup>366</sup> Rubinfeld DL and Gal MS, “Access Barriers to Big Data” [2016] SSRN Electronic Journal

<sup>367</sup> Graef I, “EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility (PhD Summary)” (*SSRN Electronic Journal*) <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3635378](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3635378)>.

It has been argued that smaller firms or new entrants in the market face a “data gap” due to which they are unable to effectively compete with larger firms on a level playing field due to their lack of access to an equivalently voluminous data set. This data gap widens the gap in quality of services between smaller and larger firms due to which the competitive rivalry between new entrants and dominant firms also diminishes.<sup>368</sup> If the quality of services of new entrants are not good, the consumers would also be reluctant in switching to them. This can consequently reduce the incentive for competitors to innovate and compete with dominant firms as dominant firms through big data analysis predict future trends and strategies to eliminate future competition.<sup>369</sup> It should be also noted that buying data from third parties like data brokers can not be as valuable as datasets curated over the years by dominant entities.<sup>370</sup> Therefore, new competitors face a disadvantage as they are unable to access data of the same quality, variety and volumes as dominant firms. For instance, in the *Nielsen-Arbitron* case, Nielsen and Arbitron, two of the best players for cross platform audience measurement services, were in the possession of highly accurate and relevant data. The commission noted that the data held by the two entities was relevant input for downstream platforms that were not developed yet and the said data could also not have been imitated/replicated by others. Therefore, the commission found that the acquisition would result in significant competitive harm to the market and noted data as a barrier to market entry but the acquisition was approved with certain conditions to protect competition and promote the development of other cross platform services.<sup>371</sup>

By now it is also clear that direct network effects can lead to the growth of an enterprise. Direct network effects can trigger positive indirect network effects and with a wider user base, firms can attract more advertisers and generate more revenue.<sup>372</sup> As highlighted earlier, due to lack of access to data, entities can face the “chicken and egg” or the “snowball effect” problem. For example, if search engines receive a limited number of queries, their algorithms will have less data to learn

---

<sup>368</sup> Comerford DDS and RE, “Antitrust and Regulating Big Data” (2016) 23 UF Law Faculty Publications.

<sup>369</sup> *ibid.*

<sup>370</sup> Kathuria V, “Greed for Data and Exclusionary Conduct in Data-Driven Markets” (2019) 35 Computer Law & Security Review 89.

<sup>371</sup> Commission FT, “Analysis of Agreement Containing Consent Order To Aid Public Comment” [2013] File No. 131 0058.

<sup>372</sup> Kathuria V, “Greed for Data and Exclusionary Conduct in Data-Driven Markets” (2019) 35 Computer Law & Security Review 89.

from and the results generated will be less accurate as a result of which users would be more unlikely to switch to their search engine.<sup>373</sup> With a small user base, the platform would also be unable to attract advertisers. Therefore, the feedback loop creates a snowball effect which can be detrimental for new entrants in the market.

In 2018, the European Commission held Google to be in violation of competition law. Three practices of Google were under scrutiny, i.e. “(1) tying of Google search and browser to Google Play Store, (2) paying manufacturers to pre-install google search app exclusively on their devices and, (3) preventing manufacturers who wished to pre-install Google apps from selling even a single smart mobile device running on alternative versions of Android that were not approved by Google”.<sup>374</sup> While adopting the theory of harm, the Commission noted that Google was trying to preserve its status quo bias among users by preventing them from switching to other search engines. It was further noted that in the search engine market, the quality and accuracy of search results plays an important role and the practices employed by Google curtailed traffic on other search engines which resulted in depriving the rivals of mass volumes of data i.e. a fundamental resource for machine learning. The exclusionary effects of Google’s conduct included (1) making it harder for i’s rival to gain search queries and improve their services, (2) increasing entry barriers and creating a shield to protect itself from competition, (3) reducing incentives for rivals to innovate and compete in the market and, (4) directly or indirectly harming consumers by reducing their choice.<sup>375</sup> Google tried to maintain the status quo bias its users and foreclosed actual and potential competition in the market. This case demonstrates how big data qualifies as essential to data driven markets.<sup>376</sup>

However, whether or not data is essential for effective competition in the market would depend from case to case and thus the assessment requires a contextual analysis. For example, in the *Google/DoubleClick* merger, the concern was whether the combination of assets of both the

---

<sup>373</sup> *ibid.*

<sup>374</sup> “Antitrust: Commission Fines Google €4.34 Billion for Illegal Practices Regarding Android Mobile Devices to Strengthen Dominance of Google’s Search Engine” (*European Commission - European Commission*) <[https://ec.europa.eu/commission/presscorner/detail/en/ip\\_18\\_4581](https://ec.europa.eu/commission/presscorner/detail/en/ip_18_4581)> accessed January 24, 2025.

<sup>375</sup> *Google LLC v. Alphabet Inc.* [2024] C-48/22 P.

<sup>376</sup> *ibid* (373).

entities resulted in an advantage that their competitors could never achieve. The Commission noted that the data in possession of both the entities i.e. data about users' search behavior, would be easily available to rivals like Yahoo and Microsoft.<sup>377</sup> For similar reasons, in the Facebook-Whatsapp case, the Commission noted that the merger would not restrict competitors from accessing data related to user behaviors.

### **2.1.2. Big Data: Pro-Competitive?**

Article 101 of the TFEU prohibits agreements between undertakings which can prevent, restrict or distort competition in the market. However, the provision is inapplicable in cases where agreements falling under Article 101 contribute to improving the distribution of goods or promoting technical and economic progress while allowing consumers a fair share of benefits.

It is argued that big data can have some pro-competitive benefits. Big data enables firms to offer free services to customers who are willing to monetise their data. Further, collection and monetisation of data is said to be a “profit maximizing” behavior that can be beneficial for consumers and not harmful to competition.<sup>378</sup> Therefore, instead of seeing free services as a competition problem, some view it as a pro-competitive benefit of big data. Big data also contributes to promoting innovation, improving quality of services and offering value added services to customers for free. Further, the nature of data is non-rival; therefore, accumulation of data by one entity does not preclude others from doing the same and does not act as an entry barrier for rivals. Rather, in digital markets, there are lesser barriers to entry wherein new entrants can rapidly collect data by entering the market with innovative goods and services.<sup>379</sup> Innovation is said to be the only way to achieve competitive success. Moreover, talent and skills can possibly be more important than merely having access to large amounts of raw data with limited analytical skills.<sup>380</sup> For instance, despite having multiple well established competitors, Whatsapp was quickly able to establish itself. Additionally, it has also been argued that the value of large volumes of data

---

<sup>377</sup> Commission Decision of 11 March 2008, Case COMP/M.4731 – Google/DoubleClick, [2008] OJ C 255/13.

<sup>378</sup> Lerner AV, “The Role of ‘Big Data’ in Online Platform Competition” [2014] SSRN Electronic Journal.

<sup>379</sup> Daniek Sukol and Roisin E.Comerford R, 'Antitrust and Regulating Big Data' (2016) 23 George Mason Law Review 119 <https://ssrn.com/abstract=2833129> accessed 04 January 2025.

<sup>380</sup> de Peyer BH, “EU MERGER CONTROL AND BIG DATA” (2017) 13 Journal of Competition Law & Economics 767.

is overestimated and what matters more is the freshness of the data.<sup>381</sup> Therefore, mere access to data is not enough for competitive advantage. On one hand it is argued that the competition promotes innovation and on the other hand, innovators require market power to fund R&D for innovation.<sup>382</sup>

The pro-competitive effects of agreements falling under Article 101 TFEU need to improve production/distribution of goods or improve technical or economic progress. This benefit however needs to be objectively measured and the benefits must outweigh the disadvantages. The Commission has previously held that research and development projects can contribute to improving efficiency and economic or technological progress. Furthermore, these agreements should not impose conditions that are not indispensable in achieving these objectives and must not eliminate substantial competition. Big data may or may not be a barrier to entry depending on the case at hand and hence a contextual approach while examining effects on competition is crucial.

Competition law enforcement can deal with the unique competition issues posed by the data market. However, it needs to adapt to the evolution of digital markets and its unique problems as it would be difficult to satisfy the traditional thresholds.<sup>383</sup> As per Judge Posner, "antitrust law is supple enough . . . to take in stride the competitive issues presented by the new economy." and, "the risk of false positives dictates a hands-off approach to digital markets."<sup>384</sup> The basics of antitrust law need not be altered or abandoned to deal with data market issues but some clarity and discourse on its applicability on the peculiar issues posed by the digital market is required.

## **2.2. Data and Abuse of Dominance**

Firms having access to valuable big data have a position of leverage in the market as big data provides a competitive edge over other competitors. Dominant firms can unilaterally create competition law issues in the market. Abuse of dominant position includes behaviors like

---

<sup>381</sup> *ibid.*

<sup>382</sup> McSweeney CT, "Data, Innovation, and Potential Competition in Digital Markets – Looking Beyond Short-Term Price Effects in Merger Analysis".

<sup>383</sup> Interview Responses

<sup>384</sup> Richard A. Posner & William M. Landes, Market Power in Antitrust Cases (1981) 94 HARV. L. REV. 937, 937.

imposition of unfair terms & conditions unilaterally, limiting production or development at the prejudice of the consumers, applying dissimilar conditions casting competitive disadvantage, etc.

In data markets particularly, instances of abuse of dominant position have been noted quite a number of times. In some cases dominant firms were made to share their data with competitors. For instance, in the *GDF Suez* case, GDF Suez was the former sole dominant player and was competing with other players in an unregulated market. The former dominant player utilized business structures, databases and resources that it had inherited from its previous monopolistic status. It was found that the inherited database contained “detailed commercial data” and was non-imitable. Therefore, the use of the database by GDF Suez had the potential to foreclose market entry for new entrants and was contrary to Article 102 of TFEU. Thus, GDF Suez was directed to mandatorily share their database in order to enable the competitors to compete on equal terms.<sup>385</sup> Similar to the GDF Suez case, the Belgian Competition Authority in the *Belgian Lottery* case found that the National Lottery was abusing its dominant position by using the contact details of customers from its business activities in public lotteries to enter into the market of sports betting by sending one-off emails to the customers.<sup>386</sup>

In the 2007 *Microsoft vs. Commission* case, the primary question before the Commission was whether Microsoft’s refusal to license “interoperability information” to potential competitors constitutes an abuse of its dominant position in the client PC operating systems market. The Court held that a refusal to license would be an abuse of dominant position in certain exceptional cases. These exceptional cases, as elaborated by the Court are (1) “*the refusal relates to a product or service indispensable to the exercise of a particular activity on a neighboring market*”; (2) “*the refusal is of such a kind as to exclude any effective competition on that neighboring market*”; and (3) “*the refusal prevents the appearance of a new product for which there is potential consumer demand.*”<sup>387</sup> The Court held that a refusal to license “interoperability information” by Microsoft

---

<sup>385</sup> Landes WM and Posner RA, “Market Power in Antitrust Cases” (1981) 94 Harvard Law Review 937.

<sup>386</sup> *GDF Suez v. European Commission* [2012] T-370/09. See also Kathuria V and Globocnik J, “Exclusionary Conduct in Data-Driven Markets: Limitations of Data Sharing Remedy” (2020) 8 Journal of Antitrust Enforcement 511.

<sup>387</sup> *Microsoft Corp. v. Commission of the European Communities* [2007] T-201/04.

would constitute abuse of dominant position since it can eliminate competition in the downstream market.

In 2016, Microsoft was under another scrutiny by the Commission for abuse of dominance. In the *Microsoft LinkedIn mergers* case, the Commission noted that a merger where datasets held by two independent firms combine, can give rise to horizontal issues in two ways. Firstly, combination of two datasets can increase the merged entity's market power and result in creation of market entry barriers for new entrants and secondly, the competition that existed between the two independent firms prior to the merger would be eliminated due to the merger. However, in this case, it was held that the concerned merger is not anti-competitive since the combination of the datasets do not constitute an entry barrier as a large amount of data valuable for online advertising would continue to exist devoid of Microsoft's control.<sup>388</sup>

In 2020, Bundeskartellamt held Facebook liable for abuse of dominant position for collecting excessive consumer data. Facebook had a two tiered data collection system i.e. "On Facebook" and "Off Facebook" for personalizing user experience. It was noted that for social networking services, collecting "On Facebook" data is sufficient and additional "Off Facebook" data collection is not necessary for satisfying Facebook's consumers' needs. Further, upon survey, the authority discovered that there are users who would like to opt out of Facebook's "Off Facebook" data collection practices but Facebook does not provide for a less data intensive feature. Therefore, users are forced to either agree with Facebook's data collection policies or fully refrain from using their services. Thus, the Court applied the theory of harm called "*aufgedra'ngte Leistungserweiterung*", which can be roughly translated as "imposed extension of services". It was noted that Facebook's decision of not providing a less "data intensive" option indicates imposition of their dominance irrespective of user preferences. This resulted in abuse of dominant position due to direct network effects that worked against competitors (the more users the more attractive it will become to advertisers).<sup>389</sup>

---

<sup>388</sup> Case M.8124 – Microsoft / LinkedIn [2016].

<sup>389</sup> Wiedemann K, "A Matter of Choice: The German Federal Supreme Court's Interim Decision in the Abuse-of-Dominance Proceedings Bundeskartellamt v. Facebook (Case KVR 69/19)" (2020) 51 IIC - International Review of Intellectual Property and Competition Law 1168.

Dominant firms like Google and Amazon have also previously faced accusations of self-preferencing. In 2012, the Competition Commission of India examined Google's conduct of preferencing its own services and partners by manipulating the search results. It was alleged that Google's search results preferred its own sites irrespective of the relevance of the search results and this conduct created an uneven playing field. It was argued that manipulated search results denied market access to other market players as it diverted the traffic to Google's preferences. It was also alleged that Google is imposing unfair conditions on third party advertisers through its Adwords program used by advertisers. The Adwords program allowed advertisers to bid for search terms for their ads but Google had access to data related to Adwords and was able to ensure that its advertisers appeared on top. The CCI held Google accountable for abuse of dominant position for imposing unfair conditions and manipulating the search results.<sup>390</sup>

In a 2021 report, it was alleged that Amazon is engaged in anti-competitive behavior by running a systematic program to create knock-off products and preferring their own knock-off products in India.<sup>391</sup> Prior to that the Commission had opened an investigation against Amazon for abuse of dominance in 2019. The Commission analyzed Amazon's conduct related to self-preferencing in 2022 and its preliminary concerns were that Amazon's data use is capable of having, and likely to have, anti-competitive effects in online retail markets on Amazon's e-commerce platforms. The Commission's concerns revolved around: (a) Amazon Retail's reliance on third party sellers data, (b) their "Buybox" criteria and, (c) the Prime label selection criteria and their anti-competitive effects on online retail markets. It was noted that Amazon Retail's decisions regarding products that it wants to sell, product pricing, inventory management, vendor decisions relied on other retailer's data and that provided Amazon with a structural competitive advantage and can foreclose other third-party sellers.<sup>392</sup> The Commission also took note of Amazon's "Buy-Box", a privileged ranking system offered by Amazon, will most likely affect the consumer's choice and behavior and have an anti-competitive effect in online retail markets. Lastly, the Commission noted that

---

<sup>390</sup> In re Matrimony.com Ltd. [2012] Case Nos. 07 and 30.

<sup>391</sup> Reports S, "Amazon Copied Products and Rigged Search Results, Documents Show" *Reuters* (October 13, 2021) <<https://www.reuters.com/investigates/special-report/amazon-india-rigging/>> accessed January 28, 2025.

<sup>392</sup> Council Regulation (EC) No 1/2003 relating to a proceeding under Article 102 of the Treaty on the Functioning of the European Union (TFEU) and Article 54 of the EEA Agreement [2022] Cases AT.40462 – Amazon Marketplace and AT.40703 – Amazon Buy Box.

Amazon's Prime Label offered free and faster delivery to buyers and therefore generated more sales. However, Amazon had a set-criteria to get the Prime label and this criteria was favoring Amazon's own retail brands. Finally, in order to address the anti-competitive behavior, Amazon committed to becoming more transparent, not using non-public seller data for Amazon's own retail activities and to handle third party data in silos. Amazon also committed to apply non-discriminatory conditions and equal conditions on all sellers while selecting the Buy Box winner to create a level playing field. With regard to the Buy Box selection criteria, Amazon also committed to display a second competing offer. Further, Amazon also committed to set non-discriminatory conditions for sellers to get Prime access, allow sellers to choose their carriers and not use their data for its own logistics purposes.

In India as per the Market Study on e-commerce conducted by the Competition Commission of India, the e-commerce sector is expected to grow at the annual rate of 51 percent. With the increased growth of the e-commerce sector in the country, competition law issues have also been soaring. Data masking by e-commerce platforms, particularly in the hospitality industry, is one the key issues identified in the report.<sup>393</sup> In the restaurant industry, it has been observed that platforms impose unfair/exploitative contract terms that are not in the interest of the service providers and do not provide any insight into the orders generated through the platforms. Platforms possess critical information related to restaurants, which they choose not to disclose to the restaurants, and instead use the said information/data to promote their own services.<sup>394</sup>

### **2.3. Data Driven Mergers**

Recently, there has been an increase in the number of mergers in the data driven industries. This rise in data-driven mergers calls for closer inspection of data-driven mergers and ignites the issue concerning suitability of current antitrust regimes' to analyze and examine their competitive impact in digital markets.

---

<sup>393</sup> Competition Commission of India, 'Market Study on E-Commerce in India: Key Findings and Observations' (8 January 2020).

<sup>394</sup> NRAI, 'NRAI Raises Concerns over Deep Discounting, Data Masking with Swiggy, Zomato, Others' *The Economic Times* (21 January 2019) <https://economictimes.indiatimes.com/industry/services/hotels-/-restaurants/nrai-raises-concerns-over-deep-discounting-data-masking-with-swiggy-zomato-others/articleshow/67442065.cms> accessed 31 January 2025 CCI report.

The ability of firms to analyze big data enables them to identify firms that would be their future competitors or “nascent firms” who have not fully grown to their full potential or their “potency as a competitor is as yet not fully developed and hence unproven.”<sup>395</sup> By identifying their competitors, firms can monitor their position in the market and trace competitive threats and take actions to eliminate them.<sup>396</sup> This practice has been termed as “acquire-copy-or-kill” or “killer acquisition”<sup>397</sup> where dominant firms eliminate their competition by acquiring or copying them or by using tactics to kill them before they achieve their true potential. For example, Facebook tried to acquire snapchat in 2013 and snapchat refused the offer. After Snapchat refused the offer, Facebook copied Snapchat’s services by introducing a similar feature on its platform Instagram. Eventually, within a small time frame, Facebook gained more users than Snapchat and succeeded in maintaining its dominance in the social media market.<sup>398</sup> Facebook also acquired Whatsapp for \$22 billion, as it feared that Whatsapp might “morph into Facebook overtime”.<sup>399</sup> Such mergers and acquisition can result in accumulation of large amounts of relevant data with one dominant entity.<sup>400</sup> If a company has access to data of this kind, they’d have better knowledge of consumer preference and would thus be able to adapt the prices of their goods and services to the customer preferences.

