The Data Centered Economy: A New Temple for a New India

Indian Council for Research on International Economic Relations

November 2019

DATA Catalyst

ICRIER

It is a capital mistake to theorize before one has data.

—Sir Arthur Conan Doyle, **The Sign of Four, A Scandal in Bohemia**
British mystery author & physician (1859 - 1930)

TABLE OF CONTENTS

List of Figures

List of Tables

1. INTRODUCTION

In the 1950s, India's first Prime Minister referred to scientific research institutes, steel plants, power plants and dams, being built in India after independence, as "*Temples of Modern India*". These projects were integral to India's development goals and combined the strengths of international cooperation. Without the latter, it is implausible that the projects would have taken off. Were India's first Prime Minister alive today, he might well have been persuaded, like we are, of using the same metaphor for data.

Indeed, data has been declared as the world's most valuable resource, beating oil in the process.[1] Like oil, it must first be extracted, refined and then delivered in a usable form to realise its full value. Data has also been variously designated as an input, an asset class, suggesting it can be traded, stored, and protected in a vault. Another description likens data to steel of the digital economy, a metaphor that simultaneously conjures up images of monopolistic practices by dominant firms in the digital space just like steel firms in their brick and mortar space.

The oil analogy has been the one that has stuck in public consciousness for at least two reasons. One, it is catchy, simple, pithy and expressive. Two, while oil generated, and continues to generate enormous economic value, it also simultaneously creates vast negative spillovers that need to be checked and internalised for true economic valuation of the oil economy.

There is however one important difference that needs all the emphasis it can get. The data centered economy does not have the luxury of time that oil had after its discovery. The industrial revolutions driven by oil created so much wealth and prosperity that very few stopped to worry about air pollution and climate change and the damage that it was causing to economic systems. Only now, when it has become a crisis, is there conscious climate action at global and local levels. On the other hand, the data economy, has been exposed much more quickly to its negative spillovers, almost simultaneous to its accelerated use by governments and businesses. It is not possible to postpone the ugly consequences of data abuse, misuse, leakages and fraud to another day. The Breach Level Index that reports real time statistics on data breaches, confirms that close to 15 billion records have been lost or stolen since 2013.[2] The average cost of every compromised record is estimated to be US$ 148.[3] Thus nations are devising standards against negative data spillovers. This is apparent in the introduction of new laws and regulations that provide consumers greater control over their personal data. 2018 was particularly significant for privacy and data protection laws. The European Union implemented the most comprehensive of such regulations called the General Data Protection Regulation (GDPR). GDPR introduced in May 2018, is a landmark legislation on how businesses ought to approach consumers' privacy with regards to data. Compliance with GDPR requires companies to define data protection policies, conduct data protection impact assessments and document their data processing mechanisms. The Brazilian General Data Protection Law of 2018 was also largely aligned to the GDPR. The California Consumer Privacy Act, to take effect in 2020, is also the first of its kind US law. It applies to businesses that collect information on Californian residents and places restrictions on the collection, use and disclosure of personal information. The Asia Pacific Economic Cooperation's (Apec) data privacy framework works on a set of privacy rules that facilitates sharing of data across borders within the region.

Some policies on data regulation are also driven by the economic interests of nations. For instance, the debate on data regulation in India conceptualises data as a sovereign asset, asserting its exclusive right to use data generated by the citizens for their welfare and development. Indonesia's policy is also driven on similar arguments. However, their Communications & IT Minister, Rudiantara, recently stated that the policy of localising data storage and use, may need to be considered in the event of its adverse impact on digital economic growth.[4] The modern view on data privacy

