ERGING NATA CAVERNANCE





existing and emerging technology and cyberspace, their implications for the society, and their regulation. Additionally, CCG organises an annual International Summer School in collaboration with the Hans Bredow Institute and the Faculty of Law at the University of Hamburg in collaboration with the UNESCO Chair on Freedom of Communication at the University of Hamburg, Institute for Technology and Society of Rio de Janeiro (ITS Rio) and the Global Network of Internet and Society Research on contemporary issues of information law and policy.

Table of Contents

I.	Introduction: Emerging Trends in Data Governance Swati Punia, Jhalak M. Kakkar, and Shashank Mohan	I
2.	An Analysis of India's New Data Empowerment Architecture Smriti Parsheera	7
3.	Group Data Rights in Law and Policy Arindrajit Basu and Amber Sinha	24
4.	Unpacking Community Data: Agency, Rights and Regulation Kritika Bhardwaj and Siddharth Peter de Souza	4 I
5.	For What it's Worth: Realising the Value of Data Mansi Kedia and Gangesh Varma	57
6.	Data Stewardship: Re-imagining data governance Astha Kapoor	73
7.	Making data count - A case for developing data stewardship models for the Indian judiciary Ameen Jauhar	84
8.	Emotion Recognition and the Limits of Data Protection Vidushi Marda	94
9.	Searching for a Room of One's Own in Cyberspace: Datafication and the Global Feministation of Privacy Anja Kovacs	III
	Notes on Contributors	125

three concepts: community, community data, and community data governance. They identify the gaps in current policy design, examine definitional challenges, and propose a taxonomy based on identity and interests. While framing the argument on broadening the horizons of the concept of 'community', they propose to look beyond the economic standpoint of extracting commercial value to account for social, cultural, and political identity and interests. They argue that communities could be formed on the basis of needs, work, or social causes and recommend that members of a community should have the agency to understand what and where the data is used, how it is used, and how it impacts them. With regard to governing such entities, they recommend acknowledging the fluid nature of communities, their heterogeneity in terms of how communities are formed, how they operate and disappear, and the ways members of the community associate with data related to them. Incorporating these factors while designing governance frameworks would ensure that the capacity of communities to organise and thrive is not constrained by needless formalisation.

Regarding the design of data governance policies, Mansi Kedia and Gangesh Varma in 'For What it's Worth: Realising the Value of Data' emphasise the need to incorporate a comprehensive understanding of the different types of values of data. They highlight different dimensions of data that assign different kinds of value to data, namely economic, social and technical. For this purpose, the urge to move the focus beyond the contours of personal data to governing non-personal data as well. They argue for a model of data governance that harnesses the potential of data without causing harm and minimising abuse and misuse. They recommend adopting a multi-stakeholder approach to designing a balanced data governance regime that considers the varying objectives and interests of stakeholders through transparent and participative processes. The essay concludes by emphasising the need to build better institutions and reliable processes that help assess the value of data in a holistic manner for developing effective data governance frameworks. Mansi and Gangesh caution against promoting one set of values alone and disregarding other drivers of the value of data as it would breed inequality and inefficiency in the ecosystem.

Astha Kapoor echoes the need for adopting an alternative approach to data governance in her essay 'Data Stewardship: Re-imagining Data Governance'. She highlights the insufficiency of the notice and consent regime to protect the privacy of individuals and also argues that existing data protection frameworks do not account for harms that arise at a community level. Astha proposes that the data stewardship model could help address some of these challenges. She discusses the concept of data stewardship to help unlock the value of data in a manner that leads to individuals' empowerment through their active and direct participation in the data economy. Consent in this regard is operationalised through 'data stewards', a class of independent intermediaries obligated to act in the interest of the data generators. She discusses some of the most popular models of data stewardship across the globe

- data cooperatives, data commons, personal data stores, and data trusts - that are attempting to enhance the decision-making powers of individuals and communities to enable them to reap economic benefits from their data. Acknowledging the challenges of actualising this bottom-up approach, she emphasises the need to record and learn from ground-level evidence and deploy sandboxes to support evolution and innovation in this space.

Ameen Jauhar explores the potential of the data stewardship model within the judicial system. In his essay, 'Making Data Count - A Case for Developing Data Stewardship Models for the Indian Judiciary', he focuses on operationalising the idea of data trusts for non-personal data within the judicial system to achieve the twin objective of accumulating digital intelligence and judicial data sharing for social innovation. Amongst stewardship models such as data exchanges, data cooperatives and data trusts, he finds data trusts as the most suitable stewardship model in the context of the Indian judiciary. He recommends that such a data trust must be independent in nature, should be a non-profit entity, and must incorporate an institutional and technological layer. Moreover, this entity should also have a framework for discharging fiduciary obligations as well as engagement protocols and mechanisms.

Vidushi Marda in 'Emotion Recognition and the Limits of Data Protection' captures the evolution and nature of emotion recognition technology. She explains that facial expressions are not solely related to emotional states and they have multiple causes and meanings. She draws from other jurisdictions like the EU to make a case for banning this form of technology for its inherent invasive and discriminatory nature and direct impact on human rights and freedoms including dignity, autonomy, privacy, etc. Vidushi explains why data protection frameworks are inadequate to regulate such technologies and the need to employ a slew of regulatory tools and levers to check the use of emotion recognition technology and its impact on society. Vidushi examines in detail why the data protection laws proposed in India as well as in the UK, US, Brazil, and China are inadequate in mitigating the harms arising from emotional recognition technology due to the wide exemptions given to the State. She attributes the growing appetite for its adoption to proliferation of surveillance technology for purposes of public security, national security, and public order across the globe and the trend of affording exemptions in the areas and relationships that require protection is not unique to India. She makes a thoughtprovoking case for why data protection frameworks are not the site for dealing with the dangers of surveillance, oppression, marginalization and criminalization of communities. According to her, the data protection frameworks are primarily concerned with efficient and safe data processing, instead of challenging the growth and ubiquity of surveillance. Therefore, she recommends that the best way to deal with emotion regulation technologies is to reject the design, development, testing, and deployment of such systems.

To do this, we developed three concepts in this essay: community, community data and community data governance. In our understanding, while these concepts are necessarily interwoven and connected, they still require to be distinguished to be able to explore how to think of community data from a regulatory perspective.