The EU probed into the *Google-Fitbit* merger case due to the concern that merger would increase Google’s market dominance in online advertising market as it can increase Google’s access to data for personalisation of services. Fitbit had consumer data related to heart rate, sleep activity, oxygen level, geolocation data, biometrics, nutrition etc. that could provide an increased data advantage to

---

<sup>395</sup> Hemphill CS and Wu T, “Nascent Competitors” [2020] SSRN Electronic Journal.

<sup>396</sup> M Case, 'Google, Big Data, & Antitrust' (2014) 162 *University of Pennsylvania Law Review* 1167.

<sup>397</sup> Dai K and Deng J, “Big Data and Antitrust Risks In Close-up: From the Perspective of Real Cases” (*California Lawyers Association*, January 22, 2021) <<https://calawyers.org/publications/antitrust-unfair-competition-law/competition-fall-2020-vol-30-no-2-big-data-and-antitrust-risks-in-close-up-from-the-perspective-of-real-cases/>> accessed January 24, 2025.

<sup>398</sup> Shinal J, “Mark Zuckerberg Couldn’t Buy Snapchat Years Ago, and Now He’s Close to Destroying the Company” *CNBC* (July 12, 2017) <<https://www.cnn.com/2017/07/12/how-mark-zuckerberg-has-used-instagram-to-crush-evan-spiegels-snap.html>> accessed January 24, 2025.

<sup>399</sup> Charlie Warzel and Ryan Mac, 'These Confidential Charts Show Why Facebook Bought WhatsApp' *BuzzFeed News* <https://www.buzzfeednews.com/article/charliewarzel/why-facebook-bought-whatsapp> accessed 31 January 2025.

<sup>400</sup> Lasserre, B., and Mundt, A., COMPETITION LAW AND BIG DATA: THE ENFORCERS’ VIEW *Italian Antitrust Review* N.1 (2017)

Google for personalized advertising.<sup>401</sup> The Commission in its assessment found that even if Fitbit and Google are not competing in the same markets, Google's access to Fitbit's data will strengthen its dominance. It was also noted that this concern needs to be assessed in the context of 4 aspects i.e. (1) the relevance of the data acquired as a result of the merger, (2) Google's position in the market, (3) Impact if the merger on the Google's position in the market and impact on its rivals and, (4) the absence of countervailing entry. While examining the impact of the merger under these aspects, the Commission noted that Fitbit's data is relevant and valuable for Google which is a dominant player in the market of supply of online search advertising services. Thus, the merger would most likely increase Google's data collection capabilities and have a negative impact on unfettered development of competition in the relevant market. The Commission also noted that the market of online advertising is characterized by entry barriers and the buyers also lack countervailing power. However, the Commission cleared the transaction and held that the final commitments eliminate anti-competitive threats concerning the compatibility of the transaction on the EU internet market. Google committed to not use health and fitness data for Google ads, to store Fitbit data separately in Silos, to maintain third party access and to not degrade interoperability and user experience of third party smart watches.<sup>402</sup>

Such mergers and acquisitions can raise competition issues. As noted by Ben Holles De Peyer, at the horizontal level, when two entities merge (for example two social platforms), the subsequent merged entity can gain access to large amounts of data that can increase its market power and create entry barriers through user feedback loops.<sup>403</sup> Feedback loops can affect the market in three ways i.e. (1) It can increase the quality of service offered by the merged entity and increase revenue, (2) It can make it harder for competitors to compete and, (3) It can increase switching costs for the users as they can be locked in by the dominant merged entity.<sup>404</sup> Additionally, if there is a merger between two entities at a vertical level, for example, a data broker company and a

---

<sup>401</sup> Anca D Chirita, 'Exclusionary and Exploitative Abuse of Consumer Data' in Maria Ioannidou and Deni Mantzari (eds), *Research Handbook on Competition Law and Data Privacy* (Elgar 2023).

<sup>402</sup> Case M.9660 – Google/Fitbit, (Only the English text is authentic) Regulation (EC) No 139/2004, Merger Procedure, Article 8(2) Regulation (EC) 139/2004 (17 December 2020).

<sup>403</sup> Ben Holles de Peyer, 'EU Merger Control and Big Data' (2017) 13 *Journal of Competition Law & Economics* 767 <https://doi.org/10.1093/joclec/nhx026>.

<sup>404</sup> de Peyer BH, "EU MERGER CONTROL AND BIG DATA" (2017) 13 *Journal of Competition Law & Economics* 767.

company in the business of data analytics, the merged entity can restrict other competitors' data access.

It has been found that in many jurisdictions, competition law's merger assessment and thresholds are not apt enough to examine mergers in digital markets<sup>405</sup> or cover mergers and acquisition cases where the company acquired is a small startup at a nascent stage and has a small turnover.<sup>406</sup> Therefore, the intersection between competition and innovation needs to be understood in analyzing merger transactions to ensure that the existing dominant firms don't solidify their dominance in the market. The 2010 US Horizontal Merger Guidelines also state that the authorities should probe into the impact of mergers on both sides of digital markets and analyze the impact of the merger on innovation. Carl Shapiro notes that when it comes to regulating digital markets it is better to tolerate some false positives that wouldn't harm competition than to allow some false negatives that could potentially reduce competition by eliminating challengers to dominant entities.<sup>407</sup>

### **3. Conclusion**

It is quite evident that exclusionary conduct and access to resources like data creates significant competition issues and the same has been dealt with by competition authorities directly or indirectly. However, balancing the interests of the data holder and other stakeholders is highly important when it comes to mandating data sharing as it can act as a disincentive for firms to create data. Remedies like mandatory data sharing of future big data give competitors access to a firm's data which is something that the competitors were supposed to compete for in the market. While the purpose of such remedies is to increase competition in the market and prevent abuse of dominant position, it can also create a disincentive for firms to create innovative products for customers and kill the innovative spirit of the entrepreneur.<sup>408</sup> Mandatory data sharing remedies can also give rise to another problem that data shared as a remedy can be used beyond the market

---

<sup>405</sup> McSweeney CT, "Data, Innovation, and Potential Competition in Digital Markets – Looking Beyond Short-Term Price Effects in Merger Analysis".

<sup>406</sup> *ibid* (398).

<sup>407</sup> Shapiro C, "Antitrust in a Time of Populism" [2017] SSRN Electronic Journal.

<sup>408</sup> Exclusionary conduct in data-driven markets: limitations of data sharing remedy-Vikas Kathuria\* and Jure Globocnik + Arul George Interview

where the abuse takes place. The competitor receiving the data can thus appropriate the dataset to gain a competitive advantage in another market. This can distort competition in the market and even if a limitation on purpose of use is imposed on the competitor, it would be impossible to monitor the use of the dataset shared.<sup>409</sup>

A crucial question that arises in the context of mandatory data sharing is who should share data? If all the firms in the data market are obligated to share their data, two issues come into play. Firstly, since data sharing comes with an administrative cost, it would create a burden on smaller firms. Secondly, bigger firms have access to other information as well that they can use to get higher marginal benefits from the information shared. Therefore, it has been suggested that bigger firms should be obligated to share more data than smaller firms (“*an asymmetric data sharing obligation*”).<sup>410</sup>

Further, dealing with issues posed by data and specifically big data, through competition law can also burden antitrust authorities. Also, ensuring that such mandated data sharing is done efficiently is a challenging task since it is difficult to identify the type, extent of the data to be shared and to quantify the same.<sup>411</sup> Furthermore, applying competition law requires not only a firm to have a dominant position in the relevant market but also an abuse of the said dominant position.<sup>412</sup> Therefore, the antitrust regime would only be able to deal with issues in data markets if its principles are tweaked to cater to the peculiarities of big data and data markets.

---

<sup>409</sup> *ibid* 377.

<sup>410</sup> Prufer J and Graef I, “Governance of Data Sharing: A Law & Economics Proposal” [2021] SSRN Electronic Journal.

<sup>411</sup> *ibid* 377.

<sup>412</sup> Interview Responses.

## **CHAPTER V: REGULATORY APPROACHES TO DATA GOVERNANCE**

### **1. Approaches to Data Governance**

The importance of data has increased exponentially in the current times and the use of big data not only provides opportunities for innovation but also provides a competitive edge to organizations in the market. The previous chapter highlighted the competitive value of big data in digital markets and the myriad of benefits and opportunities it can offer that makes it important for businesses, government, community and individuals. This competitive value makes more and more organizations want to leverage the value of big data and big data analytics.

It is believed that since data is a non-rivalrous good, it can be consumed by an unlimited number of people simultaneously. However, this non-rivalrous nature of data does not translate into practice as its use can be limited by data-holders by exercising control over its use and access. This can pose several challenges concerning data ownership, data allocation, intellectual property, free flow & access to data, market concentration and competition. Some argue that the issues raised by big data and data driven network effects in digital markets can be perceived as market failure and necessarily require government intervention through ex-ante regulation.

Regulation plays an important role in preventing market failures and maximizing consumer welfare. Regulations in the market aim to address market imperfections and market failures creating barriers to entry.<sup>413</sup> When markets are unable to create optimal outcomes, timely and effective regulatory intervention can help in maximizing economic efficiency and welfare.<sup>414</sup> Market failures can essentially be categorized into four types i.e. (1) existence of market power and the capacity of suppliers to increase prices, (2) creation of negative externalities, (3) collective goods and free riding problem and, (4) information asymmetry.<sup>415</sup> In digital markets, market power and information asymmetry are the two main types of market failures that can be observed. Digital markets are highly concentrated with a few dominant firms like Facebook, Google, Amazon etc.

---

<sup>413</sup> Spulber DF, *Regulation and Markets* (MIT Press 1989).

<sup>414</sup> Cook P, *Leading Issues in Competition, Regulation, and Development* (Edward Elgar Publishing 2004).

<sup>415</sup> Alessio M. Paces and Louis T. Visscher, "Methodology of Law and Economics" [2023] SSRN Electronic Journal.

that control large amounts of data, have strong data driven network effects and create consumer lock-ins. These platforms are two-sided or multi-sided platforms with users on both sides that generate an immense amount of valuable data (both personal and non-personal) that creates direct and indirect network effects. Such datasets make it easier for firms to gain entry into or expand to different markets/services. With little to no choice to switch to services of a different market player, consumers often get locked in and have to necessarily consent to terms and services of the dominant platform or they can't avail their services. This allows dominant platforms to gain access to large amounts of data which can act as a competitive advantage.<sup>416</sup> The take-it or leave-it opt-in contract approach used by platforms to collect free user/consumer data is directly linked to their market dominance and it also reinforces their dominance.<sup>417</sup> Dominant firms can therefore increase switching costs for consumers and create lock-ins. Furthermore, data as an input is crucial for tackling social challenges as well. As per the Expert Group Report to the EU Commission, there are five aspects concerning the importance of data for public interest i.e.; “(a) *improves situational awareness, (b) better understand the causes and variables behind the current situation, (c) more accurately predict and forecast, (d) run more rigorous impact assessments and evaluations (of any intervention) in order to better define the policy problem and identify the most effective policy options, and (e) guide public management decisions taken either by humans or automated processes.*”<sup>418</sup> These public interest dimensions as discussed below are central to the question of regulation of non-personal data.

In this context, an issue arises concerning the ability of ex post mechanisms to effectively deal with market failures in unique digital markets that the idea of ex ante regulation for data is being explored. The Committee on Digital Competition Law recently noted that feedback loops and network effects are problems peculiar to digital markets and make it prone to concentration of market power. It also noted that ex post competition enforcement would be more efficient when it is supported by ex-ante regulation in a way that “*ex ante regulation ‘sets the rules of the game’*”

---

<sup>416</sup> Carugati C, “Regulation in the Digital Economy. Is Ex-Ante Regulation of ‘Gatekeepers’ An Efficient and Fair Solution?” [2020] SSRN Electronic Journal.

<sup>417</sup> Lianos I, “Giving Away Our Data for Free Is a Market Failure” (*UCL Library*, February 1, 2021) <<https://discovery.ucl.ac.uk/id/eprint/10145818>> accessed January 28, 2025.

<sup>418</sup> Directorate-General for Communications Networks, Content and Technology (European Commission), “Towards a European Strategy on Business-to-Government Data Sharing for the Public Interest” (*Publications Office of the EU*, February 15, 2021).

*and competition authorities through ex post regulation act as ‘umpires of the game’*”. Therefore, consumer welfare can be maximized when ex post and ex ante regulation work together.<sup>419</sup>

### **1.1. An Ex-Ante Approach to NPD?**

Market failure can be dealt with in two ways, *ex-post* and *ex-ante*. Ex-post mechanisms like competition law assessment can be used to approach market and access related issues but they cannot have the same level of influence as ex-ante regulations. Both differ on timing and the scope of their application. Ex ante regulation focuses on input levels and regulates the activities of actors before an act occurs. On the other hand, ex post system focuses on output level and on activities of the players once the act has occurred. Ex ante regulations have a defined scope and specific contours of their application and enforcement and on the other hand, ex-post standards and the applicable remedy is specific only after the occurrence of a firm’s acts.<sup>420</sup>

The OECD report on Ex-Ante Regulation and Competition in Digital Markets also analyzes the importance of ex ante regulation in digital markets. The report notes the shortcomings of the application of competition law in digital markets and highlights that ex-ante regulations can help in bridging the enforcement gap. According to the report, intervention through competition law is limited to anti-competitive agreements and abuse of dominant position. Instances that don't fall under these two categories, despite being harmful for the market, cannot be remedied through competition law. Further, competition law takes a case-by-case approach and does not result in *erga omnes* effect. Therefore, it cannot be used to tackle recurring systematic problems. This approach of antitrust enforcement can be complex and time consuming. Competition law can be used as a foundational point while tackling issues in the market as it can help in identifying systematic and structural problems in sectoral markets that could possibly need to be addressed by ex-ante regulation. For these reasons, it would be more efficient to address problems related to digital markets and data through both ex-ante and ex-post regulations. When it comes to ex ante regulations, the legislature has the liberty to decide the objectives of the law to cater to the societal

---

<sup>419</sup> Committee on Digital Competition Law

<sup>420</sup> ICC, ‘Global report on antitrust enforcement in the digital economy’ [2023] <<https://iccwbo.org/news-publications/policies-reports/global-report-on-antitrust-enforcement-in-the-digital-economy/>> accessed January 30, 2025.

needs. This liberty can help in expanding the scope of law beyond the objectives of competition law that address the issues at a policy level. Furthermore, ex-ante regulations can also directly or indirectly promote competition in the market. While ex ante regulations work in the foreground, competition law acts as a safety net or a “background regime” that can be used to deal with issues that fall outside the scope of regulations.<sup>421</sup> The OECD has also noted that the intention behind ex ante regulations for digital markets is to ensure and promote fairness, contestability. Transparency, innovation and to safeguard public interests.<sup>422</sup> The Competition Commission of India has also said that ex ante regulation in the digital markets can aid in “prompt market corrections”, help in regulating the digital economy and maintain a level playing field for all the competitors. The Ministry of Corporate affairs in the report on Anti-Competitive Practices by Big-Tech Companies noted that ex ante regulations in the digital market are rooted in the rationale that it is imperative to quickly address the systemic issues in the digital market to protect competition. Even if the antitrust case-by-case approach is more nuanced and evidence based, it is a time-consuming process to collect sufficient evidence, ensure fairness in investigation and adjudication, etc. In fast paced digital markets, taking such a long route to solve systemic problems does not seem feasible. Therefore, many have recommended complementing ex post enforcement mechanisms with ex ante regulatory framework.<sup>423</sup>

The recent Report on Digital Competition Law by the Ministry of Corporate Affairs has also recommended that ex-ante measures for antitrust concerns in digital markets should be introduced to support the ex-post framework. It was observed that ex post mechanisms cannot sufficiently cater to fast paced and dynamic digital markets and must be supported ex ante regulation. Ex post instruments cannot give timely and speedy redressal of anti-competitive activities in digital markets. The CCI has also noted the impact of big data in creating entry barriers in markets for entities that don't have access to big data thereby reducing competition.

---

<sup>421</sup> “Ex-Ante Regulation and Competition in Digital Markets: BEUC Contribution to the OECD Competition Committee Meeting” (BEUC, November 25, 2021) <<https://www.beuc.eu/position-papers/ex-ante-regulation-and-competition-digital-markets-beuc-contribution-oecd>> accessed January 28, 2025.

<sup>422</sup> Ania Thiemann and Gaetano Lapenta, ‘Ex ante regulation in digital markets’ (2021) OECD Secretariat 15/2021, <[https://one.oecd.org/document/DAF/COMP\(2021\)15/en/pdf](https://one.oecd.org/document/DAF/COMP(2021)15/en/pdf)> accessed on 30 January 2025.

<sup>423</sup> Standing Committee on Finance, *Anti- Competitive practices by Big tech companies* (Ministry of Corporate Affairs, 2022) pp. 27.

Ex ante regulations can be considered when ex post remedies are ineffective or inefficient in dealing with market failures. However, ex-ante rules need to be properly drafted and enforced for it to be effective in keeping markets competitive. Poorly enacted and rigid rules can be counterproductive and can harm competition and consumers.<sup>424</sup> The OECD discussions on ex ante regulations and ex post assessment in the digital market can very well be extended to the discussion on data governance. While antitrust assessment can be applied to instances of market inefficiencies in the data market, some believe that the case-by-case approach of antitrust mechanisms cannot tackle the systematic problem regarding data access and allocation that exists in the market.