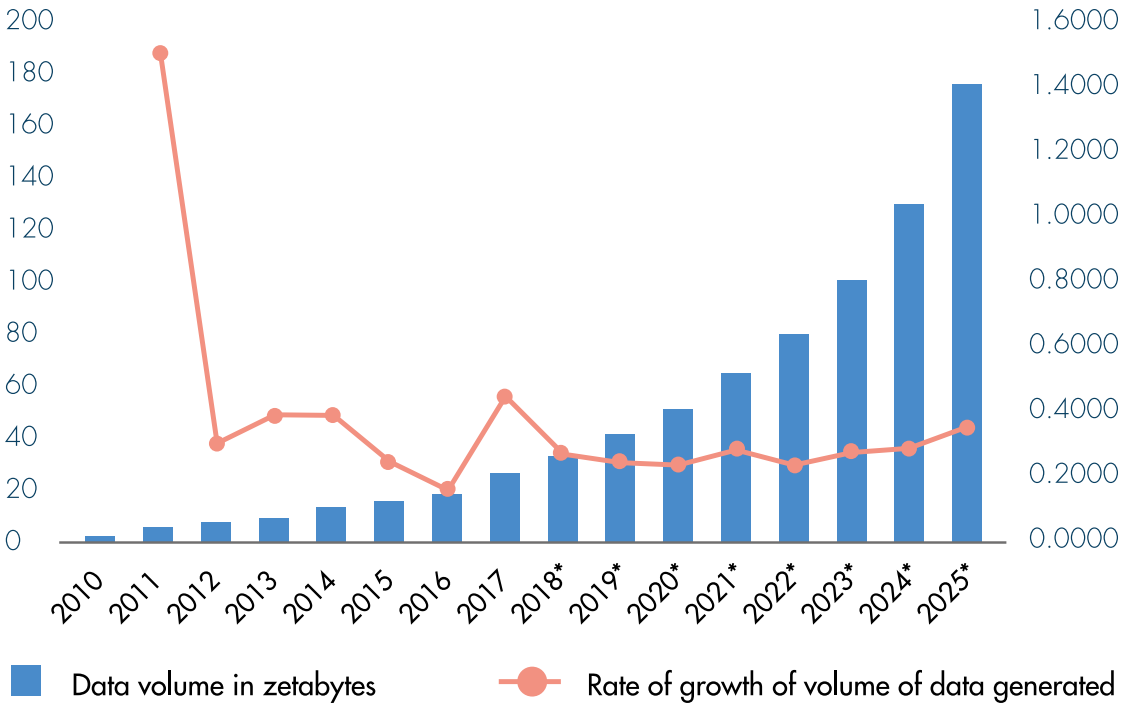
regulations is largely based on the understanding that data belongs to individuals and not to companies/ institutions that buy or use them.

There is no doubt that data needs protection and every state has the right to protect its citizens and its data. We are of the opinion that an optimal data protection regime will promote its value and what is optimal will vary by country and circumstance. In other words, it will be informed by the economic, social and cultural milieu. An overly protective, discriminatory and intrusive regime risks undermining current value and compromising future innovations.

The Indian economy is at the doorstep of a '1 trillion-dollar opportunity' anchored around its burgeoning digital economy.[5] The trillion-dollar digital opportunity is inspired by a growing market for a host of digital services, platforms, applications, content and solutions. India's digital universe is rapidly expanding. In March 2019, India's Internet user base exceeded 500 million and is set to reach 627 million by end of 2019.[6] India's growing digital wealth is now being propelled by rural India, which registered a 35 percent growth in Internet users last year, dwarfing the all India growth of 18 percent. While data driven technologies such as AI, Machine to Machine, Internet of Things, thriving on open data regimes, are creating opportunities for business growth; India's public digital-identity system, Aadhaar, also offers a platform to correct years of mismanaged governance. Designed to prevent the diversion of public funds under the direct benefit transfer (DBT) regime, Aadhaar is rapidly becoming the basis for other new services. Because data affords immense opportunities across the development spectrum and because it is set to become all pervasive, we find it compelling to borrow the phraseology to recast 'data' as the *new Temple of a New India*.