Our approach to thinking about community has involved examining how to incorporate a fluid idea of community which may emerge based on identity and interest; as well as community data, which may restrict the agency of the community that may also be generated through algorithmic classifications. In doing so, we are interested in demonstrating how such a concept can bring representation for new groups, coalitions, and alliances as a by-product of participating in a digital economy, for instance worker unions of platform workers – in addition to acknowledge existing coalitions. We have argued for thinking of group rights, while remaining mindful of hierarchies within groups and the rights of individuals within a group. Finally, we have identified the interests or outcomes that any framework for community data must - at the very least - aim to secure and the possible regulatory frameworks for articulating or securing such outcomes.

For What it's Worth: Realising the Value of Data

Mansi Kedia and <mark>Gangesh Varma</mark>t

INTRODUCTION

Economic resources, human or other, utilised to produce output were referred to by early economists as factors of production. From the classic set of four - land, labour, capital, and entrepreneurship, the factors were expanded to include other natural resources, raw materials and also information. Information has been defined by some as data affecting behaviour. The role of information (processed data) as a distinct factor of production was recognised several decades ago and was tested using guiding principles of factor markets - factor prices, sources of supply, sources of demand and ability to produce additional products. The economic value of data is thus apparent. However, the imagination and measurement of its scale and scope have been completely transformed by the digital economy. While we may have witnessed several socio-economic transitions in the past, the current digitalisation is an unimaginable scale of civilisational transformation, with data at its centre.

Data is the "raw material produced by abstracting the world into categories, measures and other representational forms, such as – numbers, characters, symbols, images, sounds, electromagnetic waves, bits, etc.". The exponential growth of data and its processing has led to ubiquitous use cases. Companies collect and use data for innovation, process efficiency, marketing, and customization of services. Governments rely on data to improve governance, quality and reach of public services that it renders. These applications cut across a range of data types - demographic data, personal data, spatial data, machine data, etc. However, the production and utilisation of data is not new. The history of such utility ranges from scientific innovations to state policy, notwithstanding the time and cost invested in generating, analysing, and interpreting it. It wouldn't be an exaggeration to say

I Mansi is Senior Fellow at the Indian Council for Research on International Economic Relations. She can be reached at mkedia@icrier.res.in. Gangesh is a Senior Associate at Saraf and Partners focusing on technology and policy. The author can be reached at gangeshvarma@gmail.com.

² Sunday Okpighe, 'The Seven Factors of Production' (2015) 5 British Journal of Applied Science & Technology 217 https://journalcjast.com/index.php/CJAST/article/view/510 accessed 12 November 2022.

³ A. M. McDonough, Information Economics and Management Systems (McGraw Hill 1963).

⁴ When data is processed, organized, structured, or presented in a given context so as to make it useful, it is called information.

⁵ Andrew Berczi, 'Information As A Factor Of Production' (1981) 16 Business Economics 14 http://www.jstor.org/stable/23482505 accessed 12 November 2022.

⁶ Rob Kitchin, The Data Revolution (SAGE Publishing 2014).

that good-quality data was a scarce resource, and therefore treated as a valuable commodity, which was either carefully shielded or traded.⁷

With the advent of the internet of things (IoT), the sources and types of data have grown exponentially. In a world of hyper digital connectivity, numerous tools, and devices, ranging from watches to washing machines are equipped to become a source of data. This paradigmatic shift is reflected in how data is produced, collected, stored, and utilised. From being scarce and limited in access, today we have multiples of quintillion (*I followed by 18 zeros*) bytes of data being generated every day. It is less costly, high frequency, targeted and much more accurate. It is also relatively open and accessible. As Rob Kitchn stated in *The Data Revolution*, "A data revolution is underway, one that is already reshaping how knowledge is produced, business is conducted, and governance is enacted."

The enormous volume of data is aptly captured by the nomenclature 'Big Data' which requires greater analytical and processing capabilities in contrast to the earlier treatment of data. Consequently, its applications are much more complex and demanding. However, big data is not the only component of the data revolution. Related initiatives include digitisation, linking together, and scaling-up of traditionally produced datasets (small data) into networked data infrastructures, and developing new indicators, targets or open datasets that has directly fed into the development discourse around the world.9

The nature and applications of data have led to the creation of institutions that attempt to standardize guidelines and policies with respect to data formats, sharing protocols and intellectual property rights regimes. The nature and function of these institutions are very diverse. For example, institutions like the Internet Engineering Task Force, founded in 1986, develop protocols and standards for the internet and that, to an extent, determines the treatment of data on the internet as a medium. The World Intellectual Property Organisation in 1997, adopted guidelines for the protection of databases. More recently, countries and regions have created their own regulatory frameworks such as the APEC Privacy Framework, the European Union's General Data Protection Regulation, and the ASEAN Framework on Personal Data Protection. Prima facie, most current institutions and regulations consider tradeoffs and prioritize one of the many aspects of data governance. The introduction to this essay might suggest that the value of data is only manifested in the economic outcomes it catalyses, a representation of monetary benefit or the importance of data in reducing costs, improving efficiency, scaling up and other factors relating to

economic growth." In our assessment, as reflected by others" this is a narrow view of the value of data and we propose that a comprehensive framework that looks at its social and technical aspects be included in the assessment of the value of data. A paper by Gunther et al (2017) recommends an integrated model for realising the value of data for individual organizations, though social well-being is measured only in terms of education, health, public safety, and security." In this essay, we propose to strengthen this idea and offer a framework for a comprehensive measurement of the value of data that includes its economic, social and technical aspects. The future of the digital economy will be better understood if we are able to measure it.

The following sections of the essay will outline the proposed framework, explaining different kinds of value drivers and their implications on data governance. It also provides an example to illustrate the integration of value drivers and an implementation model for data governance.

I. A FRAMEWORK FOR THE VALUE OF DATA AND IMPLICATIONS ON DATA GOVERNANCE

The understanding of the value of data has evolved over time, adapting to its ever-increasing applications. The current discourse on data governance mainly adopts two types of value frameworks, one that focuses on pure profit (economic value generation) and the other on human rights.¹³ Those highlighting economic profit focus on the unfettered use of data for improving the quality and reliability of business processes to maximise monetary gains, while champions of human rights emphasize on the harms of data abuse or misuse, prioritising user rights and opposing its ungoverned utilisation. However, in our opinion, this binary approach to assessing the value of data can result in a rather polarizing set of perspectives on data governance and policy development.¹⁴

⁷ ibid.