## **2. International Approaches to NPD Regulation**

### **2.1. European Union**

The European Union was the first to specifically introduce a regulatory framework dealing with non-personal data. The proposal for the EU Directive on a Framework for Free Flow of Non-Personal Data was introduced with the objective to improve mobility of non-personal data across borders, to ensure regulatory control over sharing of data and to ensure ease in data portability.<sup>425</sup> The Directive aims to address the obstacles created for data mobility by legal, contractual and technical means.<sup>426</sup> These obstacles on data mobility create serious competition issues in the market and can affect innovation, research and development.<sup>427</sup> Therefore, to promote legal certainty & uniformity and to provide a level playing field within the EU, the Directive was introduced.<sup>428</sup> The Directive is said to play an important role in promoting data driven growth and innovation.<sup>429</sup> The Directive defines NPD as data other than personal data as defined in Regulation (EU) 2016/679.<sup>430</sup> Article 4 of the Directive prohibits data localisation requirements unless on the grounds of public security. Further, it also enables competent authorities to request access to data

---

<sup>424</sup> EX-ANTE REGULATION AND COMPETITION IN DIGITAL MARKETS – NOTE BY BIAC

<sup>425</sup> Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on a framework for the free flow of non-personal data in the European Union [EUR-Lex - 52017PC0495 - EN - EUR-Lex \(europa.eu\)](#)

<sup>426</sup> EUROPEAN PARLIAMENT AND COUNCIL REGULATION on a framework for the free flow of non-personal data in the European Union [2018] L 303/59, Recital 5.

<sup>427</sup> *ibid*, [Recital 6].

<sup>428</sup> *ibid*, [Recital 7].

<sup>429</sup> *ibid*, [Recital 6].

<sup>430</sup> *ibid*, [Article 3(1)].

for the performance of their official duties and in case the competent authority is denied access to data by the concerned user, proportionate interim measures may be imposed.<sup>431</sup> The Directive also promotes competition in the market by facilitating the development of self-regulatory codes for transparency and data portability.<sup>432</sup> The Commission also issued a guidance document for the said Directive in 2019 to help users in understanding the interaction between the EU Directive on Free Flow of Non-Personal Data and the GDPR. The purpose of the Guidance Document was to enhance legal certainty and clarity regarding data processing in the EU so that data can be utilized to its full potential.<sup>433</sup>

The EU introduced the Digital Markets Act (DMA) in 2022 with the objective to selectively regulate the behavior of large digital enterprises i.e. “Gatekeepers” providing one of the ten core platform services as mentioned in the Act. The DMA provides qualitative and quantitative thresholds to designate enterprises as “Gatekeepers”. Qualitative thresholds include entities that: *“(i) have a significant impact on the internal EU market, (ii) provide a core platform service which is an important gateway for business users to reach end users, and (iii) enjoys an entrenched and durable position in its operations, or it is foreseeable that it will enjoy such a position in the near future.”* The quantitative thresholds under the DMA include an entity having “(i) a ‘significant impact and a monetary threshold of either (a) annual EU turnover of at least EUR 7.5 billion in each of the last three financial years, or (b) its average market capitalisation or fair market value amounting to at least EUR 75 billion in the last financial year; (ii) An entity’s service can be presumed to be an ‘important gateway between business users and end users’ if in the last financial year, it has at least 45 million monthly active end users and 10,000 yearly active business users in the EU and; (iii) Finally, an entity is presumed to have an ‘entrenched and durable position’ if it meets the end user and business user thresholds above for each of the last three financial years.”

The Act puts the burden on the enterprises to inform the EC within two months of hitting the quantitative thresholds. Once an entity is designated as a “Gatekeeper”, the Act imposes ex ante obligations on them. Gatekeepers are obligated to allow interoperability of services with third

---

<sup>431</sup> *ibid*, [Article 5].

<sup>432</sup> *ibid*, [Article 5].

<sup>433</sup> Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union [eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019DC0250](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52019DC0250)

parties in certain circumstances. They are also required to provide access to the data their platforms generate so exclusive control over data can be maintained, and information asymmetry can be remedied.<sup>434</sup> Gatekeepers have to provide necessary tools and information to advertisers advertising on their platforms and allow other businesses to conclude contracts with their customers outside of their platforms. These obligations will help in ensuring that entities dependent on gatekeeper's platform get a fair and equal business environment as well as get new opportunities to compete with gatekeepers.<sup>435</sup> Furthermore, gatekeepers are prohibited from engaging in activities like tying, bundling, self-preferencing and cross using data across their various services unless consent has been acquired from data subjects. Moreover, Gatekeepers are also prohibited from preventing customers from "linking up their businesses to businesses outside their platforms". With this, the DMA aims to prevent gatekeepers from engaging in any unfair practices against their users or customers and easing the ability to switch to different platforms or services.

The EU Data Act will also come into force on 12 September, 2025. The Act was proposed to improve access to data across the EU market. It was recognised that data is an essential resource in the digital economy. However, despite the large volumes of data created, its potential value remains underutilized and in order to utilize this potential value and remove barriers to the development of the EU market, this regulation was introduced. The Act would facilitate free and fair flow of data from business-to-business, business-to-government, government-to-business and government to government, as had been highlighted in the European Strategy for Data in 2021.<sup>436</sup> The Act defined "data" as " any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audio-visual recording"<sup>437</sup> and would be applicable to manufacturers, data holders, data recipients, public sector bodies and providers of data processing services.<sup>438</sup> Notably, the Act is not applicable to

---

<sup>434</sup> Centre corporate-body. JR, "The EU Digital Markets Act- A report from panel of economic experts" (*Publications Office of the EU*, February 8, 2021) <[https://op.europa.eu/publication/manifestation\\_identifier/PUB\\_KJ0221116ENN](https://op.europa.eu/publication/manifestation_identifier/PUB_KJ0221116ENN)> accessed 28 January 2025.

<sup>435</sup> "The Digital Markets Act: Ensuring Fair and Open Digital Markets" (*European Commission*) <[https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets\\_en](https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en)> accessed January 28, 2025

<sup>436</sup> European Parliament, "A European Strategy for Data" (*Shaping Europe's digital future*) <<https://digital-strategy.ec.europa.eu/en/policies/strategy-data>> accessed January 28, 2025

<sup>437</sup> European Union Data Act 2024, a 2(1).

<sup>438</sup> *ibid*, [a 1(2)].

entities designated as gatekeepers under the DMA and gatekeepers are not allowed to request access to data.

The Act provides a harmonized framework to make data generated by use of product/services available to consumers and businesses on request, without delay and free of charge.<sup>439</sup> Further, to enable easy access to data, it must be ensured that products and services are designed in such a way that data can be easily made accessible. The Act will facilitate data access and provide legal certainty for sharing valuable data on FRAND terms<sup>440</sup> while aiming to also preserve the incentive to invest in data. To balance varied interests, the Act firstly protects trade secrets and states that trade secrets should only be disclosed when necessary measures are taken to protect its confidentiality.<sup>441</sup> And secondly, it provides that data recipients should not use the data shared to create a product/service that would compete with the product/service from where the data has originated from.<sup>442</sup> Chapter 5 of the Data Act makes it mandatory for firms to share their data with the government in exceptional circumstances of public emergency.<sup>443</sup> As per Recital 56 of the Data Act, “exceptional needs are circumstances that are unforeseeable and limited in time, in contrast to other circumstances which might be planned, scheduled, periodic or frequent”.

Provisions protecting small and medium enterprises can also be traced in the Act. The provisions concerning data access and data sharing will not be applicable to small and medium enterprises.<sup>444</sup> SMEs are also not bound to accept unilaterally imposed contractual terms that are in regard to access and use of data if they are unfair i.e. terms that grossly deviate from good commercial practices in data use and access.<sup>445</sup> Data holders are required to make data available to public sector bodies when there is an exceptional need to use data. These exceptional circumstances include responding to, preventing or recovering from a public emergency. This provision is also not applicable to small and micro enterprises.<sup>446</sup>

---

<sup>439</sup> *ibid*, [a 3].

<sup>440</sup> *ibid*, [a 8].

<sup>441</sup> *ibid*, [a 4(3)].

<sup>442</sup> *ibid*, [a 6].

<sup>443</sup> *ibid*, [a 15].

<sup>444</sup> *ibid*, [a 7].

<sup>445</sup> *ibid*, [a 13].

<sup>446</sup> *ibid*, [a 14].

## 2.2. Germany

Recently, Germany introduced two new amendments to the Competition Act to modernize the provisions of the Act and make it more proactive and focused on digital markets.<sup>447</sup> The competition authorities noted that the digital economy has a very dynamic nature and its rapid pace requires effective and faster control. The amendments will allow competition authorities to take preventive measures to curb big tech companies from engaging in conducts that can be harmful to the market. The amendments introduced a new section 19a to control Abusive Conduct of Undertakings of Paramount Significance for Competition Across Markets. Whether or not an entity is of paramount significance for competition depends on certain factors, i.e. “(1) its dominant position on one or several market(s), (2) its financial strength or its access to other resources, (3) its vertical integration and its activities on otherwise related markets, (4) its access to data relevant for competition and, (5) the relevance of its activities for third party access to supply and sales markets and its related influence on the business activities of third parties”.<sup>448</sup> The competition authorities can prohibit entities of paramount significance from engaging in activities that are anti-competitive. Such entities are prohibited from favoring its own offers over other competitors’ offers by either presenting their offers more favorably than others’ or exclusively pre-installing their offers on devices. Such entities are also prohibited from impeding other competitors from carrying out their activities in markets where the entities activities are of relevance for accessing the market. Therefore, activities like exclusive pre-installation, integration of offers or preventing other competitors from advertising their offers are prohibited. Such entities are further prohibited from impeding competitors (directly or indirectly) in markets where they are not dominant yet but could easily expand to. Additionally, entities of paramount significance are prohibited from creating entry barriers in the market or impeding other competitors by processing relevant data. Other prohibited acts (unless objectively justified) under section 19a include: refusing interoperability of services or data portability, providing other competitors insufficient information regarding services commissioned and, demanding benefits for handling the offers of another undertaking which are disproportionate to the reasons for the demand, in particular.<sup>449</sup> With respect to determining market power of an entity, section 18(3a) provides that in cases of multi-sided

---

<sup>447</sup> Act against Restraints of Competition (Competition Act – GWB) 1958.

<sup>448</sup> *ibid*, [s 19(1)].

<sup>449</sup> *ibid*, [s 19(a)].

market and networks following factors should be taken into account-direct and indirect network effects, “the parallel use of several services and the switching costs for users, the undertaking's economies of scale arising in connection with network effects, the undertaking's access to data relevant for competition, the undertaking's access to data relevant for competition”.

Through the 11th amendment, the competition authority's (Bundeskartellamt) powers were expanded. The law now requires entities to notify regarding mergers even if merger thresholds are not met to deal with mergers that can plausibly impede competition in the future. This obligation applies to cases where the acquiring entity's turnover exceeds EUR 50 million.<sup>450</sup> Furthermore, in instances of significant and continuous (i.e. existing for three years) malfunction of competition, the authority (Bundeskartellamt) after a sector inquiry, can address them and impose behavioral or structural remedies to reduce the malfunctioning of competition. These remedies very importantly include obliging entities to grant access to data, interfaces or other facilities and contractual arrangements for sharing of undisclosed information.<sup>451</sup> Lastly the amendment authorizes the *Bundeskartellamt* with powers to investigate into non-compliance of the Digital Markets Act (Regulation (EU) 2022/1925).

### **2.3. United Kingdom**

In 2019, the Furman Committee on Unlocking Digital Competition recognised the new challenges posed by anti-competitive behavior in digital markets and the need to make policy changes to make competition effective in digital markets.<sup>452</sup> The Committee recommended establishing a pro-competitive unit that can recommend code of conduct for digital markets for securing and promoting effective competition in digital markets. It also recommended that the merger thresholds and guidelines need to be amended to include considerations for harm to innovation and potential competition. The Committee also recognised the need to keep a track on the developments within the realm of machine learning and artificial intelligence to ensure that its use is not resulting in anti-competitive conduct. Similarly, it also recommended that the digital advertising market and its value chain must be investigated to assess whether effective competition exists or not.<sup>453</sup>

---

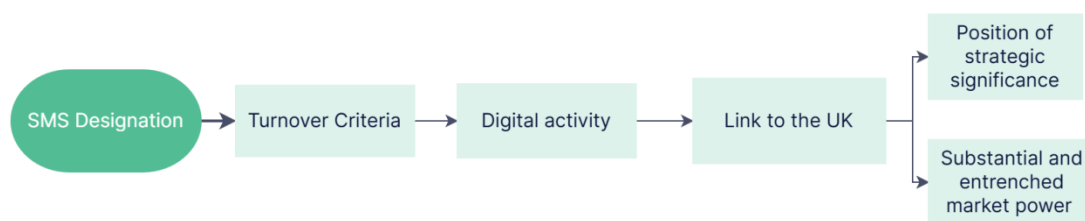
<sup>450</sup> *ibid*, [s 32f(2)].

<sup>451</sup> *ibid*, [s 32f(4)].

<sup>452</sup> Digital Competition Expert Panel, *Unlocking Digital Competition* (Open Government License, 2019).

<sup>453</sup> Digital Competition Expert Panel, *Unlocking Digital Competition* (Open Government License, 2019).

After the Expert Committee Report, the Draft Digital Markets, Competition and Consumers Bill (DMCC) was introduced with the objectives to promote effective competition in digital markets and expand the powers of competition authority. The Bill was aimed at addressing competition concerns in the digital economy particularly concerning large platforms. The Bill received royal assent in May 2024. The Act establishes a framework to regulate large platform with the aim to promote competition and protect consumer interest by regulation market power of firms with “Strategic Market Status” (SMS). It designates an entity as having SMS in respect of digital activity “linked to the UK”<sup>454</sup> if the entity has substantial and entrenched market power<sup>455</sup> and a position of strategic significance,<sup>456</sup> subject to turnover conditions.<sup>457</sup>



The Act grants Competition and Markets Authority (CMA) powers to impose Conduct Requirements (CR) on entities in order to guide their behaviour and promote competition in case the entity has been involved in anti-competitive conduct.<sup>458</sup> Conduct Requirements are guided by three core principles i.e. the fair dealing objective, the open choices objective and the trust and transparency objective. Each firm designated as SMS will get its own bespoke set of conduct requirement as opposed to EU’s blanket requirement that are common to all gatekeepers. These conduct obligations can include preventing the entity to use its position (including its access to data) to engage in self preferencing and using data unfairly.<sup>459</sup> The CMA can also intervene in cases it has reasonable grounds to suspect that digital activity is having an adverse effect on competition and can take proportionate measures to remedy, mitigate or prevent any detrimental effect on UK users.<sup>460</sup> The Act also creates a robust merger control regime for SMS entities. SMS

<sup>454</sup>Digital Markets, Competition and Consumers Act 2024, s 4.

<sup>455</sup> *ibid*, [s 5].

<sup>456</sup> *ibid*, [s 6].

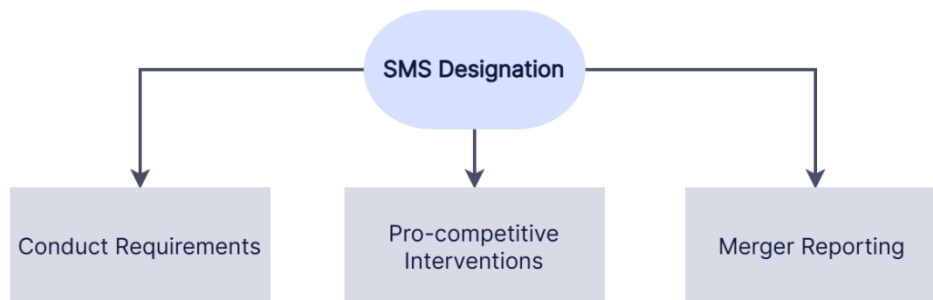
<sup>457</sup> *ibid*, [s 7].

<sup>458</sup> *ibid*, [s 6].

<sup>459</sup> *ibid*, [s 20].

<sup>460</sup> *ibid*, [s 46].

entities need to report mergers to the CMA in case they have a value of £25 million or more and is linked to the UK.<sup>461</sup> The CMA has broad information gathering powers and can request information via information notices. In cases of non-compliance of the information notice, the CMA also has a power to access the SMS entities’ premises, equipment, services, information and employees.<sup>462</sup> It has the power to impose fines up to 10% of global turnover of the firm in cases of non-compliance/breaches.<sup>463</sup>



## 2.4. Japan

The Improving Transparency and Fairness of Digital Platforms Act (TFDPA) came into force in February 2021. The Act acknowledges the increased importance of digital platforms in the field of information & telecommunication technology. It provides that the government should secure voluntary and proactive commitments from digital platform service providers that they would undertake to ensure fairness and transparency. Through this Act, Japan has taken a “co-regulation” approach wherein, along with a general framework, businesses and entities can take voluntary efforts.<sup>464</sup> The Act mandates digital platforms to disclose their terms and conditions of trading and to submit an yearly report of the voluntary actions and self-assessment. The disclosure requirement includes the requirement to disclose contract terms, scope and extent of obtaining and using data, etc. Furthermore, any change in their terms and conditions of trading has to be notified prior to their users. The Act provides that the administrative authorities under the Act i.e. The METI

<sup>461</sup> *ibid*, [s 57].

<sup>462</sup> *ibid*, [s 69].

<sup>463</sup> *ibid*, [s 86].

<sup>464</sup> “Key Points of the Act on Improving Transparency and Fairness of Digital Platforms (TFDPA) / METI Ministry of Economy, Trade and Industry” (*METI*) <[https://www.meti.go.jp/english/policy/mono\\_info\\_service/information\\_economy/digital\\_platforms/tfdpa.html](https://www.meti.go.jp/english/policy/mono_info_service/information_economy/digital_platforms/tfdpa.html)> accessed 29 January, 2025.

Minister is authorized to request the Japan Fair Trade Commission to take measures under the Anti-Monopoly Act in case a digital platform is under the suspicion of non-compliance of the Anti-Monopoly Act.<sup>465</sup>

Japan also introduced the Basic Act on the Advancement of Public and Private Sector Data Utilization in 2016. The purpose of the Act is to effectively utilize information circulated via the internet and other advanced information for the advancement of the public and move towards a society where the citizens can live safely and comfortably.<sup>466</sup> It was introduced to enable the state to effectively utilize public and private data. The Act provides that the government should make a basic plan for effective utilization of public and private sector data utilization<sup>467</sup> and the same should contribute to making Japan self-reliant, aid the local communities, create employment opportunities & new businesses and industrial development.<sup>468</sup> Companies are required to make voluntary efforts to endeavor to implement measures that would enhance public interest<sup>469</sup> and reduce the disparity among citizens in their ability to utilize information.<sup>470</sup> Under the Act, the government is required to establish a Basic Plan For Public And Private Sector Data Utilization to promote effective measures advancing utilization of public and private sector data.<sup>471</sup> This plan should include measures for data utilization, priority measures to implement data utilization and matters concerning data utilization at national administrative organs, local public entities and companies. This Basic Plan is to be reviewed by the government every fiscal year to consider changes in data utilization in light of effective measures.<sup>472</sup>

The Ministry of Agriculture, Forestry and Fisheries (MAFF) has also introduced Guidelines on Contracts Regarding Utilization of AI and Data in the Agricultural Sector with the intention to

---

<sup>465</sup> “Key Points of the Act on Improving Transparency and Fairness of Digital Platforms (TFDPA) / METI Ministry of Economy, Trade and Industry” (*METI*) <[https://www.meti.go.jp/english/policy/mono\\_info\\_service/information\\_economy/digital\\_platforms/tfdpa.html](https://www.meti.go.jp/english/policy/mono_info_service/information_economy/digital_platforms/tfdpa.html)> accessed 29 January, 2025.