2. DECONSTRUCTING DATA

We have been dealing with data for years, however, with continuous and fast paced innovation, everything 'digital' has become increasingly pervasive and complicated. Ubiquitous adoption of Internet and mobile technologies has resulted in abundant data being generated, collected, processed, managed and analysed every day. It has come to form the core of the global digital economy which uses this data to generate economic and social value. Industry, as we understand today, is characterised by deep digitisation of scientific discovery and knowledge.[7] Businesses are becoming progressively reliant on data for activities like monitoring production systems, managing a global workforce, managing supply chains etc.[8]Figure 2.1 is a representation of the rise in volume of data generated worldwide. Table 2.1 breaks down this estimate to capture data generated per minute on some popular applications. See: Cory (2017)



Source —Compiled from Statista

Figure 2.1: Volume of Data Generated Worldwide (in Zetabytes)

Application	Data Generated Every Minute
Giphy	Serves up 1,388,889 GIFs
Netflix	Users stream 97,222 hours of video
Snapchat	Users share 2,083,333 snaps
YouTube	Users watch 4,333,560 videos
Twitter	Users send 473,400 tweets
Instagram	Users post 49,380 photos
Spotify	Streams over 750,000 songs
Uber	Users take 1389 rides
Tinder	Users match 6,940 times
Google	Conducts 3,877,140 searches
Tumblr	Users publish 79,740 posts
Amazon	Ships 1,111 packages
LinkedIn	Gains over 120 new professionals
Reddit	Receives 1,944 new comments

Source – Data Never Sleeps 6.0 (DOMO), available at:
https://web-assets.domo.com/blog/wp-content/uploads/2018/05/18_domo_data-never-sleeps-6verticals.pdf

Table 2.1: Data Generated Every Minute in 2018

Given its centrality in the digital economy and in digital life in general, data is now the "word of the moment",[9] with terms like 'Big Data', data science, data flows, data protection, data mining, data governance, among a range of other terms pervading our everyday vocabulary. These are relatively recent, although the term 'data' itself has a history[10] that goes back to at least 100 BCE. The term *datum* is defined as "A thing given, a gift delivered or sent". The modern understanding of the term "data", began with the Euclidian revolution in geometry. The use of the word data first featured in an English language dictionary, which defines 'data' as "A term in Geometry for something proposed or known, in order to find other things unknown.

While data consists of observations, ideas, meanings, etc., it is always *about* some thing or things. There is a tendency to use data interchangeably with other terms such as 'information', 'knowledge', and 'wisdom'. There is a distinction between what data, as defined above is, and what 'knowledge' and 'wisdom' may mean. However, it may be maintained

that the term 'data' can be strongly associated with the term 'information', which can be understood as a form of aggregated data, or vice-versa.

While the definition may continue to evolve, there is no denying that data is now everywhere. It pervades all sectors of the economy. In the case of India, there are examples of data-driven business models in agriculture, manufacturing and services sectors. Within manufacturing, electronics, heavy electricals, and automobiles have seen an uptake of data driven processes that have partially replaced assembling, manual testing, repair and maintenance services. The adoption is much more wide spread in the services sector using data for customer analytics, predictive decision making and fraud detection.[11] Data is also being leveraged for several applications critical to citizen welfare services such as health, education, agriculture, and governance. For instance, Apollo Hospital in India uses big data analytics for its microbiology and infection control team. The hospital is better able to identify the cause of dengue through the use of big data.[12] There are also a host of agri-tech startups that are providing data-based solutions to tackle challenges facing Indian agriculture. These startups are leveraging satellite imagery, weather and soil data to advise farmers on input allocations and increase crop productivity.[13] Big data technology also holds immense potential for the future of teaching and learning by providing customised curriculum, improving student performance and creating new learning possibilities. Data based applications are also proving to be transformative for governance in India. Data mining techniques for instance are being deployed for effective law enforcement, disaster relief efforts, targeting of public services and social welfare schemes.[14]

A physicist would have told us a few centuries ago that matter composed of atoms is the stuff of the universe. In the same vein, a 'techie' and several others might tell us today, that data is the stuff of the digital economy. In fact, the digital economy would cease to exist in its current form without data. But this is a recent reality and while the use and impacts of data are visible and in some cases tangible, society has yet found it hard to put a numerical value to it. In India, traditional 'value' systems have made it difficult to assign a commercial value to even formal data exchange. Government has been loath to share data and businesses have benefitted more from the breach rather than safeguarding of data. All this is changing and the future might yield a robust valuation model.