⁸ ibid.

Independent Expert Advisory Group on Data Revolution for Sustainable Development, 'A World That Counts: Mobilising the Data Revolution for Sustainable Development' (2014) https://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf accessed 12 November 2022; World Bank, 'The World Bank Group Supports the Data Revolution for Sustainable Development' https://www.worldbank.org/content/dam/Worldbank/Statcap/HDRSD/WBG-support_data_revolution.pdf accessed 14 November 2022.

¹⁰ Andrew McAfee and Erik Brynjolfsson, 'Big Data: The Management Revolution' [2012] Harvard Business Review https://hbr.org/2012/10/big-data-the-management-revolution accessed 12 November 2022.

II James Wilson and others, 'The Value of Data: Applying a Public Value Model to the English National Health Service' (2020) 22 Journal of Medical Internet Research e15816 https://www.jmir.org/2020/3/e15816/ accessed 14 November 2022; Jaap Wieringa and others, 'Data Analytics in a Privacy-Concerned World' (2021) 122 Journal of Business Research 915 https://linkinghub.elsevier.com/retrieve/pii/S0148296319303078 accessed 14 November 2022.

Wendy Arianne Günther and others, 'Debating Big Data: A Literature Review on Realizing Value from Big Data' (2017) 26
The Journal of Strategic Information Systems 191 https://linkinghub.elsevier.com/retrieve/pii/S0963868717302615 accessed 14 November 2022.

¹³ Bill Schmarzo and Dr. Mouwafac Sidaoui, 'Applying Economic Concepts To Big Data To Determine The Financial Value Of The Organization's Data And Analytics, And Understanding The Ramifications On The Organizations' Financial Statements And IT Operations And Business Strategies' (2017) https://www.dell.com/wp-uploads/2017/03/USF-The-Economics-of-Data-and-Analytics-Finalz.pdf; HM Treasury, 'The Economic Value of Data: Discussion Paper' (2018) <a href="https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/731349/20180730_HMT_Discussion_Paper_-_The_Economic_Value_of_Data.pdf.

¹⁴ Ivana Kottasová, 'These Companies Are Getting Killed by GDPR' (CNNMoney, 11 May 2018) https://money.cnn.com/2018/05/II/technology/gdpr-tech-companies-losers/index.html accessed 14 November 2022 (several companies have had to shut down operations due to GDPR); Rebecca Janßen and others, 'GDPR and the Lost Generation of Innovative Apps' (National Bureau of Economic Research, May 2022) https://www.nber.org/papers/wyoo28 accessed 14 November 2022 (These findings which sate that the GDPR has has an adverse impact on innovation should not be overstated); Joseph Jerome, 'The GDPR's Impact on Innovation Should Not Be Overstated' (*Center for Democracy and Technology*) https://ct.org/insights/the-gdprs-impact-on-innovation-should-not-be-overstated/ accessed 12 November 2022; Nicholas Martin and others, 'How Data Protection Regulation Affects Startup Innovation' (2019) 21 Information Systems Frontiers 1307 https://ct.org/insights/the-gdprs-impact-on-innovation Accessed 14 November 2022. (For a mixed perspective where research results show both the stimulation and stifling of innovation).

Our purpose here is to illustrate the need for a comprehensive value framework that helps harness the potential of data without causing harm and minimising abuse. This idea is not new but needs reiteration and mainstreaming for policymakers to recognise its importance in designing data governance policies. We examine conceptions of both data and value.

While the definition of data was introduced earlier in this essay, its varied conceptions have not been explored. There are several metaphors and analogies used to describe data and make sense of it. Comparisons range from oil¹⁵ to sunshine,¹⁶ avocado¹⁷ to nuclear power,¹⁸ but these rarely offer consolation to those trying to make sense of the role and value of data. While there may be several more analogies that may be developed to explain the exploitation of data, each one becomes inadequate in a fresh context. However, these diverse conceptions reflect on its extensive utility and value drivers. From the economics point of view, the value of data can be paralleled to its conceptualisation as capital,¹⁹ infrastructure,²⁰ currency,²¹ asset²² or generally as an economic good. Arguing that it is non-rivalrous and non-excludable, data is predominantly considered as a public good that should be made accessible and used to deliver real value to society.²³

Greater utilisation of data and demands for regulating its use led to the emergence of privacy guidelines that offered multiple taxonomies for data. With digitalisation there is a fundamental change in the data itself. Data that originates from observations today are less obvious to the individual and are a product of processing itself.²⁴ More recently, regulators have delineated data into personal data and non-

personal data. This was arguably a middle-path that protected individuals as well as facilitated innovation especially since it demarcated specific types of data and did not create a blanket restriction on the use of all types of data. Probably, the potential utility of non-personal data for developing both better products and better policies increased in its perceived value, and thereby brought more regulatory attention to the category of non-personal data. This scrutiny also highlights that there exists a certain degree of fluidity between these types of data where the anonmyisation of personal data could convert it to non-personal data whereby the restrictions could be diluted without threatening an individual's privacy. However, with the evolution of technology and data use, we are forced to question its efficacy.25 Researchers have shown how they were able to de-anonymise data and render the protection provided by anonymisation pointless. Even with anonymisation the risk to privacy persists.²⁶ Regulations on data sharing infrastructure and data exchanges respond to the various types of data such as geo-spatial, financial, health, industrial, and demographic that are broadly categorised into personal and non-personal data. Examining different value dimensions of data is essential to illustrate the limitations of using an either-or approach.27

We believe that value from data can be classified into three main categories - social, economic, and technical. These can manifest in various permutations and combinations and could also result in overlaps. While notoriously difficult to measure, the concepts of economic and social value have often been discussed, ²⁸ whereas technical value requires familiarization.