<sup>466</sup> Basic Act on the Advancement of Public and Private Sector Data Utilization 2016, a 1.

<sup>467</sup> *ibid*, [a 8].

<sup>468</sup> *ibid*, [a 3].

<sup>469</sup> *ibid*, [a 11(2)].

<sup>470</sup> *ibid*, [a 14].

<sup>471</sup> *ibid*, [a 8].

<sup>472</sup> *ibid*, [a 8(7)].

promote smart agriculture in Japan. The guidelines provide model contracts for data receivers (like smart agriculture companies) and data providers (like farmers) so that data can be used efficiently and at the same time the threat of know-how leakage can be minimized. Data providers like skilled farms are afraid of data and know-how leakage and this fear can restrain them from sharing agricultural data. These guidelines enable them to feel comfortable while sharing their data.<sup>473</sup> Japan has also launched an Agricultural Data Collaboration Platform “WAGRI” in 2019 to improve the utilization of data in agriculture. The term WAGRI is a coined word denoting a platform that links various data and services to promote harmonization of various communities in Japan to promote innovation in the agricultural field. The primary intention of these steps is to fully and effectively realize the utility of data to increase productivity and sustainability in agriculture.<sup>474</sup> Further, Japan is also pursuing “Society 5.0” which is also called ‘super-smart’ society that envisions a sustainable, inclusive socio-economic system which is powered by digital technologies like big data analytics, artificial intelligence (AI), the Internet of Things and robotics.<sup>475</sup> While pursuing “Society 5.0” Japan is promoting “connected industries” with the aim to achieve industrial inter-connectedness that would provide new value and solutions for societal challenges. The initiative includes attempts to establish mechanisms for data sharing and data utilization across industries and national boundaries. The Ministry of Economy, Trade and Industry (METI) has named this initiative as the “Ouranos Ecosystem”.<sup>476</sup> The initiative is representative of an ecosystem where stakeholders can come together to collaborate to create new value.<sup>477</sup>

## 2.5. USA

The US deals with competition issues in the market through the Sherman Act, the Clayton Act and the Federal Trade Commission Act. Apart from these legislations, the US has introduced 12 new

---

<sup>473</sup> Policies : MAFF” (*MAFF-Ministry of Agriculture, Forestry and Fisheries*) <<https://www.maff.go.jp/e/policies/index.html>> accessed 30 January, 2025.

<sup>474</sup> *ibid.*

<sup>475</sup> “Japan Pushing Ahead with Society 5.0 to Overcome Chronic Social Challenges” (*UNESCO*, November 13, 2024) <<https://www.unes.co.org/en/articles/japan-pushing-ahead-society-50-overcome-chronic-social-challenges>> accessed 29 January, 2025.

<sup>476</sup> “Japan’s Initiatives for Interoperable Data Infrastructures Officially Named ‘Ouranos Ecosystem’” (*METI*, April 29, 2023) <[https://www.meti.go.jp/english/press/2023/0429\\_001.html](https://www.meti.go.jp/english/press/2023/0429_001.html)> accessed 29 January, 2025

<sup>477</sup> JIN Staff Writer, “Japan Launches Ouranos Initiative for Cross-Border Data Sharing and Collaboration” *Japan Industry News* (May 1, 2023) <<https://www.japanindustrynews.com/2023/05/japan-launches-ouranos-initiative-for-cross-border-data-sharing-and-collaboration/>> accessed 29 January, 2025.

bills to regulate competition in digital markets, a few of which are briefly discussed in this section. The Ending Platform Monopolies Bill aims to “promote competition and economic opportunity in digital markets by eliminating the conflicts of interest that arise from dominant online platforms’ concurrent ownership or control of an online platform and certain other businesses”. The Bill defines “covered platform” as a platform that (1) has at least 50,000,000 United States-based monthly active users on online platform; or at least 100,000 United States-based monthly active business users on the platform; (2) is owned or controlled by a person with net annual sales, or a market capitalization greater than \$600,000,000,000; and (3) is a critical trading partner for the sale or provision of any product or service offered on or directly related to the online platform.<sup>478</sup> As per the Bill, it is unlawful for covered platforms to own, control, or have beneficial interest in a line of business other than the covered platform that “(1) utilizes the covered platform for the sale or provision of products or services, (2) offers a product or service that the covered platform requires a business user to purchase or utilize as a condition for access to the covered platform, or as a condition for preferred status or placement of a business user’s product or services on the covered platform; or (3) gives rise to a conflict of interest.”<sup>479</sup> Noncompliance with the provisions of this Act would result in civil penalties. The Bill can prohibit big platforms from offering their own products alongside third-party sellers offering similar services or products.

The American Innovation and Choice Online Act aims to prohibit discriminatory conducts by covered platforms.<sup>480</sup> Unlawful conduct covered under the Act include conducts that provide an advantage to the covered platform’s own business over others, excluding or disadvantaging other business users or, discriminating among similarly situated business users. Some of the other prohibited acts include restricting other business users’ access to the features (hardware or software) being accessed by the covered platform for their own services, imposition conditions like purchasing other services of the covered platform to access the platforms, using non-public data to the benefit of their own services and products, etc.<sup>481</sup> Overall, the Act aims to prohibit digital platforms from preferencing their own products or services over other business users’ on

---

<sup>478</sup> Ending Platform Monopolies H.R. Bill (2021-2022) 3825, s 5(5).

<sup>479</sup> *ibid*, [s 2].

<sup>480</sup> American Innovation and Choice Online H.R. Bill (2021-2022) 3816.

<sup>481</sup> American Innovation and Choice Online H.R. Bill (2021-2022) 3816, s 2.

their platforms. It also prohibits platforms from self-preferencing and using non-public business information derived from third party apps to compete with them.<sup>482</sup>

The Open App Markets Act was introduced to promote competition by keeping a check on gatekeepers and by improving quality, decreasing consumer costs and increasing consumer choice. The Act aims to protect competition by prohibiting tying and bundling in app markets.

The U.S. also introduced the REPAIR Bill with the objective to ensure that all tools, equipment and information related to repair and maintenance of vehicles is available to the independent repair industry on FRAND terms. Since the automobile industry has progressed substantially and has integrated complex technologies in vehicles, the manufacturers possess substantial relevant vehicle data that is necessary for independent repairers. The Bill was introduced to ensure that independent repair providers are not denied access to vehicle data and other crucial information. The Act does this in three ways. Firstly, it prohibits manufacturers creating impediments for vehicle owners in accessing their data. Secondly, it requires manufacturers to provide access to the relevant data. Thirdly, the Bill provides that the safety, security and privacy of the consumers must be maintained while handling information.<sup>483</sup> The United States also enacted the Federal Energy Administration Act in 1974. The objective of the Act was to conserve energy supplies and ensure fair and efficient distribution of, and maintenance of free and reasonable consumer prices for energy sources. One of the provisions of the Act requires public disclosure of information that is necessary to keep the public informed regarding the nature, extent and duration of shortage of energy supplies and the relevant steps that are being taken to minimize its impact.<sup>484</sup>

## **2.6. China**

In order to develop the digital economy and protect competition, China introduced Guidelines for Internet Platform Categorizing and Grading in 2021. On the bases of the main functions of all the platforms and their linkage attributes, they were categorized into the following six categories.<sup>485</sup>

---

<sup>482</sup> American Innovation and Choice Online H.R. Bill (2021-2022) 3816, s 3(c) and 3(e).

<sup>483</sup> The Right to Equitable and Professional Auto Industry Repair (REPAIR) H.R. Bill (2021-2022) 6570.

<sup>484</sup> The United States Code 1926, 15 USC Ch. 16B.

<sup>485</sup> Graham Webster, Lorand Laskai, Rogier Creemers, Johanna Costigan, “Translation: Guidelines for Internet Platform Categorization and Grading (Draft for Comment) – Oct. 2021” (*DigiChina*, February 28, 2022)

| <b>CATEGORY</b>                | <b>LINKAGE ATTRIBUTES</b>           | <b>KEY FEATURES</b>            |
|--------------------------------|-------------------------------------|--------------------------------|
| Online Sale Platform           | Connecting people and goods         | Trading Features               |
| Life Service Platform          | Connect people to services          | Service Features               |
| Social Entertainment Platform  | Connect people                      | Social entertainment features  |
| Information Platform           | Connect people with information     | Information Function           |
| Financial Services Platform    | Connecting people with money        | Financing Function             |
| Computing Application Platform | Connect people with computing power | Network Computing Capabilities |

The online sales platform connects people with goods to facilitate trade. This may include several sub platforms like comprehensive commodity trading, vertical commodity trading or supermarkets. Life service platforms connect people to services to facilitate availing of services like travel services, tourism services, housekeeping services, etc. Social entertainment platforms connect people and facilitate social interaction, game and leisure activities, audio and video services, etc. Information platforms link people with information through news portals, search engines, new organizations, etc. Financial services platforms link people with money and serve a financing function. Its main functions include providing payment services, financial information, financial wealth management services, etc. Lastly, computing application platforms link people with computing power and this category may include platforms engaged in information management, research and development of operating systems, cloud computing services, etc.<sup>486</sup>

The Guidelines also classify platforms into Super Platforms, Large Platforms and Small and Medium Sized Platforms. This classification is done considering the platform's scale of users, types of services, market value and their capacity to restrict merchants' ability to reach consumers.

---

<<https://digichina.stanford.edu/work/translation-guidelines-for-internet-platform-categorization-and-grading-draft-for-comment-oct-2021/>> accessed 29 January, 2025.

<sup>486</sup> *ibid.*

Super platforms have an exceedingly large amount of user base, operate on exceedingly broad categories of operations, have high economic value and have a strong ability to restrict merchants' access to consumers. Large platforms have large user scale and have relatively broad range of operations and economic value and have a strong ability to restrict merchants' access to consumers. Lastly small and medium platforms have a certain user scale with limited operations, economic value and limited ability to restrict merchants' ability.<sup>487</sup>

Furthermore, the Cyber-security Administration of China (CAC) in 2023 introduced certain provisions for promoting and regulating Cross-border Data Flows. These provisions specify the conditions and processes that data processors need to comply with to transfer data out of China also known as "outbound data transfer". These provisions will have a significant impact on cross border flow of data. Under these provisions data processors have to undergo a security assessment before transferring certain data outside of China. This includes transfer of important data, transfer of personal information by critical information infrastructure operator (CIIO), transfer of personal data by data processor that has processed personal information of 1,000,000 or more subjects or, transfer of sensitive personal information of 10,000 or more subjects.<sup>488</sup> The self-assessment procedure requires data processors to conduct a self-assessment and prepare a report that is to be submitted to the CAC for review. The CAC while reviewing will consider factors like the data processor's legal compliance and whether the data protection laws and policies of the country in which data is to be transferred meet with the standards that apply in China. Once reviewed, the CAC will decide on the outbound transfer. If the application for outbound data transfer is approved by the CAC, the same remains valid for two years. If unsatisfied with the review, data processors can reapply for assessment, but the decision of the CAC remains final.

In 2023, China introduced four new regulations in furtherance of the recent amendments to the Anti-Monopoly Act. These regulations are: Regulation Prohibiting Monopoly Agreements, Regulation Prohibiting Conduct Abusing a Dominant Market Position, Regulation on the Review of Concentrations between Business Operators and Regulation Preventing Conduct Abusing

---

<sup>487</sup> *ibid.*

<sup>488</sup> Tim Hickman, Bob Li, Joe Devine, "New requirements for outbound data transfers from China" (*White & Case LLP*, 2 February 2, 2023) <<https://www.whitecase.com/insight-alert/new-requirements-outbound-data-transfers-china>> accessed 29 January, 2025.

Administrative Powers to Eliminate or Restrict Competition. Through these changes, the authorities have shown an intention to focus on and regulate the digital sector. The updates focus on digital platforms and have updated the ways in which relevant markets can be defined. The provisions state that the relevant markets may be designated on the basis of one side of the platform, on the basis of the goods involved in the platforms overall or on the basis of separate multiple relevant markets. Further, while determining cases of abuse of dominance by digital platforms, factors like their business model, user scale, network effects, lock in effects and the ability of the platform to master and process data can be considered.

## **2.7. Australia**

In 2020, the Australian Competition and Consumer Commission was directed by the Australian Government to inquire into digital markets. The Direction inquiry included search engine services, social media services, online messaging services, digital content aggregation services, media referral services, e-marketplace services and their data practices<sup>489</sup>. The ACCC was directed to provide biannual reports from 2020 to 2025 identifying practices in digital markets. The Digital Platform Survey Inquiry report released in September 2023 identified three risks to competition from the expanding digital markets, these are (1) leveraging position of market power (bundling, tying and self-preferencing (2) exclusionary data practices & its impact on innovation and, (3) strategic mergers and acquisitions.<sup>490</sup> The Report recommended that there is a need to consider a prohibition on economy wide unfair trading practices and the need to develop a new and stronger competition law regime with targeted obligations in service-specific codes of conduct with targeted obligations addressing issues like self-preferencing, tying, lock-ins and interoperability that would be applicable to designated digital platforms.<sup>491</sup>

In 2021, the Treasury Laws Amendment (News Media and Digital Platforms Mandatory Bargaining Code) Act 2021 was enacted in Australia. The Act is a mandatory code of conduct for

---

<sup>489</sup> “Digital Platform Services Inquiry 2020-25” (*Australian Competition and Consumer Commission*, November 27, 2023) <<https://www.accc.gov.au/inquiries-and-consultations/digital-platform-services-inquiry-2020-25>> accessed 29 January, 2025.

<sup>490</sup> “September 2023 Interim Report” (*Australian Competition and Consumer Commission*, September 30, 2023) <<https://www.accc.gov.au/inquiries-and-consultations/digital-platform-services-inquiry-2020-25/september-2023-interim-report>> accessed 29 January, 2025.

<sup>491</sup> *ibid.*

news businesses and designated digital platforms. The code enables the treasurer to designate digital platforms subject to certain obligations if there is a significant imbalance of power between the platform and news businesses. Till date, none of the platforms have been designated, however, it is said to act as a deterrent for digital platforms as since the enactment of the Act, platforms like Google and Facebook have entered into agreements with news businesses.<sup>492</sup>

In 2022, Australia also took a step towards regulating data by enacting the Data Availability and Transparency Act. The objectives of the Act are to promote availability of public sector data while complying with the privacy standards provided by the Privacy Act 1988. The Act aims to enhance integrity and transparency and build confidence in sharing of public sector data for public interest. The Act regulates sharing of public sector data and defines it as “lawfully collected, created or held by or on behalf of a Commonwealth body, and includes ADSP-enhanced data.” Therefore, it only covers data collected, created or held by commonwealth bodies (as defined under Public Governance, Performance and Accountability Act 2013)<sup>493</sup> and leaves data collected by private bodies and universities outside its scope. As per section 15 of the Act, data can be shared for purposes like delivery of government services, informing government policy and programs and research and development. The purpose for which the data is being shared must be specifically mentioned in the data sharing agreement.<sup>494</sup>

## **2.8. Other Ex Ante Initiatives**

Estonia has introduced the Information System Data Exchange Layer or the “X-Road” with the objective of enabling public and private organizations to exchange information online. X-Road is an interoperability platform<sup>495</sup> that facilitates use of data in a safe, secure and standardized manner. It connects different systems and different services and enables transfer of large amounts of data across multiple systems simultaneously.<sup>496</sup>

---

<sup>492</sup> “News Media Bargaining Code” (*Australian Competition and Consumer Commission*, May 2, 2022) <<https://www.accc.gov.au/by-industry/digital-platforms-and-services/news-media-bargaining-code/news-media-bargaining-code>> accessed 29 January, 2025.

<sup>493</sup> Data Availability and Transparency Act 2022, s 10.

<sup>494</sup> Data Availability and Transparency Act 2022, s 15.

<sup>495</sup> Hart P ’t and Compton ME, *Great Policy Successes* (Oxford University Press 2019).

<sup>496</sup> Lars E, “X-Road” (*e-Estonia*, June 10, 2024) <<https://e-estonia.com/solutions/x-road-interoperability-services/x-road/>> accessed January 29, 2025.

In France, the Open Data Legislation puts an obligation on companies to share the data they hold in public interest. The data may include “data used for procurement, commercial data for the establishment of official statistics, certain electricity and gas production and consumption data held by transmission and distribution systems operators for re-use by any other party, and certain data relating to changes in real estate ownership for re-use by certain third parties”. Such data is being defined as “public interest data”. The purpose of the proposed legislation is to “enhance the circulation of data and knowledge”.<sup>497</sup>

The Finnish Act on Transport Services attempts to promote competition and bring changes to the transport market. It puts an obligation on service providers to share essential data regarding mobility services. Providers of transport services must ensure that essential data is available freely, openly and in a computer readable format. This data should be inclusive of route information, stops, timetables, prices, availability and accessibility.<sup>498</sup> The Act also requires traffic control and management service providers to provide, via an open interface, information related to traffic-related meteorological data and weather forecasts, measurement data on traffic amounts, information on traffic flow and traveling times, information on disturbances and exceptional circumstances, other public information produced through public funding, etc.<sup>499</sup>

### **3. Indian Approach to Ex-Ante Regulation of Digital Market**

The Report of the Committee on Digital Competition Law observed that the current ex-post framework needs to be supported by an ex-ante regulatory framework. It noted that the Competition Act is not well suited to timely and speedily address anti-competitive conduct of digital enterprises. It also noted several issues that are peculiar to digital markets i.e. a) data as an asset and access to data as an entry barrier, b) market dominance of tech giants and issues like self-preferencing and, c) algorithmic and network effects. It recommended introducing the Digital

---

<sup>497</sup> Act on Transport Services 2017.

<sup>498</sup> Act on Transport Services 2017, Ch 2 s 1.

<sup>499</sup> Chapter 2 (579/2018) Information management of traffic control and management service providers: [Section 3](#) (579/2018) Disclosure of information via an open interface

Competition Act (DCA) that would aim to address the unique challenges posed by digital markets.<sup>500</sup> The DCA aims to identify “Systemically Significant Digital Enterprises” and subject them to additional scrutiny and regulation. The Act applies to a pre-defined list of core digital services that include core digital services, online search engines, video-sharing platforms, interpersonal communication services, operating systems, web browsers, cloud services, advertising services and online intermediation services.<sup>501</sup> The Act identifies SSDE based on a twin test: 1) the financial strength test: India-specific turnover (INR 4000 Cr), global turnover (USD 30 billion), global market capitalization (USD 75 billion), or gross merchandise value (INR 16000 Cr). and; 2) the significant market spread test: Enterprises with at least 1 crore end users or 10 lakh business end users.<sup>502</sup>

It is notable that even if the twin test is not satisfied, it can still be designated as SSDE if it has significant presence due to factors like number of users, economic powers, presence of entry barriers (including data barriers), market structure, etc. As per the draft Act, firms have an obligation to self-report their status if it fulfils the SSDE criteria.<sup>503</sup> Furthermore, the entities are prohibited from segmenting their services in order to circumvent the regulatory requirements.<sup>504</sup> The provisions of the draft Act obligate the SSDE’s to allow data portability to enable the users to transfer their data freely across platforms. It prevents SSDEs from favouring their own products and services over third party offerings.<sup>505</sup> SSDEs are prohibited from using or sharing the non-public data of business users to compete with them and they can also not mix datasets without the consent of the business users.<sup>506</sup> SSDEs are obligated to allow installation and use of third party applications without the imposition of unnecessary restrictions.<sup>507</sup> Furthermore, SSDEs cannot mandate tying and bundling of unrelated products or services unless it is integral to the digital

---

<sup>500</sup> Ministry of Corporate Affairs, *The Report of the Committee on Digital Competition Law* (2024).

<sup>501</sup> Draft Digital Competition Bill 2024, Sh 1.

<sup>502</sup> *ibid*, [s 3].

<sup>503</sup> *ibid*, [s 4].

<sup>504</sup> *ibid*, [s 5].

<sup>505</sup> *ibid*, [s 11].

<sup>506</sup> *ibid*, [s 12].

<sup>507</sup> *ibid*, [s 13].

service being provided.<sup>508</sup> Non-compliance of obligations under the Act would result in monetary penalties of up to 10% of the global turnover of the SSDE.

The Draft Act introduces a proactive ex ante regulatory model that aims to prevent anti-competitive behavior in digital markets. Although a step towards regulating the digital economy, the feedback from the stakeholders on the report was a mixed one. Amazon and uber raised concerns of over regulation, increased compliance cost and the risk of potentially stifling innovation. Some emphasized the privacy risks and the need to include data protection within the draft Act. Some stakeholders very stringently opposed having an ex-ante regulation at all.<sup>509</sup> The Act represents an effort to address antitrust issues in digital markets. It proposes measures to target anti-competitive conduct of dominant tech firms concerning issues related to data use, market power, self preferencing, etc.

- **Comparing the EU, UK and the Indian Approaches**

The Indian Digital Competition Bill, the UK's Digital Markets, Competition and Consumers Act 2024, and the EU's Digital Markets Act (DMA) each aim to address the challenges posed by digital markets with focus on promoting fair competition, preventing monopolistic behaviour and protecting users. Each of these have common goals however, are tailored to their respective regions. The primary goal of India's Draft Bill is to identify Systematically Significant Digital Enterprises (SSDEs) and Associate Digital Enterprises (ADEs) and regulate their conduct in core digital services markets on the basis of principles of contestability, fairness and transparency. It aims to promote foster innovation and protect user interests by regulating the conduct of SSDE entities providing core digital services that are identified in Schedule I of the Draft Bill.<sup>510</sup> The UK's DMCCA focuses on regulating the behaviour of entities providing core digital services with Strategic Market Status (SMS) with the aim to regulate competition in digital markets and protect consumer rights.<sup>511</sup> It designates powers on the Competition and Markets Authority (CMA) to

---

<sup>508</sup> *ibid*, [s 15].

<sup>509</sup> *ibid* n 501.

<sup>510</sup> Draft Digital Competition Bill 2024.

<sup>511</sup> Digital Markets, Competition and Consumers Act 2024.

impose conduct requirements and make pro-competitive interventions. The EU DMA targets “gatekeepers” that provide core digital services (pre-identified)<sup>512</sup> with the goal to ensure fair and contestable digital market, protect consumers and business users and ban unfair practices by large platforms.<sup>513</sup>

As per the Indian Bill, a twin test is used to identify SSDEs that considers both financial strength and significant spread test. It also provides that even if the twin test is not satisfied, an entity can be designated as SSDE due to factors like market structure, barriers to entry, network effects, etc. In the UK, the SMS status is designated if the digital activity of the firm is linked to the UK and it holds substantial & entrenched market power and holds a position of strategic significance. However, there is an initial threshold test that needs to be satisfied as a pre-requisite to designating an entity as SMS. In the EU, the DMA designates platforms as “gatekeepers” based on a turnover criteria of (annual turnover of €7.5 billion in the EU or a market valuation of up to €75 billion), if they provide a core platform services which acts as a gateway for business users to reach consumers, if they control core platform services in at least 3 EU countries, have a strong economic position and significant impact on the market, and have an entrenched market position.<sup>514</sup>

The SSDE’s in India are prohibited from engaging in anti-competitive conduct like self preferencing, restricting third party software/applications. They are also required to allow data portability & interoperability and refrain from using non-public data of business users to compete with them. SSDEs among other things are also expected to allow businesses to promote their services outside of the SSDE platform. In the UK, the DMCCA requires the CMA to impose bespoke conduct requirements on SMS firms that are driven by the principles of fair dealing, open choices and trust & transparency. Some specific obligations include ensuring data portability, interoperability and not engaging in self preferencing conduct. Along with conduct requirements, the SMS firms may also be made subject to pro-competitive intervention and merger reporting obligation. Similarly, the DMA also imposes these obligations to ensure fair competition and market contestability.

---

<sup>512</sup> *ibid*, [a 2(2)].

<sup>513</sup> Digital Markets Act 2022.

<sup>514</sup> *ibid*, [a 3].

All the three frameworks share a common goal to regulate digital market and ensure fair competition by preventing anti-competitive conduct, however the approaches do slightly differ. The Indian approach stresses on SSDEs with a two-pronged approach that combines financial and market strength criteria. The UK with its SMS designation focuses on imposition of bespoke conduct requirements and pro-competitive interventions by the CMA. The EU places similar emphasis on gatekeepers. All three approaches recognise the importance of controlling the conduct of digital enterprises and tackling market dominance. Each recognises the importance of data as an asset and subsequent data access issues.

| <b>Aspect</b>  | <b>India (Digital Competition Bill)</b>   | <b>UK (Digital Markets, Competition and Consumers Act 2024)</b>   | <b>EU (Digital Markets Act)</b>   |
|----------------|---|---|---|
| <b>Purpose</b> | Regulate <b>“Systemically Significant Digital Enterprises”</b>  | Ensure fair digital markets by regulating firms with <b>Strategic Market Status (SMS)</b> .   | Ensure fair and competitive digital markets by regulating <b>“gatekeeper”</b> firms.  |
| <b>Scope</b>   | Regulates <b>SSDEs</b> in core digital services to promote fair competition and prevent anti-competitive practices. | Regulates firms with <b>SMS</b> i.e. firms qualifying the turnover criteria and that have significant market power and impact competition in digital markets. | Regulates <b>Gatekeepers</b> , firms providing core platform services that act as critical gateways between businesses and consumers. |

| Aspect                                  | India (Digital Competition Bill)  | UK (Digital Markets, Competition and Consumers Act 2024)   | EU (Digital Markets Act)   |
|---|---|--|--|
| <b>Identification of Targeted Firms</b> | <p><b>Twin-test mechanism:</b></p> <p>Significant financial strength test (INR 4000 Cr turnover or USD 30 billion globally)</p> <p>+ Significant spread test (1 crore end users or 10 lakh business users).</p> <p>An entity can be designated as SSDE even if these requirements are not met if it has significant presence.</p> | <p>Turnover criteria</p> <p>+ Digital Activity</p> <p>+ Linked to the UK</p> <p>+ Position of strategic significance</p> <p>+ Substantial and Entrenched Market Power (SEMP)</p> | <p>Significant impact on the internal market (global turnover of €8 billion or €1 billion in the EU)</p> <p>+ Core platform service (an important gateway for business users to reach end users)</p> <p>+ Entrenched and durable position.</p> |
| <b>Reporting &amp; Transparency</b>     | <b>Self-reporting</b> by SSDEs  | CMA designates upon investigation.   | <b>Self-reporting</b> upon reaching the threshold.   |

#### **4. Indian Approach to NPD Regulation**

The Ministry of Electronics and Information Technology (MeitY) constituted an expert committee under the chairmanship of Mr. Kris Gopalankrishan to deliberate on a governance framework for Non-Personal Data in India. The committee released its first report on July 12, 2020. After receiving feedback and suggestions from the public, a revised report was released on December 16, 2020. Considering the increased importance and value generation capacity of data economy, the committee proposed a single national level regulation for non-personal data. A single national level regulation, according to the committee, will help in realizing the economic value & the utility of non-personal data that would help in generating economic value for the Indian community. Regulation of NPD will also lead to proper allocation of benefits accrued from processing of NPD, create an incentive for innovation, and address issues related to privacy and re-identification of anonymised datasets.

##### **4.1. Highlights from the Report**

The Report has defined non-personal data as data which is not ‘Personal Data’ (as defined under the PDP Bill, now DPDP Act), or the data is without any Personally Identifiable Information (PII). Non-personal data is the data which was either never related to an identified or identifiable natural person or data which was initially personal data, but was later anonymised using various data transformation techniques.<sup>515</sup> The Report provides that once the personal data has been anonymised, it will not fall within the scope of the DPDP Act. However, if the anonymised data is de-anonymised, it will fall within the purview of the DPDP Act.

The Committee Report has proposed a new horizontal classification for business called “Data Business”. A Data Business includes both private and public organizations that collect, process, store, or manage personal or non-personal data.<sup>516</sup> The Committee suggests that data businesses above a certain threshold would share the meta-data about the data being collected, stored and processed by them<sup>517</sup> and the same will be stored in a meta data directory. The submitted meta data

---

<sup>515</sup> MeitY, Revised Report by the Committee of Experts on Non-Personal Data Governance Framework (24(4), 2020), para 4.1.

<sup>516</sup> *ibid*, para 6.1.

<sup>517</sup> *ibid*, para 6.3.

will be analyzed by data trustees so that they can identify data from different data businesses that can be combined to create greater community benefit.<sup>518</sup>

The Committee Report proposes a mechanism to establish rights over non-personal data that is being collected and created in India. These rights include the right to derive economic and other value from NPD to maximize community benefit and the right to eliminate/minimize harms to the community.<sup>519</sup> This right will be exercised by the community. A community here means “*any group of people that are bound by common interests and purposes and involved in social and/or economic interactions.*” The objective behind allocating benefits to the community is to ensure that the benefits accrued from processing of non-personal data do not remain limited or restricted to the organizations collecting and processing the said data.<sup>520</sup> In furtherance of this, the committee has recommended creation of data custodians, data processor, high value datasets, data trustee, and a NPD authority. A data custodian is defined as an entity that collects, stores or processes any type of data and has a duty of care towards the community while handling non-personal data.<sup>521</sup> Data custodians have been said to have a duty of care towards the community to ensure that re-identification of individuals does not harm them or the community and that technology is being used with care.

A data processor is defined as a company that processes data on behalf of a data custodian and under the framework is not required to share data.<sup>522</sup> Further, high value datasets (HVD) are defined as those datasets that are beneficial to the community and should be shared as a public good.<sup>523</sup> Organizations responsible for creating and maintaining HVDs in India have been termed as data trustees and owe a duty of care towards the community in handling the data.<sup>524</sup> The committee recommended creating a separate NPD authority that would ensure benefits of NPD are shared with the community and would address other issues concerning NPD in India.

---

<sup>518</sup> *ibid*, para 6.4.

<sup>519</sup> *ibid*, para 7-7.1.

<sup>520</sup> *ibid*, para 7.2.

<sup>521</sup> *ibid*, para 7.4.

<sup>522</sup> *ibid*, para 7.5.

<sup>523</sup> *ibid*, para 7.6.

<sup>524</sup> *ibid*, para 7.7.

Finally, the Committee suggested that data may be shared for three purposes. Firstly, sovereign purposes like national security, law enforcement, infrastructure security, etc. Secondly, for public good purposes like research and innovation, policy development, better delivery of public services, etc. Thirdly, business purposes where data would be shared between two private entities.

#### **4.2. Criticisms of the Report**

While the objective behind establishing a committee and creating a draft framework for use of NPD for public welfare is commendable and welcomed, it fails to properly address nuanced issues surrounding the subject. The expert committee report was appreciated and welcomed by some but was criticized by many. Commenters highlighted that there are multiple definitional issues in the framework. Concepts like NPD, community NPD, private NPD, data trustee, etc. were termed as unambiguous and unclear. The report does not correctly grasp the intersection between data, intellectual property and proprietary rights which could result into unintended negative effects for businesses.<sup>525</sup>

The definition of NPD as provided in the Governance Report suffers from the same definitional problems as discussed in one of the previous chapters. The Report defines NPD as data which is not personal data (as described under the PDP bill) or when the data is without any personally identifiable information (PII).<sup>526</sup> The instability of the character of data as personal or non-personal can be traced in the Report. Personal data, when anonymised, falls outside the purview of PDP bill. However, since it is possible to re-identify anonymised data, the moment data is re-identified, it falls within the purview of PDP bill. This fluidity in the character and definition of the dataset raises concerns related to application of the appropriate law. Therefore, there is a need to take a contextual approach while defining NPD.

---

<sup>525</sup> Alliance for an energy efficient economy, *Comments on Report by the Committee of Experts on Non-Personal Data Governance Framework* (CL&ES, 2020), Also see Anupriya Dhonchak, 'Revised Non-Personal Data Governance Framework and Intellectual Property Implications – Part I' (SpicyIP, 2021) <https://spicyip.com/2021/02/revised-non-personal-data-governance-framework-and-intellectual-property-implications-part-i.html> accessed 3 February 2025.

<sup>526</sup> MeitY, Report by the Committee of Experts on Non-Personal Data Governance Framework (24(4), 2019).

Further, it is also being said that open access to data is not enough of a solution to offset the existing power imbalance in the digital sector. Certain organizations like Amazon and Facebook have natural monopolies and data sharing does not essentially offset the effect of such key players on competition. Moreover, critics argue that the report takes a flawed approach towards competition law. The competition authorities have imposed the obligation the “duty to supply” in cases where the competitors are unable to duplicate the facility and access to the facility is crucial for the competitors to compete in the market. If either of these two factors are not met with, the “duty to share” cannot be imposed. The report essentially targets dominance and not abuse of dominance and the same is being perceived as penalizing success. Further, data sharing will not remove the incentive for organizations to collect data which can result in data maximization.<sup>527</sup>

The report proposes that data can be shared for three purposes i.e. sovereign purposes, public good purposes and business purposes. The report allows the government to collect data purposes relating to national security, legal purposes etc.<sup>528</sup> However, the language of the report is very broad and can raise issues regarding state surveillance and government overreach.<sup>529</sup> Data can be requested for community uses/benefit, research and innovation, policy development, better delivery of public delivery services, *etc.* The report again uses the term “etc” while noting the purposes for which data can be shared.

Furthermore, the Report’s recommendation to create an independent NPD authority has also attracted some criticism. It has been argued that creation of a regulatory agency is not an appropriate step at this point and should be re-visited once there is more clarity and consensus concerning NPD governance. It has also been argued that in case an NPD authority is established it must be independent from the government, accountable and transparent, separate from judicial and enforcement functions and inter-regulatory body coordination must be ensured. India has

---

<sup>527</sup> Aman Nair, Pallavi Bedi and Anubha Sinha, Comments on the Report on the Non-Personal Data Governance Framework (The Centre for Internet Society, 2021).

<sup>528</sup> Report by the Committee of Experts on Non-Personal Data Governance Framework (24(4), 2019).

<sup>529</sup> Anupriya Dhonchak, 'Revised Non-Personal Data Governance Framework and Intellectual Property Implications – Part II' (SpicyIP, 2021) <https://spicyip.com/2021/02/revised-non-personal-data-governance-framework-and-intellectual-property-implications-part-ii.html> accessed 3 February 2025.MeitY, Also see Author G, “Five Key Concerns with India’s Non-Personal Data Report” (*MEDIANAMA*, July 22, 2020) <<https://www.medianama.com/2020/07/223-five-key-concerns-with-indias-non-personal-data-report/>> accessed 29 January, 2025.

introduced many laws and policies for open data and data sharing. These initiatives include Digital India Policy, National Data Sharing and Access policy, the Open Government Data Platform, etc. Even though these policies exist, they have failed to effectively achieve their objectives. The data shared by the government often lacked in quality and quantity. The causes behind the failure of these policies should have been explored by the expert committee.<sup>530</sup> Before introducing new regulations for data sharing, causes for failure of previous laws must be examined in order to ensure that new laws have a strong foundation. Due to these reasons and complexities within the framework, many suggested that there is a need for more consultation to create a regulation that would efficiently address NPD regulation in India. There is a need for a balance approach that can support smaller data businesses and avoid excessive regulatory burden that can stifle innovation.

## **5. Conclusion**

The debate on whether to adopt ex ante or ex post regulatory mechanisms to effectively deal with complex problems related to data and data sharing is gaining traction. Ex ante regulations, which proactively address potential issues, may be necessary to complement ex post mechanisms, which react to existing problems. A hybrid approach that combines both strategies could provide a more comprehensive framework for managing the complexities of big data in the digital economy. Several jurisdictions including Germany, EU, Japan, China, etc. have enacted regulations on data and data sharing. India's approach to non-personal data regulation represents a forward-thinking strategy to promote data sharing while ensuring economic and public benefits. The Ministry of Electronics and Information Technology (MeitY) formed an expert committee on Non Personal Data Governance Framework in India. The committee explored the regulatory aspects of Non Personal Data and emphasized the importance of data sharing for public interest and economic purposes. Among other suggestions, the committee also proposed to create a community right over data with the aim to democratize the benefits of data processing, ensuring that they are not confined to data holders alone. However, the report faced criticism for definitional ambiguities and its approach to competition law. Critics argue that the focus should be on preventing the abuse of

---

<sup>530</sup> Pastora Valero, "This Is Why We Should Care about the Free Flow of Data between Countries" (*World Economic Forum*, July 6, 2016) <<https://www.weforum.org/stories/2016/07/free-flow-of-data-between-countries/>> accessed 29 January, 2025.

dominance rather than targeting dominance itself. This distinction is crucial for developing a fair and effective regulatory framework that promotes competition without stifling innovation.