I.I. Economic Value of Data

As discussed above, data metaphors are often used to capture the economic value of data speaking of some part of its character, but rarely do they ever provide a complete picture. Data has distinctive traits such as its reusability. This also means that the often used "data is the new oil" is a poor metaphor. Its reusability and inexhaustible character renders it very different as a resource from that of a fossil fuel. Data can be reused endlessly and in domains like

¹⁵ Charles Arthur, 'Tech Giants May Be Huge, but Nothing Matches Big Data' *The Guardian* (23 August 2013) https://www.theguardian.com/technology/2013/aug/23/tech-giants-data accessed 14 November 2022.

Shona Ghosh and Jake Kanter, 'Google Says Data Is More like Sunlight than Oil, Just 1 Day after Being Fined \$57 Million over Its Privacy and Consent Practices' (Business Insider, 22 January 2019) https://www.businessinsider.in/google-says-data-is-more-like-sunlight-than-oil-just-1-day-after-being-fined-57-million-over-its-privacy-and-consent-practices/articleshow/67640224. cms> accessed 12 November 2022.

¹⁷ Dr. Deborah Elms, 'Data Is the New Avocado?' (Asian Trade Centre, 9 April 2019) https://asiantradecentre.org/talkingtrade//data-is-the-new-avocado accessed 12 November 2022.

¹⁸ James Bridle, 'Opinion: Data Isn't the New Oil — It's the New Nuclear Power' (ideas.ted.com, 17 July 2018) https://ideas.ted.com/opinion-data-isnt-the-new-oil-its-the-new-nuclear-power/ accessed 12 November 2022.

¹⁹ Jathan Sadowski, 'When Data Is Capital: Datafication, Accumulation, and Extraction' (2019) 6 Big Data & Society 205395171882054 http://journals.sagepub.com/doi/10.1177/2053951718820549> accessed 14 November 2022.

²⁰ Peter Kawalek and Ali Bayat, 'Data as Infrastructure' https://nic.org.uk/app/uploads//Data-As-Infrastructure.pdf accessed 12 November 2022.

²¹ Knowledge at Wharton Staff, 'Data as Currency: What Value Are You Getting?' (Knowledge at Wharton) https://knowledge.wharton.upenn.edu/article/barrett-data-as-currency/ accessed 14 November 2022; William D Eggers, Rob Hamill and Abed Ali, 'Data as the New Currency: Government's Role in Facilitating the Exchange' [2013] Delloit Review <a href="https://www2.deloitte.com/content/dam/insights/us/articles/data-as-the-new-currency/DRJ_data_as_the_new_currency_pdfs accessed 12 November 2022; Guillaume Desjardins, 'Your Personal Data Is the Currency of the Digital Age' (The Conversation) https://theconversation.com/your-personal-data-is-the-currency-of-the-digital-age-146386 accessed 14 November 2022.

²² Peter Lake and Paul Crowther, 'Data, an Organisational Asset' in Peter Lake and Paul Crowther, Concise Guide to Databases (Springer London 2013) http://link.springer.com/10.1007/978-1-4471-5601-7_1 accessed 12 November 2022.

²³ 'The Data Centered Economy: A New Temple for a New India' (Indian Council for Research on International Economic Relations ²⁰¹⁹) http://icrier.org/pdf/The-Data-Centred-Economy-A-New-Temple-for-a-New-India.pdf accessed 12 November ²⁰²².

²⁴ Martin Abrams, 'The Origins of Personal Data and Its Implications for Governance' [2014] SSRN Electronic Journal http://www.ssrn.com/abstract=2510927 accessed 14 November 2022.

²⁵ Luc Rocher, Julien M Hendrickx and Yves-Alexandre de Montjoye, 'Estimating the Success of Re-Identifications in Incomplete Datasets Using Generative Models' (2019) 10 Nature Communications 3069 http://www.nature.com/articles/s41467-019-10933-3 accessed 14 November 2022.

²⁶ Alex Hern, "Anonymised" Data Can Never Be Totally Anonymous, Says Study' *The Guardian* (23 July 2019) https://www.theguardian.com/technology/2019/jul/23/anonymised-data-never-be-anonymous-enough-study-finds accessed 14 November 2022; Michele Finck and Frank Pallas, 'They Who Must Not Be Identified—Distinguishing Personal from Non-Personal Data under the GDPR' 10 International Data Privacy Law II.

²⁷ In some ways, the convergence of the regulatory scope on personal and non-personal data by the Joint Parliamentary Committee in its review of the Indian Data Protection Bill 2019 and then later the withdrawal of the bill by the government to reconsider the approach reflects the importance of carrying out a multi-dimensional value assessment at the pre-liminary stages of policy development.

²⁸ Federico Cabitza, Angela Locoro and Carlo Batini, 'Making Open Data More Personal Through a Social Value Perspective: A Methodological Approach' (2020) 22 Information Systems Frontiers 131 http://link.springer.com/10.1007/s10796-018-9854-7 accessed 14 November 2022; Günther and others (n 11); 'World Development Report 2021: Data for Better Lives' (The World Bank Group 2021) https://www.worldbank.org/en/publication/wdr2021; HM Treasury (n 12).

artificial intelligence and machine learning models, its value grows greater with scale. Data creates positive and negative externalities. These often cross-over to socio economic consequences. An individual's car journey when used with digital maps, saves them from traffic congestion, is an example of a positive externality. On the other hand, the use of an individual's personal data to exclude and discriminate against communities, especially those who did not share their data, is an example of a negative externality.

Mariana Mazzucato, professor in the economics of innovation and public value at University College London, thinks that profits in the digital technology era have become confused with value. She draws parallels to critiques of GDP as a misleading indicator when looked at through different lenses and asks important questions - What are the market participants doing? This is similar to the broader macroeconomic literature that challenges the use of GDP as a measure of welfare. Several empirical studies found that mean welfare stagnated and even deteriorated in many developed countries despite steady rise in GDP.²⁹ The parallel in the digital ecosystem could be an exponential increase in scale of digital adoption, accompanied by exploitative business models that profit a tiny group of organisations, while flawed algorithms and poor governance raise serious ethical concerns of consumer harm.³⁰ This highlights the limitation of using a single lens to measure the value of data.