## **CHAPTER VI: REGULATORY AUTONOMY FOR NON- PERSONAL DATA: NAVIGATING THE PATCHWORK OF GLOBAL DATA GOVERNANCE**

### **1. Introduction**

The digital age has become a core part of economic growth, innovation and trade. The fast-paced evolution of data driven technology has however outpaced the development of a comprehensive legal framework. The transfer of data across borders plays a crucial role in trade, economic growth and innovation.<sup>531</sup> However, measures like restriction of cross border data flows and data localization can create trade barriers that may violate WTO obligations.

The previous few chapters explored the complex interplay between data and its interface with intellectual property rights, antitrust issues in digital markets, and the omnipresent challenges of data theft & cyber security issues. Laws concerning Intellectual property, antitrust and cyber security are predominantly domestic resulting in varied approaches across jurisdictions. For instance, the Trans-Pacific Partnership Agreement has included provisions for cross border data flow and a prohibition on data localisation measures. However, such measures are not universal. India's digital trade policies include data localization laws that can enhance the government's ability to control data and the same can be in conflict with GATS obligations. Further, countries are also adopting measures to reduce cyber threats but these measures may also act as trade barriers and be GATS inconsistent if the GATS exceptions can't be invoked to justify them sufficiently. In this context, this chapter aims to address the divergence in data governance and the patchwork caused by national laws and FTAs. It further examines the pressing need for harmonization of these laws at a multilateral level for a unified global framework for data governance. The importance of international cooperation in creating a cohesive system to address the challenges of data governance is emphasized.

---

<sup>531</sup> "WTO" (*The WTO in brief*) <[https://www.wto.org/english/thewto\\_e/whatis\\_e/inbrief\\_e/inbr\\_e.htm](https://www.wto.org/english/thewto_e/whatis_e/inbrief_e/inbr_e.htm)> accessed 29 January, 2025.

## 2. WTO Framework and its Limitations

The World Trade Organization at a multilateral level has struggled to keep up with technological developments in the digital world. Existing WTO instruments like the General Agreement on Tariff and Trade (GATT), the General Agreement on Trade in Services (GATS) and the TRIPS Agreement are not equipped to address the distinctive challenges related to cross border data flows, data localization and digital trade. This chapter aims to highlight the possible inconsistencies of national laws with WTO obligations by pointing out how measures like mandatory data sharing and data localization can conflict with WTO obligations. By exploring the regulatory autonomy of nations it highlights the need for an effective framework that can balance national interests and global trade.

The World Trade Organization (WTO) was established in 1995 and is the only organization dealing with international rules for trade aiming to provide assurance, stability in trade and eliminate trade barriers.<sup>532</sup> The fundamental principles guiding the WTO multilateral trading system are trade without discrimination (most favored-nation (MFN)), national treatment, reducing trade barriers & increasing predictability, transparency and promoting fair competition.<sup>533</sup> WTO law has a substantial impact on the domestic laws of the member states. Unlike other organizations, the WTO laws are referred to as “hard” law due to its effective dispute resolution system wherein the decisions of the WTO panel are considered enforceable and breach of WTO obligations is termed as “punishable”.<sup>534</sup> The obligations under the WTO, can impact national laws of the member states but at the same time, it also provides them with general exceptions to derogate from WTO obligations.

The WTO, at the heart of its multilateral trading system has three main pillars i.e. General Agreement on Tariffs and Trade (GATT), General Agreement on Trade in Services (GATS) and Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS). The fundamental

---

<sup>532</sup> *ibid.*

<sup>533</sup> “Principles of the Trading System” (WTO) <[https://www.wto.org/english/thewto\\_e/whatis\\_e/tif\\_e/fact2\\_e.htm](https://www.wto.org/english/thewto_e/whatis_e/tif_e/fact2_e.htm)> accessed 30 January, 2025.

<sup>536</sup> Mira Burri, “The Governance of Data and Data Flows in Trade Agreements: The Pitfalls of Legal Adaptation” (2017) 51 UC Davis Law Review 65.

<sup>534</sup> *ibid.*

principles of the WTO can be traced in each of these agreements and in the context of digital trade and data flows, all the three agreements have their own relevance. However, as of now, there is a lack of certainty concerning data governance at the WTO level. Additionally, as discussed in previous chapters, laws governing intellectual property, antitrust activities and cyber security across jurisdictions differ from one another. This divergence in legal approaches across jurisdictions raises a question concerning harmonization of laws at a global level for uniformity in data governance and cyber security. This section discusses the divergence in data governance across jurisdictions in light of WTO principles and attempts to address the inability of WTO to deal with cyber security issues in a technologically advanced world.

## 2.1. GATT and Digital Trade

The WTO Work Programme on E-Commerce was adopted in 1998 and since then the WTO members have engaged in the practice of not imposing custom duties on electronic transmissions. The moratorium applies to “electronic transactions” and covers anything transmitted by telecommunications, including communications made over the internet (from emails to video games). Significant benefits of the moratorium have been noted by the WTO. It has helped in reducing the cost of trade, increasing access to knowledge & software tools and providing new opportunities for businesses and consumers.<sup>535</sup> A 2019 study found that if custom duties were to be imposed in countries like India, South Africa, Indonesia and China, it would result in an increase in prices and a reduction in consumption which would eventually bring down the GDP growth, tax revenues<sup>536</sup> and increase unemployment.<sup>537</sup> However, it has been observed that other than the “recurring decision” to extend the moratorium, nothing much has been done<sup>538</sup> and despite the recognition of the importance of digital trade by the WTO, most of the focus has been on the

---

<sup>535</sup> Op-ed, “Understanding the WTO E-Commerce Moratorium” (*World Commerce Review*, March 26, 2024) <<https://worldcommercereview.com/understanding-the-scope-definition-and-impact-of-the-wto-e-commerce-moratorium/>> accessed January 30, 2025.

<sup>536</sup> “Electronic Transmissions and International Trade - Shedding New Light on the Moratorium Debate” (*OECD*) <[https://www.oecd.org/en/publications/electronic-transmissions-and-international-trade-shedding-new-light-on-the-moratorium-debate\\_57b50a4b-en.html](https://www.oecd.org/en/publications/electronic-transmissions-and-international-trade-shedding-new-light-on-the-moratorium-debate_57b50a4b-en.html)> accessed January 29, 2025.

<sup>537</sup> Timothy, “WTO Plurilateral Negotiations on Trade-Related Aspects of Electronic Commerce - ICC Baseline Position” (*ICC - International Chamber of Commerce*, June 18, 2019) <<https://iccwbo.org/news-publications/policies-reports/wto-plurilateral-negotiations-trade-related-aspects-electronic-commerce-icc-baseline-position/>> accessed January 29, 2025.

<sup>538</sup> Janow ME and Mavroidis PC, “Digital Trade, E-Commerce, the WTO and Regional Frameworks” (2019) 18 *World Trade Review* S1.

service paradigm under the General Agreement on Trade in Services (GATS) since digital trade does not involve tangible products. The GATS, like GATT aims to enhance competition in global trade but in the service paradigm. However, unlike GATT's general MFN and national treatment obligation, GATS has specific or positive commitments that are accepted individually by the member states in their "schedule of specific commitments".

## 2.2. GATS and Digital Trade

The General Agreement on Trade and Tariffs (GATS) provides a schedule of commitments to member states and when members schedule a services commitment under GATS, they shall accord national treatment and market access for the delivery of the service. Similarly the MFN principle applies unless the member states schedule an exception.<sup>539</sup> GATS covers all service sectors and of which, the ones relevant for digital trade are telecommunication services, computer and related services, audio visual and financial services.<sup>540</sup> However, it has been suggested that owing to the technological developments and development of new services, there might be a need to revisit the classification of services in order to avoid blurring of lines between services.<sup>541</sup> Among the above mentioned services, computer and related services have the most potential to regulate digital trade. As per the schedule of services, computer and related services include data processing services and database services.<sup>542</sup> Member states have made "far reaching commitments for both market access and national treatment under GATS" since during the time of negotiations, computer and related services was not a contentious issue because not many trade barriers and regulations on computer and related services existed during that time.<sup>543</sup> Due to these commitments, the member states don't have much room to make different domestic regulations and creation of contemporary

---

<sup>539</sup> Meltzer JP, "Cybersecurity, Digital Trade, and Data Flows: Re-Thinking a Role for International Trade Rules" *Brookings* (November 13, 2019) <<https://www.brookings.edu/articles/cybersecurity-digital-trade-and-data-flows-re-thinking-role-for-international-trade-rules/>> accessed January 29, 2025.

<sup>540</sup> Government of India, 'List of Services Sectoral Classification' (Ministry of Commerce & Industry, November 2020) <https://commerce.gov.in/wp-content/uploads/2020/11/List-of-Services-Sectoral-Classification-1.pdf> accessed 30 January 2025.

<sup>541</sup> Lee Tuthill & Martin Roy, 'GATS classification issues for information and communication technology services' (2012) in M. Burri & T. Cottier (Eds.), 'Trade Governance in the Digital Age: World Trade Forum' Cambridge University Press, (pp. 157–178).

<sup>542</sup> Ibid.

<sup>543</sup> Mira Burri, "The Governance of Data and Data Flows in Trade Agreements: The Pitfalls of Legal Adaptation" (2017) 51 *UC Davis Law Review* 65.

trade barriers like data localisation could be GATS inconsistent.<sup>544</sup> It is notable that GATS violative requirements like data localisation have not been examined by the WTO Panel yet. Some WTO members like the EU have justified this by arguing that digital services would fall under the category of audiovisual services because they “inherently function as content platforms”.<sup>545</sup> Further, this is also motivated by the fact that, in the audiovisual service sector, not many WTO members have made commitments and therefore, there is a lot more room for them to make regulations that can possibly be discriminatory or otherwise violative of WTO obligations.<sup>546</sup>

### 2.3. Lingered WTO Developments

The WTO laws, despite being hard laws governed by principles of non-discrimination and free trade, have not yet provided any legal certainty regarding data and data flows. It is in this lack of development on a multilateral level, new models like national laws, foreign trade agreements, bilateral, plurilateral and regional arrangements are developing to address new trade barriers and facilitate digital trade across the globe.<sup>547</sup> Post the development of the digital economy and commodification of data, data is being increasingly seen as an asset and nation states have started to assert their control and sovereignty over the digital world. The approaches of nations for governing the digital economy differ. For example, the EU has a right-based approach wherein it protects personal data of the citizens while encouraging and facilitating the growth of the digital markets. On the other hand, China has a “centralized governance style” or a “state based approach” wherein data is treated as a part of broader policies like national security, infrastructural autonomy and general socio-economic goals to improve the life of the citizens.<sup>548</sup>

Further, since there is a lack of clarity on an international level regarding standards governing data and data flows, the new instruments have introduced measures like data localisation that require

---

<sup>544</sup> *ibid.*

<sup>545</sup> Sen N, “Understanding the Role of the WTO in International Data Flows: Taking the Liberalization or the Regulatory Autonomy Path?” (2018) 21 *Journal of International Economic Law* 323.

<sup>546</sup> *ibid* n 545.

<sup>547</sup> Hodson S, “Applying WTO and FTA Disciplines to Data Localization Measures” (2018) 18 *World Trade Review* 579.

<sup>548</sup> Castellano GG, Selga and ĘK and Arner DW, “The Emergence of Financial Data Governance and the Challenge of Financial Data Sovereignty,” See also *Data Sovereignty* (Oxford University Press New York 2023) <<https://doi.org/10.1093/oso/9780197582794.003.0009>> accessed January 29, 2025.

the data to be stored within the national borders.<sup>549</sup> Such restrictive measures restrict trade flows and create modern barriers to digital trade (Explained later in this chapter). These measures raise two crucial questions for WTO members. Firstly, whether these restrictions are covered under the WTO regime and would they be in violation of the WTO principles or can they be covered under WTO exemptions. Secondly, whether there is a need to develop a multilateral framework governing data and data flows.<sup>550</sup> These initiatives act as a patchwork to the lack of clarity on an international level. However, this also results in fragmentation of a global data governance framework.<sup>551</sup>

#### **2.4. WTO Exceptions and Data Governance**

Data localisation measures are one of the most common measures adopted by member states domestically and through FTAs. Data localisation refers “*to a mandatory legal or administrative requirement directly or indirectly stipulating that data be stored or processed, exclusively or non-exclusively, within a specified jurisdiction*”.<sup>552</sup> It is considered to be one of the most restrictive measures for cyber security as it can restrict cross border data flow and impose trade barriers on foreign companies by imposing unreasonable and excessive compliance requirements that can curtail market access.<sup>553</sup> For example, China’s cybersecurity law imposes restrictions on cross border data flows and requires data localisation. It also requires foreign and development facilities to be placed in China to obtain market access. The US flags it as imposition of trade barriers for foreign entities that are unwilling to share their data with countries like China where cybersecurity risks are high. The U.S. Chamber of Commerce noted that imposition of data localisation requirements on companies to simply conduct their business will put the data at the risk of

---

<sup>549</sup> Hodson S, “Applying WTO and FTA Disciplines to Data Localization Measures” (2018) 18 World Trade Review 579.

<sup>550</sup> Ibid n 545.

<sup>551</sup> Castellano GG, Selga and ĘK and Arner DW, “The Emergence of Financial Data Governance and the Challenge of Financial Data Sovereignty,” *Data Sovereignty* (Oxford University Press New York 2023) <<https://doi.org/10.1093/oso/9780197582794.003.0009>> accessed January 29, 2025.

<sup>552</sup> “Data Localisation Trends and Challenges” (*OECD*, December 22, 2020) <[https://www.oecd.org/en/publications/data-localisation-trends-and-challenges\\_7fbaed62-en.html](https://www.oecd.org/en/publications/data-localisation-trends-and-challenges_7fbaed62-en.html)> accessed January 29, 2025.

<sup>553</sup> Mishra N, “The Trade: (Cyber) Security Dilemma and Its Impact on Global Cybersecurity Governance” (2020) 54 Journal of World Trade 567.

misappropriation.<sup>554</sup> Similarly in 2019 Japan expressed its concern for Vietnam and China’s cyber security measures including data localisation and restriction on cross border data flows at the WTO “cluster” of service meeting.<sup>555</sup> In the context of such measures being taken, a question arises concerning their validity under WTO norms.

As mentioned previously, GATS does provide member states with some flexibility under Article XIV and XIV *bis* and it is very probable that member states would defend cybersecurity measures using these exceptions. Article XIV *bis* provides member states with a security exception wherein they cannot be required to furnish information the disclosure of which they believe to be against their security interests. In the Russia Transit Case, the WTO Panel laid down a comprehensive test for Article XXI.<sup>556</sup> The Panel notes that “Essential Security Interests” refer to “quintessential functions of states” i.e. to protect its territory from external threat and to maintain internal law and order. A state’s security interests keep changing and it can decide its own security interest and therefore, the provision has a “self-judging” element. However, the Panel clarified that the security exception under GATT Article XXI is not entirely “self-judging” because “*the phrase ‘it considers’ is qualified by essential security interests limited to specific scenarios, namely, related to military facilities, nuclear facilities and measures taken in time of war or other emergency in international relations*”.<sup>557</sup> Further, the Panel interpreted the term “war” as “an armed attack” and, “emergency in international relations” as a “situation of armed conflict” or “heightened tension or conflict”. It is also settled that the security exception has to be adopted in good faith and not to avoid WTO obligations or to set disguised barriers to trade.<sup>558</sup>

---

<sup>554</sup> “Office of Trade & Manufacturing Policy Report: ‘How China’s Economic Aggression Threatens the Technologies and Intellectual Property of the United States and the World’ – The White House” (*The White House*) <<https://trumpwhitehouse.archives.gov/briefings-statements/office-trade-manufacturing-policy-report-chinas-economic-aggression-threatens-technologies-intellectual-property-united-states-world/>>.

<sup>555</sup> “WTO Members Hold Latest ‘Cluster’ of Services Meetings” *2019 News items - WTO members hold latest “cluster” of services meetings* <[https://www.wto.org/english/news\\_e/news19\\_e/serv\\_21mar19\\_e.htm](https://www.wto.org/english/news_e/news19_e/serv_21mar19_e.htm)> accessed January 29, 2025.

<sup>556</sup> General agreement of Tariffs and Trade 1947, Art XXI.

<sup>557</sup> WTO, *Russia- Measures Concerning Traffic in Transit* [2017] WT/DS512.

<sup>558</sup> Panos Delimatsis and Olga Hrynkiv , ‘Security Exceptions under the GATS – A Legal Commentary on Article XIVbis GATS’ (2020) TILEC Discussion Paper No. 2020-026 <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3757455#](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3757455#)> accessed 27 January 2025.

While examining the validity of cyber security measures under the National security exception, it has been observed that cyber security measures that are trade restrictive do not fall within the scope of the security exception. For cyber security measures to qualify under the security exception, the resulting damage of cyber security risks should be of a similar nature to a conventional military war. A critical issue that arises regarding cyber security measures taken in situations of war is that cyber-security risks are permanent and can happen at any point in time. This raises a peculiar question that can the exception be permanently available to members because they are always anticipating an attack. When it comes to cyber security measures, it becomes slightly peculiar to distinguish genuine measures from protectionist measures. Cyber security policies are generally adopted for long term goals. Restricting data flows for preventing cyber-attacks can possibly be for protection of essential security interests but cyber-attacks are generally not imminent therefore, it can become difficult to show that the measure was taken "in the time of emergency". Furthermore, with the anticipation of a likely threat of cyber-attack at any time, the exception is unlikely available.<sup>559</sup> On the other hand, if a broad interpretation is given to "war" or "emergency in international relations" while understanding the idiosyncrasies of the nature of data and the uncertainties of cybersecurity threats, restriction on cross border data flows and data localisation measures can qualify under the exception. Therefore, if the exception is given a broad interpretation it may apply to cyber security measures. However, it is unlikely that the exception can be applied to protective measures that are motivated to protect economic interests in digital markets or disproportionately benefit domestic industries.<sup>560</sup>

The GATS also provides certain General Exceptions under Article XIV which include measures: (a) protecting public morals/maintaining public order, (b) protecting human, animal or plant life or health, (c) secure compliance with laws or regulations relating to prevention of deceptive and fraudulent practices or to deal with the effects of a default on services contracts; protection of the privacy of individuals in relation to the processing and dissemination of personal data and; safety, (d) to ensure imposition or collection of direct taxes and (e) to avoid double taxation.<sup>561</sup> However,

---

<sup>559</sup> Joshua P Meltzer, "Cybersecurity, Digital Trade, and Data Flows: Re-Thinking a Role for International Trade Rules" Global Economy & Development WP 132, (2020) <<https://www.brookings.edu/articles/cybersecurity-digital-trade-and-data-flows-re-thinking-role-for-international-trade-rules/>> accessed 29 January 2025.

<sup>560</sup> *ibid* n 555

<sup>561</sup> General agreement of Tariffs and Trade 1947, Art XIV.

these measures must not be applied arbitrarily or discriminate between members unjustifiably or be disguised trade barriers to trade in services. To qualify under the general exception it needs to be ensured that the measure adopted is designed to achieve the objectives mentioned above. There must be a causal connection between the measure and the objective. Once it is established that the measure is taken to achieve one of the objectives of the Article, the Panel is required to take a “weigh and balance” approach to determine the importance of the objective, its contribution in achieving the objective, its overall impact on restriction of trade and the availability of alternative less trade restrictive measures.<sup>562</sup> Once this is established, it is necessary to determine whether the chapeau of good faith is satisfied.<sup>563</sup> This entails two queries: Whether the measure is applied in an arbitrary manner or is unjustifiably discriminatory and is it a disguised restriction on trade.

In the context of cyber security measures, member states would most likely submit that the measure is necessary to protect public morals/maintaining public order or is necessary to secure compliance with laws or regulations relating to prevention of deceptive and fraudulent practices. Against this backdrop, it has to be demonstrated that the cybersecurity measure is “necessary” to achieve the objective.<sup>564</sup> In doing so, the panel needs to conduct the process of weighing and balancing the relevant factors, contribution of the measure in achieving the end pursued and its impact on international commerce.<sup>565</sup> Evidence of data localisation measures or restrictions on cross border data flows improving cyber security can be of relevance here.<sup>566</sup> However, if such measures are not reducing cybersecurity threats<sup>567</sup> or are instead driven by the motivation to disproportionately benefit the domestic industries, they can't be deemed as “necessary”.<sup>568</sup>

---

<sup>562</sup> Andrew D. Mitchell and Glyn Ayres, “General and Security Exceptions Under the GATT and the GATS” [2011] INTERNATIONAL TRADE LAW AND WTO.

<sup>563</sup> United States — Import Prohibition of Certain Shrimp and Shrimp Products [1998] DS58.

<sup>564</sup> WTO, Brazil — *Measures Affecting Imports of Retreaded Tyres* [2007] DS332.

<sup>565</sup> WTO, Brazil — *Measures Affecting Imports of Retreaded Tyres* [2007] DS332.

<sup>566</sup> Meltzer JP, “Cybersecurity, Digital Trade, and Data Flows: Re-Thinking a Role for International Trade Rules” *Brookings* (November 13, 2019).

<sup>567</sup> *ibid.*

<sup>568</sup> WTO, Brazil — *Measures Affecting Imports of Retreaded Tyres* [2007] DS332.

### 3. TRIPS and Non-Personal Data

The relationship between the TRIPS Agreement and non-personal data is a complex and evolving subject within the realm of intellectual property law. TRIPS, established by the WTO, primarily focuses on protection of intellectual property rights including patents, copyright and trademarks. The TRIPS Agreement was established to reduce international trade barriers, promote effective and adequate protection of intellectual property and to ensure that the protection of intellectual property does not itself become a trade barrier.<sup>569</sup> TRIPS mainly has threefold objectives. Firstly, it lays down the standards and principles for the scope of trade related intellectual property rights. Secondly, it provides means for effective enforcement of trade related intellectual property rights while. Thirdly, it provides for effective procedures for settlement of disputes between different governments.<sup>570</sup>

The advent of the digital age has introduced complexities in the traditional intellectual property framework that are difficult to address. Managing and protecting data raises issues as it struggles to fit into existing IP categories. The TRIPS Agreement sets minimum standards of protection for IP. However, it does not provide for or explicitly address the digital economy or data.

#### 3.1. Copyright and Big Data

Article 2 of the Berne Convention enumerates protection for collection of literary and artistic works. It states “*collections of literary or artistic works such as encyclopedias and anthologies which, by reason of the selection and arrangement of their contents, constitute intellectual creations shall be protected as such, without prejudice to the copyright in each of the works forming part of such collections.*” It must be noted that when this provision was set in place, electronic databases or big data was not a phenomenon and these “collections” referred to anthologies and encyclopedias. The rights in collections had two layers; first is the right in each individual entry in the collection and second, the right in the collected work that was accorded to it due to selection or arrangement of individual works. The emergence of electronic databases in the 1990s and its recognition can be traced in Article 10 of the TRIPS Agreement that protect (1)

---

<sup>569</sup> The Agreement on Trade-Related Aspects of Intellectual Property Rights 1995

<sup>570</sup> Peter K. Yu , “Data Exclusivities and the Limits to Trips Harmonization ” (2019) 46 Florida State University Law Review 641

computer programs, whether in source or object code<sup>571</sup> and (2) compilations of data or other material, which by reason of their selection and arrangement of their contents constitute intellectual creations, shall be protected. However, this protection does not extend to the data itself.<sup>572</sup>

An understanding of Article 10.1 of the TRIPS Agreement, indicates that human generated softwares used for collecting, processing and analyzing big data can be given copyright protection. At the same time, an issue surfaces in the context of Article 10 as it imposes the condition of originality. A reading of Article 10 of the TRIPS Agreement suggests that some level of protection can be granted to databases that are original in their selection or arrangement.<sup>573</sup> However, this protection cannot extend to individual pieces of raw data that form part of a database. In furtherance of this, some authors argue that copyright protection for big data extends only to processed data compiled in a database.<sup>574</sup>

### **3.2. Mandatory Data Sharing vis-a-vis Article 13 of the TRIPS Agreement**

It is important to consider measures like mandatory data sharing within the boundaries of the TRIPS Agreement. Mandating sharing of original databases that can be protected under copyright law can raise concerns in relation to potential TRIPS violation. Article 13 of the TRIPS provides for circumstances in which member states can provide exception to exclusive rights granted under copyright protection. The provision aims to strike a balance between exclusive rights and public interests. Any limitations or exceptions provided by member states must ensure that the said limitations or exceptions comply with the three-step test. The three-step test was first enumerated in Article 9 of the Berne Convention and then again in Article 13 of the TRIPS Agreement. The test requires fulfillment of three separate tests for copyright limitations and exceptions i.e. a) be confined to certain special cases, b) do not conflict with normal exploitation of the work, and c) do not unreasonably prejudice the legitimate interests of the right holder.<sup>575</sup> Each of these are

---

<sup>571</sup> The Agreement on Trade-Related Aspects of Intellectual Property Rights 1995, a 10.1.

<sup>572</sup> The Agreement on Trade-Related Aspects of Intellectual Property Rights 1995, a 10.2.

<sup>573</sup> Daniel J. Gervais, "TRIPS Meets Big Data" [2021] SSRN Electronic Journal.

<sup>574</sup> Chandni Raina, *Protection Of Data Generated By E-Commerce Market Space As Intellectual Property* (Centre for WTO Studies, Policy Brief 2, 2020).

<sup>575</sup> The Agreement on Trade-Related Aspects of Intellectual Property Rights 1995, a 13.

separate requirements and each of these must be fulfilled separately. The WTO Panel on Section 110(s) of the US Copyright Act has elaborated on the three-step test. The same is discussed below particularly in relation to mandatory data sharing recommendations by the Expert Committee on Non-Personal Data Governance Framework.

The first requirement of “*be confined to certain special cases*” has been interpreted to mean that an exception must have a clearly and narrowly defined scope and reach. The exception thus needs to be limited or exceptional in its application and scope.<sup>576</sup> The WTO panel report provides that “the public policy purpose behind a limitation/exception can be useful in inferring the scope and reach of the limitations/exception”. However, specifically in India, the exception carved out by the Expert Committee for public good purposes is very widely worded. The grounds for mandatory sharing, as prescribed in the report i.e. sovereign, public good and business purposes, leaves a lot of scope open for interpretation. Therefore, due to lack of clear scope and boundaries of the grounds for mandatory sharing of NPD, the first part of the three-step test cannot be said to be fulfilled.

The second step (*do not conflict with normal exploitation of the work*) has also been elaborated in the WTO report. The term “exploitation” literally implies “making use of” or “utilizing for one’s own ends,” or the author’s right to extract economic value from their own work. The term “normal” has two different connotations, it could be of an empirical nature wherein it is referring to something “common” or it could also be referring to the normative type or standard. Both these connotations have been deemed relevant and an attempt has been made to harmoniously and effectively interpret both the connotations of the term “normal”. Normal exploitation of the work therefore implies that the right holders exercise all the exclusive rights conferred upon them. This step is crucial for new and emerging businesses like big data and big data analytics since they are not yet “*common*”.<sup>577</sup>

---

<sup>576</sup> WTO Report of the Panel WT/DS160/R of 15 June 2000 on United States - Section 110(s) of the US Copyright Act, See also Gervais DJ, “Exploring the Interfaces Between Big Data and Intellectual Property Law” (2019) 10 Journal of Intellectual Property.

<sup>577</sup> *ibid.*

While interpreting the third step i.e. “*not unreasonably prejudice the legitimate interests of the right holder*”, the WTO report explains the meaning of the terms “interests”, “legitimate”, “prejudice” and, “not unreasonable”. The general meaning of the term “interests” can be “a legal right or title to a property” or “a potential detriment or advantage” or even something that could be of some importance to a person. The term “legitimate” means “sanctioned or authorized by law or it can also refer to “something normal or regular”. The term “prejudice” means “damage, harm or injury”. And the term “unreasonably prejudice” indicates that “some level of prejudice is justifiable”. Considering that under this step some level of prejudice is not considered to prejudice the legitimate interests of the right holder, the Panel took a conservative approach and stated that legitimate interest could be looked at from the economic perspective and the economic value of rights conferred exclusively on the copyright holder. However, this is an incomplete opinion and legitimate interests are not necessarily always understood in economic terms. According to the Panel, “the prejudice to the legitimate interests of right holders reaches an unreasonable level if an exception or limitation causes or has the potential to cause an unreasonable loss of income to the copyright holder”.<sup>578</sup>

All the three components of the test have to be separately fulfilled by any limitations or exceptions imposed by the member states. The crucial question in context of non-personal data that falls within the scope of copyright protection and mandatory sharing is whether it qualifies the three-step test and does it amount to expropriation of data by the state that deprives data holders or investors from reaping the fruits of their investments. Mandatory sharing of data can affect a company's trade secrets and intellectual property.<sup>579</sup> Therefore, it is crucial for policymakers to determine the scope, duration and limitations of rights over data and avoid any unnecessary complications and incompatibilities with already existing international obligations. There is little possibility that the Berne Appendix on non-voluntary uses can in service of access to copyrighted materials for the big-data industries.

---

<sup>578</sup> “WT/DS160 - United States - Section 110(5) of US Copyright Act” (*Trade*) <[https://policy.trade.ec.europa.eu/enforcement-and-protection/dispute-settlement/wto-dispute-settlement/wto-disputes-cases-involving-eu/wtds160-united-states-section-1105-us-copyright-act\\_en](https://policy.trade.ec.europa.eu/enforcement-and-protection/dispute-settlement/wto-dispute-settlement/wto-disputes-cases-involving-eu/wtds160-united-states-section-1105-us-copyright-act_en)> accessed January 29, 2025.

<sup>579</sup> ICC (2023), Report and recommendations on the effective and efficient use of Requests for Information in competition investigation and studies. <https://iccwbo.org/newspublications/policies-reports/10-recommendations-to-make-requests-for-information-in-competition-investigation-more-efficient/>.

### 3.3. Trade Secret Protection and Big Data

If a database cannot be protected as a copyright due to lack of originality, it can still fall within the ambit of trade secret protection. Article 39 of the TRIPS Agreement provides protection for undisclosed information. It states that information that has been kept secret in the sense that it is not readily available to the public, has commercial value because it has been kept a secret and reasonable efforts have been made to keep the information a secret.<sup>580</sup> Secrecy is not an absolute criteria but a relative one and secrecy can be maintained even if the information is shared with third parties by the information holder. Whether a trade secret confers an IP right or protection akin to tort or unfair competition law is highly debated. Some scholars argue that it is a “quasi-IP right” that is a mix between liability and property rules.<sup>581</sup>

TRIPS and other legal instruments like the EU Trade Secret Directive refer to trade secrets as “information” which raises a concern that data might not be information. Legal scholars have suggested that a distinction between semantics (data as information conveying meaning) and syntactic (data as sequences of zeros and ones) must be drawn for the purposes of trade secret protection for big data.<sup>582</sup> They also suggest that only semantic level data should be protected and accordingly, unprocessed and unorganized data remains outside the scope of trade secret protection. Syntactic data does not constitute valuable information unless it is processed and analyzed. Therefore, some argue that unprocessed raw data is excluded from trade secret protection.<sup>583</sup> Trade secrets do not grant exclusive rights to the information holder. Rather, it merely grants protection against misappropriation of undisclosed information. It entitles the information holder to take action against the breaching party but cannot restrict them from accessing or making use of the information thus disclosed.<sup>584</sup>

---

<sup>580</sup> The Agreement on Trade-Related Aspects of Intellectual Property Rights 1995, a 39.

<sup>581</sup> Fia T, “Resisting IP Overexpansion: The Case of Trade Secret Protection of Non-Personal Data” (2022) 53 SSRN Electronic Journal 917.

<sup>582</sup> Determann and Drexl-No one owns data, resisting IP over expansion

<sup>583</sup> *ibid.*

<sup>584</sup> Wolfgang K, “A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis” (2016) 37 MAGKS Joint Discussion Paper Series in Economics.

#### 4. FTAs and Non-Personal Data Governance

Data governance and cybersecurity issues that have remained unresolved at a multilateral level have increasingly become a part of FTA negotiations. One of the first FTAs that banned data localisation requirements was the 2015 Japan-Mongolia FTA. The FTA specifies that neither party should require “to use or locate computing facilities in that area as a condition for conducting its business.”<sup>585</sup> A similar ban was also imposed under the TPP/CPTPP. The CPTPP has a chapter on E-commerce and it provides that the parties shall allow cross border transfer of information by electronic means, including personal information.<sup>586</sup> The CPTPP explicitly deals with cross border data flows and also provides exceptions to this obligation for achieving “legitimate public policy objectives”. Measures inconsistent with cross border flow obligation can be permitted in cases where the measure is to “achieve legitimate public policy objectives, is not applied in a manner of arbitrary and unjustifiable discrimination or a disguised trade restriction, and does not restrict the data transfer more than necessary.”<sup>587</sup> The CPTPP Framework also prohibits one of the key aspects of data localisation measures i.e. the requirement of domestic facility installation. It provides that the parties shall not require to use or locate computing facilities in their territory as a precondition to for conducting business.<sup>588</sup> The CPTPP also leaves room for exceptions wherein parties can defend data localisation requirements with the objective to achieve “a legitimate purpose of public policy”. A notable aspect of the CPTPP framework is that it expressly recognises data localisation as trade barrier issues<sup>589</sup> and its e-commerce chapter is not limited to service sectors and hence it would have a wider scope and application when it comes to data localisation measures.<sup>590</sup> However, only 12 out of 164 WTO members have been a part of CPTPP Framework and the GATS framework will continue to give baseline direction to other nations for addressing data localisation measures until legal clarity is provided to interpret WTO rules.

---

<sup>585</sup> Article 9.10 Japan- Mongolia FTA and, Creating Data Flow Rules through Preferential Trade Agreements Mira Burri

<sup>586</sup> Comprehensive and Progressive Agreement for Trans-Pacific Partnership 2018, a 14.11.2.

<sup>587</sup> Comprehensive and Progressive Agreement for Trans-Pacific Partnership 2018, a 14.11.3.

<sup>588</sup> Comprehensive and Progressive Agreement for Trans-Pacific Partnership 2018, a 14.13.

<sup>589</sup> Hodson S, “Applying WTO and FTA Disciplines to Data Localization Measures” (2018) 18 World Trade Review 579.

<sup>590</sup> Data Localization Measures and International Economic Law: How Do WTO and TPP/CPTPP Disciplines Apply to These Measures? ABE Yoshinori

Many FTAs now include chapters on IP, data flows and e-commerce which consequently can result in fragmentation in standards governing data and data flows internationally. The conflict between approaches of different countries can be traced not only in different FTAs but also in the national laws of countries participating in FTAs. For example, China has a rigorous data governance framework mandating data localisation and business to government data disclosure which can potentially be incompatible with its obligations under international investment agreements (IIA). China's domestic regulations aim to achieve data localisation and mandatory data sharing and on the contrary, China's international obligations push for free flow of data in the digital economy. The incompatibility between national laws and international obligations is the most prominent and profound in the case of China.<sup>591</sup> China's governance regime for data sharing and data flows have invoked criticisms from investors due to concerns surrounding confidentiality of data. Further, investors are also concerned if the implementation of the laws would be discriminatory against them.

The requirement of mandatory business to government data sharing has been criticized by foreign investors as a breach of fair and equitable treatment (FET) in the investment agreements if such requirements "lack transparency, are arbitrary, lack due process of enforcement, lack legal certainty or affect the investors' legitimate expectations".<sup>592</sup> It has also been argued that the practice of mandatory data sharing can be challenged for expropriation (direct or indirect). Expropriation is "the taking of foreign property by a state, whether for public purposes or other reasons". Direct expropriation is outrightly taking away an investment by the state and indirect expropriation is where the totality of the state's actions results in taking of an asset which deprives the investor of the ability to benefit from the investment. Post World-War II, minimum standards for lawfully taking foreign property developed. These standards included taking away the property for public purposes, in a non-discriminatory manner, with due process of law and accompanied by adequate compensations.<sup>593</sup>

---

<sup>591</sup> Bian C, "Data as Assets in Foreign Direct Investment: Is China's National Data Governance Compatible with Its International Investment Agreements?" (2022) 13 *Asian Journal of International Law* 342.

<sup>592</sup> *Ibid.*

<sup>593</sup> Baetens F, "Expropriation in International Investment Law," *International Law* (Oxford University Press 2017) <<https://doi.org/10.1093/obo/9780199796953-0159>> accessed January 29, 2025.