The debate on the economic identity of data is also one that remains unsettled. As mentioned earlier, one of the various conceptions of data is as capital. Many businesses, especially in the digital ecosystem consider data as their single biggest asset. Proponents of this view regard data as a form of capital just like financial capital, given its ability to generate new products and services. Data also creates new data. For those treating data as labour, the argument is centered around data as a measure to reduce inequality. Data intensive companies do not compensate consumers for their data, a resource that is monetised and exploited by businesses for financial gain. The economic treatment of data varies with differing types of data and the purposes they serve. A cookie cutter approach in determining the economic value of data, without considering the context in which data is being collected and used, would be unfair and inefficient.

1.2. Social Value of Data

Social value of data arises from its ability to be utilised by and for communities or society at-large. The key distinction being that its use is not restricted to the economic considerations and benefits to an individual or singular entities in society. Open data sets are a great example of the potential social value of data beyond the economic gains it offers. Social value has been interpreted as the ability of data to be used for the benefit of and by the sources of data (citizens themselves).33 This understanding is built around the public good argument and the non-rivalrous and non-excludable characteristics of data use. For instance, data on the health status of individuals in the pandemic or crowd sourced traffic updates create positive spill overs benefitting societies as a whole. This has featured in India's policy discourse at multiple stages, but most prominently as the discussion around 'community data' in the Report of the Expert Committee on Non-Personal Data which may be perceived as a call for 'distributive justice in a digital economy'.34 The existence of informational externalities and the non-rival character of data immediately imply that private markets uses and market prices (if they exist), will not deliver social value. What's more, the value of any given data set is also fundamentally determined by the value of the uses to which it can be put, which are likely unknown until after the fact.35 Yet for public controllers of data concerned to maximise social welfare, methods based on realised financial values in market transactions are insufficient. Social value could potentially be gained from more data collection, wider access, or the scope to join information from different data sets with varying types of data records (noting also the need to manage the negative externalities of potential privacy loss and security breaches).36 In some sense social value of data is an aggregation of the positive and negative externalities, or the net social welfare.³⁷ A non-integrated approach can lead to sub-optimal outcomes as illustrated by Coyle & Diepeveen (2021) using the application of geospatial data and the transport sector in the UK.38

1.3. Technical Value of Data

According to MIT's recent Tech Review Insights, valuing data means understanding who participated in its creation. Data's value is also a product of the input and participation of digital users with complex consent protocols,

²⁹ Jeroen Van den Bergh and Antal Miklós, Evaluating Alternatives to GDP as Measures of Social Welfare/Progress (Vienna: WWWforEurope 2014).

³⁰ MIT Technology Review Insights, 'Fair Value? Fixing the Data Economy' (MIT Technology Review) https://www.technologyreview.com/2020/12/03/1012797/fair-value-fixing-the-data-economy/ accessed 13 November 2022.

³¹ MIT Technology Review Custom in partnership with Oracle, 'The Rise of Data Capital' (2016) https://files.technologyreview.com/whitepapers/MIT_Oracle+Report-The_Rise_of_Data_Capital.pdf accessed 13 November 2022.

³² Wendy C.Y. Li, Makoto Nirei, and Kazufumi Yamana, 'Value of Data: There's No Such Thing As A Free Lunch in the Digital Economy' https://www.oecd.org/site/stipatents/programme/ipsdm-2018-5-2-li-nirei-yamana.pdf accessed 13 November 2022.

³³ Cabitza, Locoro and Batini (n 27).

³⁴ Aniruddh Nigam, '[Vidhispeaks] Exploring Community Data Rights over Non-Personal Data' (30 October 2021) https://www.barandbench.com/columns/policy-columns/vidhispeaks-exploring-community-data-rights-over-non-personal-data-accessed 13 November 2022.

³⁵ Diane Coyle and Annabel Manley, 'Working Paper - Potential Social Value from Data: An Application of Discrete Choice Analysis' (Bennett Institute for Public Policy, University of Cambridge, UK 2021) https://www.bennettinstitute.cam.ac.uk/publications/social-value-data/ accessed 12 November 2022.

³⁶ Diane Coyle and Stephanie Diepeveen, 'Creating and Governing Social Value from Data' [2021] SSRN Electronic Journal https://www.ssrn.com/abstract=3973034 accessed 13 November 2022.

³⁷ ibid.

³⁸ ibid.

from granting permissions to platforms to access their data, to labelling and digitization of work conducted during processes like reCAPTCHA, etc.39 We conceptualise technical value as the efficiency gains from the architecture of technology using which data is collected, processed, and stored. With dramatic increases in the volume of data structuring, data centre networks, and interconnecting architectures become critical to allow for the efficient use of data. Data management and data centre networks must choose from competing priorities routing efficiency, high capacity, low power consumption, flexibility, etc. Technical value is also grounded in the context that technology is not always agnostic to social or economic values and is capable of regulation by virtue of its architecture and design. Often referred to through the concept that 'code is law', 40 it highlights that value choices are made in the development of technology and its capabilities. Thus, technical value is high where the product or technology in question enables further innovation and behaves as a medium for further growth. This may be referred to as the 'generative' nature of the technology41 that plays a role in its ability to become a general-purpose technology much like electricity, or the internet as we see it today.42

There are of course other concepts of value such as cultural, moral, political, geo-political, etc. However, the categorisation of social, economic, and technical subsumes the broadest range of issues and would suffice to deliberate on developing balanced regulations. While these categories are not water-tight and do influence each other, in policy discourse they present distinct objectives which are often lost when one gets prioritized over another. Some recent data governance regulations or proposals such as those restricting cross-border data flows,⁴³ emphasise the socio-political value of data and the geopolitical risks arising out of its misuse when located outside territorial jurisdictions. However, ICRIER's study on the economic implications of data flows found techno-economic reasons driving the choice of location for data storage.⁴⁴ The importance of data flows is amplified with the growth of data value chains spanning across organizations and countries. This has even prompted thinking along the lines of "global data value chains" and their implications.⁴⁵ The increase

in operational costs, compliance burden, the risk to privacy through potential surveillance of localised data etc. are consequences that are side-lined because of the prioritization of geo-political factors over techno-economic factors. On the other hand, where antitrust regulations for data centered businesses are more focused on economic value creation, social objectives of privacy violations and harm may become second-order priorities. An optimal data governance regime will promote collective value - an outcome of a country's economic, social, cultural, and political context. Necessarily, optimal data governance regimes will differ for countries.

Similarly, the selective assessment of values and binary choice can be seen in the debate between privacy and national security. For instance, in the context of encryption, tough adversarial views can result in an either-or type of policy options. ⁴⁶ The friction between value perspectives was also visible through the ongoing Covid-19 pandemic. Contact tracing tools confronted stakeholders with a dilemma between the protection of individual privacy and protection against the pandemic ⁴⁷ While several solutions emerged across the world, all of them dealt with tough trade-offs.

2. WHAT SHOULD DATA GOVERNANCE IN INDIA LOOK LIKE?

Given that data governance has implications for diverse sections and sectors of society, and priorities that may range from economic growth, social development, and national security, it is not uncommon to see the presence of different ministries and departments⁴⁸ that try to regulate data with various competing objectives. Some of these regulations or policies are not specifically focused on data governance but impact the treatment of data significantly. For example, the e-commerce policy proposed by the Department for Promotion of Investment and Internal Trade while developing policy for the e-commerce sector had several proposals relating to data and its utility. It attempted to address the alleged unfair advantage that big-tech players had over small and upcoming players.⁴⁹ Proposals designed to facilitate

MIT Technology Review Insights (n 29).

Lawrence Lessig, 'Code Is Law' [2000] Harvard Magazine https://www.harvardmagazine.com/2000/01/code-is-law-html accessed 13 November 2022.

Jonathan Zittrain, 'Law and TechnologyThe End of the Generative Internet' (2009) 52 Communications of the ACM 18 https://dl.acm.org/doi/10.1145/1435417.1435426 accessed 14 November 2022.

⁴² George RG Clarke, Christine Zhenwei Qiang and Lixin Colin Xu, 'The Internet as a General-Purpose Technology: Firm-Level Evidence from around the World' (2015) 135 Economics Letters 24 https://linkinghub.elsevier.com/retrieve/pii/S0165176515002773 accessed 14 November 2022.

⁴³ Reserve Bank of India, 'Storage of Payment System Data - RBI/2017-18/153' https://www.rbi.org.in/scripts/NotificationUser.aspx?Id=11244 accessed 14 November 2022.

⁴⁴ Rajat Kathuria and others, 'Economic Implications of Cross-Border Data Flows' (Indian Council for Research on International Economic Relations 2019) http://hdl.handle.net/11540/11375 accessed 13 November 2022.

⁴⁵ Jeremmy Okonjo, 'Legal Constitution of Global Value Chains in the Digital Economy' (Afronomicslaw.org) https://www.afronomicslaw.org/2020/11/11/legal-constitution-of-global-value-chains-in-the-digital-economy accessed 12 November 2022.

⁴⁶ Yashovardhan Azad, 'Data Bill: The Security vs Privacy Debate' *Hindustan Times* (23 January 2021) https://www.dnaindia.com/analysis/data-bill-the-security-vs-privacy-debate-101611406252249.html accessed 14 November 2022; Dipshikha Ghosh, 'The Big National Security versus Privacy Debate' *DNA India* (28 March 2017) https://www.dnaindia.com/analysis/column-the-big-national-security-versus-privacy-debate-2371072 accessed 13 November 2022.

⁴⁷ Bernard Marr, 'Why Contact Tracing Apps Will Be The Biggest Test Yet Of Data Privacy Versus Public Safety' Forbes
https://www.forbes.com/sites/bernardmarr/2020/06/01/why-contact-tracing-apps-will-be-the-biggest-test-yet-of-data-privacy-versus-public-safety/ accessed 13 November 2022; Ashkan Soltani Bergstrom Ryan Calo, and Carl, 'Contact-Tracing Apps Are Not a Solution to the COVID-19 Crisis' (Brookings, 27 April 2020) https://www.brookings.edu/techstream/inaccurate-and-insecure-why-contact-tracing-apps-could-be-a-disaster/; Pierfrancesco Lapolla and Regent Lee, 'Privacy versus Safety in Contact-Tracing Apps for Coronavirus Disease 2019' (2020) 6 DIGITAL HEALTH 205520762094167 https://www.brookings.edu/techstream/inaccurate-and-insecure-why-contact-tracing-apps-could-be-a-disaster/; Pierfrancesco Lapolla and Regent Lee, 'Privacy versus Safety in Contact-Tracing Apps for Coronavirus Disease 2019' (2020) 6 DIGITAL HEALTH 205520762094167 https://www.brookings.edu/techstream/inaccurate-and-insecure-why-contact-tracing-apps-could-be-a-disaster/; Pierfrancesco Lapolla and Regent Lee, 'Privacy versus Safety in Contact-Tracing Apps for Coronavirus Disease 2019' (2020) 6 DIGITAL HEALTH 205520762094167 <a href="https://www.brookings.edu/techstream/inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccurate-and-inaccura

⁴⁸ For example, the Ministry of Electronics and Information Technology (MEITY) with the proposed data protection bill, the recently passed intermediary rules etc. also has numerous departments focusing on different areas, Ministry of Communications through its telecommunications regime, the Reserve Bank of India through its regulations for financial data, The Department of Consumer Affairs with proposed amendments for consumer protection on e-commerce platforms.

⁴⁹ Ministry of Commerce and Industry, 'Draft National E-Commerce Policy: India's Data for India's Development' <a href="https://doi.org/10.237/bb/-10.2

different objectives of regulation are therefore laden with specific value judgments. The emergence of new regulatory instruments such as - mandating privacy by design, data minimization shows the prioritization of privacy rights over economic efficiency, or instruments such as data portability and interoperability, reflect the prioritization of market competition and economic opportunity. Many of these are still in proposal stages in India, while in other jurisdictions like Europe, are already in force through the GDPR and allied regulations.