In *Einarsson v Canada*, indirect expropriation claims and breach of FET claims were made in respect of mandatory sharing of data for the first time in the international investment law regime. The case involves data owned by geophysical services incorporated (GSI) which was established under the Canadian laws. GSI licensed its data to third parties and obliged them to maintain confidentiality of the data shared. Canada's regulations required the sharing of the GSI data related to seismic operations. In compliance with Canadian regulations on environmental safety, GSI shared the data with the expectation that the information will be kept confidential. However, the data was allegedly shared by Canada with third parties without GSI's information or consent. The dispute is an ongoing dispute and the award is awaited. The final award will have an impact on the understanding of data as assets and would help in clarifying the interface between data and investment.<sup>594</sup> It has been argued that a claim for expropriation lies in cases where disclosure of data is mandatory will depend on the nature of the disclosure.<sup>595</sup> If the disclosure of data impacts the investor's ability to benefit from the investment, then a claim for expropriation would be successful. Data driven businesses rely heavily on their data for profits and disclosure of data to third parties without information or consent can substantially affect the business' capacity to earn profits from the data. However, in this case if the disclosure was made to another government entity (whose function is not similar to that of the disclosing entity), the investor's ability to derive profits from the data would remain unaffected and a viable claim for expropriation would not be made out. Therefore, mere disclosure of data is not enough for an expropriation claim and the nature of disclosure plays a crucial role in determining indirect expropriation.<sup>596</sup>

From the discussion above, it is evident that many new instruments are being used to govern data and data flows. FTAs, as compared to WTO norms, have more comprehensively tried to address data, data flows and barriers to data flows. However, FTAs can't replace an effective multilateral system governing the international standards for data and data flows. FTAs benefits and advantages at the best can do a patchwork in the context of international digital economy. Notably,

---

<sup>594</sup> Lachmann N, "Einarsson v Canada and Data as Asset in Investor-State Dispute Settlement" (*Kluwer Arbitration Blog*, October 9, 2023) <<https://arbitrationblog.kluwerarbitration.com/2023/10/09/einarsson-v-canada-and-data-as-asset-in-investor-state-dispute-settlement/>> accessed January 29, 2025.

<sup>595</sup> Tara Peramatukorn, "Potential Expropriation Claims Against Data Sharing Requirements — Guarini Global Law & Tech - NYU Law" (*Guarini Global Law & Tech - NYU Law*) <<https://www.guariniglobal.org/peramatukorn-expropriation>>.

<sup>596</sup> *ibid.*

the advantages of FTAs would be offset by the inconsistencies and overlaps it would create globally.<sup>597</sup> The former WTO Director General has observed that “*proliferation [of plurilateral trade agreements] is breeding concern about incoherence, confusion, exponential increase of costs for business, unpredictability and even unfairness in trade relations.*” Therefore, the value of a multilateral governance system cannot be undermined.<sup>598</sup>

## **5. India’s Data Localization Measures in Light of International Obligations**

Concerns surrounding exclusive control over big data have motivated certain countries to take initiatives to effectively regulate big data. The Ministry of Electronics and Information Technology published an expert committee report on Non-Personal Data Governance Framework. The expert committee report, while highlighting the role of big tech companies and the vast amounts of valuable data held by them, recommended mandatory sharing of raw non-personal data without any remuneration. In cases where value is added to private data through processing, data will be required to be shared on fair, reasonable and non-discriminatory (FRAND) terms. An entity can make a request for sharing of data for sovereign purposes, public interest/public good purposes and business purposes<sup>599</sup>. These recommendations were made by the committee with the objective to promote and foster innovation in the Country. However, some argue that it can stifle innovation instead of fostering it. More importantly, the approach of the expert committee report towards regulating non-personal data is being seen as a concern vis a vis India’s international obligations under TRIPS and other International instruments. While the concentration of big data in the hands of big tech companies or dominant market players is a rising concern, it is crucial to ensure that an effective balance is maintained and already existing international obligations are duly regarded.

The Report introduces a new category of data that is beneficial for the community called the High Value Data (HVD) that should be treated like a public good. According to the report, HVD can be beneficial for creating new jobs, helping in research and education, alleviating poverty, healthcare,

---

<sup>597</sup> Mira Burri, “The Governance of Data and Data Flows in Trade Agreements: The Pitfalls of Legal Adaptation” (2017) 51 UC Davis Law Review 65.

<sup>598</sup> Pascal Lamy, Dir.-Gen., World Trade Org., Opening Remarks at the Conference on “Multilateralizing Regionalism” in Geneva (Sept. 10, 2007), [http://www.wto.org/english/newse/sppl/sppl67\\_e.htm](http://www.wto.org/english/newse/sppl/sppl67_e.htm) [<https://perma.cc/6CTS-2XNV>].

<sup>599</sup> MeitY, Report by the Committee of Experts on Non-Personal Data Governance Framework (24(4), 2019).

urban planning, environmental planning, etc. Further, the report has suggested that in order to create checks and balances on the creation and sharing of HVD, data localisation measures should be imposed. Since a lot of non-personal data is derived from anonymised personal data, the chances that such data can be anonymised and misused are high. Therefore any non-personal data that once was personal data would inherit the sensitivity of personal data and must comply with the storage requirements specified in the PDP Bill (now DPDP Act). The Digital Personal Data Protection (DPDP) Act was passed in 2023. The DPDP Act in section 16 specifies that the government can restrict the transfer of personal data for processing outside the territory of India.<sup>600</sup>

Arguments for data localisation are majorly divided into three categories i.e. civil liberties, government functions and economic reasons. The expert committee on non-personal data framework has relied on the reason that sensitive data and super-sensitive data should not leave India.<sup>601</sup> However, a reading of the entire report suggests that the expert committee's primary focus is on the economic and strategic value of data with privacy being more of an ancillary concern. Even the DPDP Act's data localisation requirements have attracted criticism on similar lines that data localisation does not have "tangential relationship to data privacy" and has "become a proxy for debates on issues such as data sovereignty, something that, again, is not directly related to the issue of data privacy".<sup>602</sup>

India's data localization measures were motivated to defend sovereignty, ensure timely enforcement of laws, facilitate access to data and promote local enterprises against big competitors. These measures attracted criticisms from many including the US and the EU. Both India and China were warned by the US for creating restrictions on digital trade flows through their data localisation practices and for violating privacy and intellectual property protection.<sup>603</sup> The US has

---

<sup>600</sup> Digital Data Protection Act 2023, s 16.

<sup>601</sup> Bhalla K, "Data Sovereignty Cannot Be Compromised: Ravi Shankar Prasad" *Inc42 Media* (September 23, 2019) <<https://inc42.com/buzz/countrys-data-sovereignty-cannot-be-compromised-ravi-shankar-prasad-on-data-localisation/>> accessed January 29, 2025.

<sup>602</sup> Anirudh Burman, "Understanding India's New Data Protection Law" *Carnegie Endowment for International Peace* (October 3, 2023) <<https://carnegieendowment.org/research/2023/10/understanding-indias-new-data-protection-law?lang=en>> accessed January 29, 2025.

<sup>603</sup> Press Trust of India, "India Counters Donald Trump On Digitisation, Calls Data 'New Form Of Wealth'" *NDTV* (June 28, 2019) <<https://www.ndtv.com/india-news/india-counters-donald-trump-on-digitisation-calls-data-new-form-of-wealth-2060832>> accessed January 29, 2025.

strictly advocated against data localisation requirements as such measures in the long term can affect economic growth and competition and may not be consistent with WTO rules.<sup>604</sup> Concerns have also been raised regarding the expert committee's disregard for the role cross border flows of data has played and will continue to play in the digital economy.

With India's current trajectory to regulate personal and non-personal data, it is crucial to ensure that the new regulatory framework is coherent and consistent with the existing international legal framework. Data localisation and mandatory data sharing requirements being introduced domestically across various jurisdictions have the potential to violate existing WTO principles. There is a need for a multilateral level framework or international standards for data sharing to ensure that there is no fragmentation and non-uniformity in laws and agreements governing data and data flows in the digital economy.

## **6. Conclusion**

The complexities of data governance in the digital economy pose several challenges globally. Measures for cyber security like data localisation and mandatory sharing are often justified by goals concerning national security and privacy. However, these measures can conflict with international trade obligations under GATS, GATT and TRIPS Agreement. Several attempts have been made to address these issues but it has created a fragmented regulatory landscape that complicates cross border trade and data flows. Although these measures may pass under GATS exceptions if interpreted broadly, these exceptions are not self-judging and call for a careful assessment to prevent misuse as disguised protectionism.

The inconsistencies between international obligations and domestic data governance laws underscore the need for a global harmonized framework. While some FTAs attempt to address data issues, the need for a more comprehensive framework cannot be undermined. International standards that provide for cybersecurity measures, IP protection and promoting fair competition without creating trade barriers are essential. Additionally, it is essential to address the

---

<sup>604</sup> "Localization Barriers to Trade" (*United States Trade Representative*) <<https://ustr.gov/trade-topics/localization-barriers>> accessed January 29, 2025.

fragmentation of laws at the domestic level and their potential to conflict with international trade obligations.

Given the rapid evolution of the digital economy and the increasing importance of data as a resource, there is a pressing need for greater international cooperation to harmonize data governance rules for uniformity. A multilateral approach for clear and consistent rules for data governance are crucial to ensure that unnecessary restrictions to trade are not mounted. This will require international dialogue to develop international standards that can align domestic regulations with international obligations. Therefore, domestic data governance frameworks need to be crafted carefully to comply with international obligations so that both national and global objectives can be balanced.

## CONCLUSION

With the revolution in Artificial Intelligence (AI) and its need to feed on data for producing outcomes based on deep learning, the exploration of big data and its regulatory landscape underscores a critical intersection of technology, law, and economics in the digital age. As big data continues to be a catalyst for innovation, its regulation poses unique challenges and opportunities for policymakers, businesses, and society at large.

Big data's transformative potential lies in its ability to drive innovation, improve decision-making, and create new business models. Its volume, velocity, and variety enable organizations to derive valuable insights and gain a competitive edge over their rivals. In the AI and digital economy, big data is often termed as the “new oil” due to its ability to drive innovation. However, the analogy does not go beyond metaphors as it often overlooks the significant differences between data and oil. The same attributes that make big data valuable also introduce complexities in its regulation. Due to the benefits and advantages big data can offer, its access has become a debatable issue. The challenge is to strike a balance between fostering innovation and implementing robust regulatory frameworks that protect data and information privacy, ensure fair competition, and safeguard public interest. However, there are other definitional challenges involved in the data, particularly when there is convergence between personal and non-personal data as explored in chapter one.

Since data creation, collection, processing and storage entails substantial opportunity costs, firms tend to fiercely protect it from their competitors. This is done primarily through four ways. Intellectual property (IP) law plays a pivotal role in protecting big data. The EU's Database Directive, with its dual protection of original and non-original databases, exemplifies how IP law can incentivize data creation and sharing. The sui generis database right, in particular, aims to protect substantial investments made in obtaining, verifying or presenting the contents of the database, thereby encouraging innovation. The EU also introduced a novel Data Producer's right on raw and machine generated data with the objective to improve access to data and to promote data sharing. However, owing to certain challenges and massive criticism, the thought remained a mere unimplemented experiment. Intellectual property protection for data raises complex issues that call for a balance between creating a property right in data and ensuring optimal allocation of data along with information dissemination for the greater social and economic good.

Datasets can also be protected through common law remedies for protection of undisclosed information in the form of breach of confidence for misappropriation of undisclosed information.

However, with regard to trade secret protection for big data, an issue arises whether protection can be extended to individual datum. The limitation of these protections to datasets with commercial value raises questions about the scope and effectiveness of existing legal frameworks. As data becomes increasingly central to business strategies, there is a need to reassess and possibly expand the legal definitions and protections for data. Contract law is also a fundamental tool for governing data transactions, with non-disclosure and confidentiality agreements being standard practices. However, contractual protection has its pitfalls, including issues related to data ownership, party autonomy, and unequal bargaining power. These challenges necessitate careful consideration of the terms and conditions in data sharing agreements to prevent unfair practices and ensure equitable access to data.

It is critical to note that although copyright law in India does not extend protection to raw data, nor does Indian law extend protection to non-original databases, it does not essentially mean that data is in the public domain or that there can be free-riding of such data. Data that is in de-facto possession of the firm and measures to protect confidential information/trade secrets, including measures in IT laws can be used to protect data against misappropriation by providing both a legal layer against unfair competition and a technology layer against hacking and data theft. Contract law offers another legal layer for securing effective and efficient transactions involving data. However, as seen from the failed attempts in the E.U. creation of a sui generis right for protecting non-personal data, is unnecessary as the current legal protection seems adequate.

In furtherance of this, standard rules for regulating data sharing contracts that can help in providing foundation for effective and fair contractual sharing of data can be a great tool for instilling confidence in data sharing without the risks of data breaches. Technological measures are equally crucial in protecting data from cyber threats. Cybersecurity attacks such as hacking, phishing, and malware can compromise the integrity and value of data, leading to significant economic and reputational damage that may also come with hidden and indirect costs for data holders. In India, the Information Technology Act, 2000, and the IT Rules provide a comprehensive legal framework for data protection, emphasizing the importance of reporting cyber incidents promptly so that the harms of data breach incidents can be minimized at the earliest. Once data is leaked or hacked it loses its value. Thus, appropriate remedial measures in cases of data breach can create deterrence and strengthen cybersecurity measures and promote best practices that are essential for safeguarding big data.

In the digital economy, the benefits of big data are increasing exponentially, and it can drive innovation and provide a competitive edge in the market. Amidst this digital transformation of the market, concerns regarding the conduct of big data firms emerge in the context of competition law. The concept of data as an essential facility highlights its strategic importance in the digital economy. Firms possessing big data can make predictions from the past and create self-reinforcing “feedback loops” wherein the more data they collect, the more accurately they can predict the market. Therefore, firms that control large datasets can create barriers to entry, potentially leading to anti-competitive practices. While the Competition Commission of India has yet to invoke the essential facilities doctrine in the context of data, other jurisdictions have recognized its relevance. The *Nielsen Arbitron* case in the United States and the *Google/DoubleClick* merger case in the European Union illustrate how control over critical data can influence market dynamics. However, effective caution must be exercised in treating data as an essential facility under competition law since market wide effects of harm to competition must prevail over the needs of individual companies who need access to the data. Such effects-based tests will go a long way in balancing the requirements of significant opportunity costs incurred by data producers and the incentives required to generate such data to feed into the growing AI and big data industry.

In this context, the debate on whether to adopt ex-ante or ex post regulatory mechanisms to effectively deal with complex problems related to data and data sharing is gaining traction. Ex ante regulations, which proactively address potential issues, may be necessary to complement ex post mechanisms, which react to existing problems only when there is a market-failure. A hybrid approach that combines both strategies could provide a more comprehensive framework for managing the complexities of big data in the digital economy when such market failure is shown. Current literature does not show industry wide market failure in access to data and that voluntary mechanisms involving data transaction are always preferable over mandatory data sharing requirement. Of course, measures involving transparency in such transactions, including how similarly situated parties may be treated remains a major question where regulation may step-in. Several jurisdictions including Germany, EU, Japan, China, etc. have enacted regulations on data and data sharing. India's approach to non-personal data regulation although represents a forward-thinking strategy to promote data sharing while ensuring economic and public benefits, it needs to promote voluntary sharing at its core with minimal interventions with high triggers. Since data is in the physical possession of the firm and without effective data localisation, a remedy in the form

of data sharing may remain ineffective. However, data localisation in itself remains controversial since there are arguments that this may effectively prevent data companies from engaging in significant investments in India due to higher entry barriers and costs in the form of excessive and fragmentation in legal and regulatory requirements in different jurisdictions due to multiple data centres, reduced data security, higher entry for small firms and its effect on competition, less incentives to innovate if corresponding remedies of data sharing are executed due to data localisation, risk of government surveillance and non-transparent methods involved in data sharing, or at worse, the risk of data theft.

The Ministry of Electronics and Information Technology (MeitY) formed an expert committee on Non-Personal Data Governance Framework in India. The committee explored the regulatory aspects of Non-Personal Data and emphasized the importance of data sharing for public interest and economic purposes. Among other suggestions, the committee also proposed to create a community right over data with the aim to democratize the benefits of data processing, ensuring that they are not confined to data holders alone. However, the report faced criticism for definitional ambiguities and its relationship to competition law. Critics argue that the focus should be on preventing the abuse of dominance rather than targeting dominance itself. This distinction is crucial for developing a fair and effective regulatory framework that promotes competition without stifling innovation. It remains to be seen if a hard-edged regulatory approach will be preferred in light of some data sharing norms already in place since 2012 through the National Data Sharing and Accessibility Policy (NDSAP) and also within the framework of the DPDP Act, 2023 and 2024 SEBI circular to share anonymized data for research purposes in light of the 2022 SEBI circular on “Approach to securities market data access and terms of usage of data provided by data sources in Indian securities market”. In the light of piece-meal data sharing requirements, it is doubtful if an overarching regulation for mandatory data sharing, particularly affecting private players, will be legislated.

Finally, Internationally, the regulatory landscape for big data is evolving, with various jurisdictions experimenting with different approaches. The EU, Japan, China, and Germany have enacted regulations that address data and data sharing, reflecting diverse legal and cultural contexts. These global perspectives offer valuable lessons for India as it navigates its regulatory aspects. While other jurisdictions have not yet invoked hard-edged data sharing regulations, the Expert Committee in India in its 2021 report has specifically recommended that data sharing must be allowed for

public interest purposes and for economic purposes to provide a level playing field or for monetary consideration or in certain cases (community data, raw data etc.) based on non-remuneration. In such a situation a cautious approach needs to be exercised in order to incentivize investment in Big-Data creation and its monetization. The regulatory dimension of requiring data-owning firms to share data based on public interest and market-failure considerations is complex. While such regulations can promote fairness and competition, they must be designed carefully to avoid disincentivizing investment in big data and the digital economy. A nuanced approach that balances public policy objectives with private interests is essential for fostering a vibrant digital economy. While data is being equated to oil, it's surprising that not much has been done for making a robust and effective data sharing framework. Navigating the regulatory landscape of big data is a multifaceted challenge that requires a delicate balance between innovation, competition, and public interest. This also necessarily calls for careful consideration of India's international obligations and commitment to free trade and under WTO principles. The interplay between intellectual property law, contractual and technological measures, and competition law highlights the complexity of regulating big data. India's proactive stance on non-personal data regulation offers a promising path forward, but it must address criticisms and refine its approach to ensure effectiveness. However, it is doubtful that any overarching global framework for data-sharing will shape up through international multilateral agreements. The regulatory framework for big data must evolve to keep pace with technological advancements and market dynamics. There is a need for a governance system that can facilitate data re-use, data repurposing, data sharing along with data protection in order to encourage holistic development. Policymakers, businesses, and stakeholders need to collaborate to create a balanced and forward-thinking regulatory environment that promotes innovative incentives, ensures fair competition, and maximizes public benefits. An outcome-based approach must be eschewed in favor of nuanced regulation, when and only needed where there is established market-failure. By doing so, the full potential of big data to drive economic growth through economic efficiency, enhance societal well-being, and shape a sustainable digital future can be ensured.