India's digital journey and its existing heterogeneities, support the need for a comprehensive data governance regime that must accommodate varying and sometimes competing objectives. With an enormous population that is connected online issues such as privacy, security, consumer protection become important priorities. Yet, with an equally large population that is offline or unable to use the internet - issues of access, digital inclusion, and digital literacy become equally important demands on policy. The diversity of objectives can result in a disparate approach to policy making that will have sub-optimal outcomes. Similarly, divergent positions of stakeholders or stakeholder groups such as the interests of small business as against large establishments, foreign and local businesses, private and public enterprises also complicate policy formulation. While several parts of the government machinery have responded to these diverse needs, the related regulations developed in India are neither coordinated nor comprehensive. In the last year alone, India has seen attempts by various sectoral regulators or different ministries that have introduced regulations either with a very narrow view such as RBI's new guidelines for recurring payments through credit cards or broad-based regulations that may be counterproductive. 50 For example, e-commerce rules under the Consumer Protection Act that tried to accommodate both competition concerns and consumer protection objectives, instead created potential contradictions with other regulations.⁵¹ While the eagerness to address problems is commendable, the lack of patience seen through ad-hoc unpredictable policy development processes and a lack of a larger vision to plan and implement strategically is disappointing. A simple case in point being the recently implemented intermediary guidelines⁵² which has created a chimera of the Ministry of Information and Broadcasting and the Ministry of Electronics and Information Technology for regulating online

content. The guidelines attempt to regulate not only social media intermediaries but also digital news media in the same stroke. This trend of knee-jerk regulations implemented so far is not only complex and overlaps in the jurisdiction but can result in competing or conflicting objectives and outcomes. The most confounding aspect in this quagmire is that all these developments occur while the country awaits a primary data protection legislation. The proposed personal data protection bill has been deliberated for years, languished between parliamentary committees, and tabled in the parliament, only to be withdrawn by the government. At the time of writing this essay, the Indian government has stated that it's reworking the data protection bill along with a slew of legal reforms for "contemporary and future challenges and catalyse Prime Minster Narendra Modi's vision of India Techade." While India has been very active in its regulatory responses, the fundamental approach is still unclear.

For better inspiration, we may look to Singapore that showcased an excellent example of balance of values, nuanced regulatory clarity and cooperation between regulatory agents and other stakeholders. A well-synthesised and potentially optimal data governance regime is visible in Singapore's proposed data portability framework. A significant amendment in 2020 to the Singapore Personal Data Protection Act introduced a data portability requirement among others in its first set of significant changes to its legislative framework since 2012.55 This was jointly developed by the Personal Data Protection Commission and the Competition and Consumer Commission of Singapore. This collaboration is a reflection of a comprehensive approach that has laid down economic, social and technical requirements in the design of the regime. The discussion paper issued provides a thorough examination of different dimensions of introducing a data portability requirement ranging from competition implications, and data protection concerns. It captured the potential costs involved in compliance, and potential barriers to entry that such provisions may create. The discussion paper was followed by public consultation on the introduction of data portability requirement and other amendments to the PDPA of Singapore.⁵⁶ Following the nearly two-month comment period that closed in July

Ashwin Manikandan and Saloni Shukla, 'New Norms Put Auto Debits in No Man's Land' The Economic Times (21 October 2021) https://economictimes.indiatimes.com/tech/technology/new-norms-put-auto-debits-in-no-mans-land/articleshow/87169930.cms accessed 14 November 2022.

As a platform, e-commerce entities would come under the ambit of 'intermediary' as envisaged under the Intermediary Guidelines under the IT Act and imposing personal liability would conflict with safe harbour provisions under section 79. On issues of mis-selling and mis-leading advertisements, the proposed amendments tend to conflict with provisions of the parent legislation i.e., the Consumer Protection Act, 2019. Specifically, the parent legislation under s.21 (6) provides some defence for potentially mis-leading advertisements that may have been part of 'ordinary course of business'. Similarly, the inclusion of liability for innocent misrepresentation through the proposed amendments are not aligned with the objectives of protecting consumers against malicious and fraudulent misrepresentations by the parent act.

⁵² Ministry of Electronics and Information Technology, 'Notification Dated, the 25th February, 2021 G.S.R. 139(E): The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021' https://www.meity.gov.in/content/notification-dated-25th-february-2021-gsr-139e-information-technology-intermediary.

⁵³ Aashish Aryan, 'Explained: Social Media and Safe Harbour' *The Indian Express* (27 May 2021) accessed 13 November 2022; Malavika Raghavan, 'India's New Intermediary & Digital Media Rules: Expanding the Boundaries of Executive Power in Digital Regulation' (https://fpf.org/ accessed 14 November 2022; Internet Freedom Foundation, 'Explainer: Why India's New Rules for Social Media, News Sites Are Anti-Democratic, Unconstitutional' Scroll.in (25 February 2021) https://stroll.in/article/988105/explainer-how-indias-new-digital-media-rules-are-anti-democratic-and-unconstitutional-accessed 13 November 2022.

⁵⁴ BS Reporter & PTI, 'Govt Withdraws Data Protection Bill, 2021, Will Present New Legislation' *Business Standard* (3 August 2022) https://www.business-standard.com/article/economy-policy/centre-withdraws-personal-data-protection-bill-2019-to-present-new-bill-122080301226_L.html accessed 13 November 2022.

^{55 &#}x27;PDPC | Amendments to the Personal Data Protection Act and Spam Control Act Passed' https://www.pdpc.gov.sg/News-and-Events/Announcements/2020/II/Amendments-to-the-Personal-Data-Protection-Act-and-Spam-Control-Act-Passed-accessed 14 November 2022.

⁵⁶ Personal Data Protection Commission, Singapore, 'Public Consultation on Review of the Personal Data Protection Act 2012 -Proposed Data Portability and Data Innovation Provisions' (2019) accessed 14 November 2022.

2019, the PDPC published all responses received during the public consultation⁵⁷ and in January 2020 provided a detailed document which curated inputs and the point-by-point responses of the PDPC.⁵⁸ It took into account feedback received from stakeholders and communicated the intent to reduce the scope of data that would come under the scope of the portability provision while retaining the requirement itself. Furthermore, the amendments also excluded the exemptions previously available to private actors acting on behalf of government. A notable shift is also seen in the approach to consent – with greater clarity on consent requirements, new exceptions, and development of the concept of deemed consent. Furthermore, the process factored in economic considerations involved in facilitating data portability including different approaches for valuing data.⁵⁹ What we wish to highlight here is not specific provisions to mimic but a reliable process that assess diverse value conceptions and inspires balanced regulations. A predictable and transparent process also contributes to a sense of accountability of actors that are dealing with citizens' personal information.

Some may point to India's recently initiated Account Aggregator (AA) Network as a good example of balanced value realisation. The framework is considered to have managed to create a financial data sharing entity that facilitate seamless sharing of data for a consumer to avail financial services based on a strong consent requirement. The AA as an entity would not be able to create user profiles nor have access to the contents of the data they transfer and would be under RBI's strict regulations. Yet, there are strong criticisms that identify several concerns relating to ethics, consent friction and fatigue, lack of specificity in associated guidelines to prevent abuse, storage of data after revocation of consent etc., clearly highlighting that the choice of solutions has prioritized one set of values over another or failed to meaningfully address specific concerns.

These instances highlight how issues of regulation in the digital economy are no longer disjoint or completely separable. The overlap of jurisdictions has become inevitable. There is an evident need for integration for law making and regulatory responses. This is no easy task. A converged regulator brings many benefits but also poses questions of effective enforcement and efficiency. This becomes especially challenging in the context of India's institutional history that is rife with turf battles. Given a regulatory landscape that is dotted with several key stakeholders and their varying objectives, a multistakeholder approach to governance becomes necessary. A statutory warning that come with the support for this approach is that it is deeply detrimental in the absence of specific measures for inclusion of stakeholders, transparent and participative processes, and accountability of actors. Without these factors ensured, any multistakeholder approach would merely become a conduit for the will of the loudest or the most powerful voice in the ecosystem.

India's approach to multistakeholder participation in policymaking is mixed. Some institutions like the Telecom Regulatory Authority of India (TRAI) are by design consultative. Yet the lack of an institutionalized and predictable process of multistakeholder policymaking is clearly felt. The policy journey in the development of the data protection legislation for example began with the Justice Srikrishna Committee's consultative processes, however, upon moving to the stage of the bill being reviewed by a Joint Parliamentary Committee, it became invite-only with closed-door consultations. While there are arguments both for and against open or closed consultative processes, the trouble is a lack of consistency and predictability of the process in addition to the lack of accountable institutions that facilitate policy development. Further, this problem was only amplified by the subsequent withdrawal of the bill. In the absence of consistent, predictable, and transparent processes, stakeholders are left in a cliff-hanger of suspense relying on media reports quoting unnamed sources⁶² and statements from ministers⁶³ to guess both the unpredictable timelines and undecided scope of the future of data protection laws in India. In this context, it is important to be reminded that the benefits of an institutionalized, consistent, and predictable process include better transparency and accountability of stakeholders in the process. This can also improve trust not only among stakeholders involved in the regulatory ecosystem but also among general citizens and motivate wider civic engagement. It would facilitate a democratic and reliable process to identify and deliberate value(s) that need to be considered holistically in making policy choices. Eventually, the goal is not a utopian realization of all values or achieving an average in a comprehensive regulation. It is to provide the opportunity

⁵⁷ ibid.

Personal Data Protection Commission, Singapore, 'Response to Feedback on the Public Consultation on Proposed Data Portability and Data Innovation Provisions' (2020) accessed 14 November 2022.

⁵⁹ INFOCOMM Media Development Authority, 'Guide to Data Valuation for Data Sharing' https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Guide-to-Data-Valuation-for-Data-Sharing.pdf accessed 14 November 2022.

George Matthew, 'Account Aggregators: New Framework to Access, Share Financial Data' (*The Indian Express*, 6 September 2021) https://indianexpress.com/article/explained/account-aggregators-new-framework-to-access-share-financial-data-7490966/ accessed 14 November 2022; 'Account Aggregator, a Game Changer' (9 September 2021) https://www.thehindubusinessline.com/opinion/account-aggregator-a-game-changer/article36385799.ecce accessed 14 November 2022; ET Bureau, 'A New Digital Revolution in India' *The Economic Times* (8 September 2021) https://economictimes.indiatimes.com/opinion/et-editorial/a-new-digital-revolution-in-india/articleshow/85955574.cms accessed 14 November 2022; Rajat Deshpande, 'How Data Can Change Lending-experience in India'? '(cnbctv18.com, 22 December 2021) https://www.cnbctv18.com/finance/how-data-can-change-lending-experience-in-india-11902902.htm accessed 14 November 2022.

⁶¹ Rohan Jahagirdar and Praneeth Bodduluri, 'Digital Economy: India's Account Aggregator System Is Plagued by Privacy and Safety Issues' (2020) 55 Economic and Political Weekly https://www.epw.in/engage/article/digital-economy-indias-account-aggregator-system accessed 26 October 2022.

⁶² Aishwarya Paliwal, 'New Data Protection Bill to Be Stringent, Tech Giants in India Need to Strictly Follow Rules: Govt' *India Today* https://www.indiatoday.in/technology/features/story/new-data-protection-bill-to-be-stringent-tech-giants-in-indianeed-to-strictly-follow-rules-govt-1986127-2022-08-10> accessed 14 November 2022.

⁶³ PTI, 'Significant Work Done, Draft Digital India Act Framework by Early 2023: MoS IT' *The Hindu* (6 November 2022) -https://www.thehindu.com/business/Economy/significant-work-done-draft-digital-india-act-framework-by-early-2023-mos-it/article66103357.ece> accessed 14 November 2022.

for adequate representation and consideration of stakeholder interests through transparent deliberations. This would facilitate a review of different permutations and combinations of value frameworks and enable informed decision-making in a nuanced manner especially when it comes to trade-offs. Furthermore, developing regulations for the digital ecosystem has significant repercussions globally - for people, businesses, and the technology itself. Given India's quest for leadership in global affairs, it would need to improve its policymaking processes and institutions to factor not only national perspectives but build a collaborative approach with different jurisdictions and their value frameworks.

CONCLUSION

We reiterate the undercurrent of this essay that echoes several other scholars and continues to be attempted by various stakeholders - move away from the binaries in assessing the value of data while developing the data governance ecosystem. In India's pursuit of a balanced data governance regime so far, the dominant focus has been on the specific substance of regulation that often fail to accommodate diverse value drivers. Whereas the need of the hour is to invest in better institutions and reliable processes that can holistically approach the value(s) assessment of data. Championing one set of values while disregarding another is a recipe for unsustainable ecosystems that breed inequality and inefficiency. We are at an inflection point in history that asks several questions and raises the stakes so high that decisions today will determine the kind of future we build in a world where lines between digital and analog fade.