2.1 The Role of Data in the Economy and Society

According to the Austrian thinker Friedrich Hayek the central economic question facing any society is how best to allocate scarce resources to their most efficient use. In the context of this paper, it is means addressing the fundamental problem of how to best utilise data and knowledge for social and economic gains. Market prices provide a useful solution by meaningfully aggregating data (information) about commodities in ways that enable efficient exchange. As articulated by Ronald Coase in his famous article, "The Nature of the Firm", firms appear as islands of pocket-sized planned economies,".[15] Discovering relevant prices and information, comes at a cost and if these costs became substantial, the market would be superseded by an organisational form called the firm, in which allocation of resources is done by fiat and not by market forces. Firms are thus a reaction to the high cost of using markets implying that information (data) comes at a significant cost. The data economy is still to test Coase's thesis. Yochai Benkler (2002)[16] proposed a new mode of common-based peer production for a digitally networked environment which is different from both property and contracts based modes of firms and markets. Since the production capabilities in this mode are widely distributed, the informational advantages allow for better assignment and utlisation of resources.

When the first impacts of mobile telephony were made, the overwhelming idea at least for emerging markets such as India was improvements in communications *(read voice)*. Mobile phones did not substitute but replaced fixed lines as the first point of connectivity. The GDP increase initially for India varied between 1.2 percent to 1.5 percent for every 10 percent increase in mobile phones.[17] These impacts were justified largely on the grounds that mobile phones substituted for other poor infrastructure in emerging markets and provided a boost to productivity. But we now know there was much more hidden in the background. Mobile phones were becoming not only points of connectivity and modes of voice communication, they were increasingly becoming devices on which information was stored, exchanged, processed and analysed. To a large extent therefore, the current estimates of 3.3 percent increase in GDP for every 10 percent increase in mobile telephony for India largely captures the new economics of data.[18] It could be also viewed as a proxy of its value.

Another counter intuitive exposition of the power of data to shape society is made by Viktor-Mayer Schönberger and Thomas Ramge in their book 'Reinventing Capitalism in the Age of Big Data' (2018).[19] They contend, controversially, that the abundance of data flows would mean that 'prices' as we know it will no more be the conveyor of market information. They explain the ongoing transition from money-based markets into data rich ones. In traditional exchanges, information on preferences were condensed into price, which albeit simplistic, failed to capture more valuable details. With technological advances that are able to tame such complexities in information processing, we can move away from our reliance on price. An outdated yet powerful example in this context is Jensen's seminal work on the role of mobile technology in enhancing the market performance and welfare of the fisherman in Kerala.[20] He illustrates the role of mobile technology in expanding the scope of information available to buyers and sellers. For instance, data on quality, quantity, price, buyer and seller, became much more easily available, minimising wastage, price dispersion, etc. Cut to today, data-driven algorithms, processing varieties of data attributes are building predictability in purchase of goods and services, winning political campaigns and generally improving lives of human beings.

Indian society, for one, has witnessed path-breaking transformations owing to the diffusion of mobile internet and the applications of big data analytics. One of the most interesting examples is that of the matchmaking industry, where mobile connectivity has enabled scores of people, even from rural areas to access a vast database of potential suitors.[21] Indian society is truly an example of converting, what could simply be a form of regular social engagement, into a flourishing industry with a revenue stream. The marriage market in India, which was earlier dependent on eager informal negotiators within ever-increasing family chains and advertisements in newspapers, has now been taken over by online portals like Matrimony.com, Shaadi.com etc.[22] According to a research report by Ken Research in 2016, this industry could earn US$ 318 million.[23] The Indian society is now witnessing a tectonic shift in the matchmaking domain, with the percolation of online dating apps. These apps use data from social media accounts which are fed into deep-learning algorithms to assess compatibility.[24]

Another technology led social transformation is visible in the healthcare sector, particularly services related to elderly care. AI algorithms are now embedded in the care-delivery mechanism, ranging from intelligent tracking of biometric information to early tracking of diseases.[25] AI based solutions have made it easier to monitor health at home, which is particularly useful for senior citizens. For example, voice-based virtual assistants such as Amazon Echo and Orbita Health are helping the elderly to adhere to medication and coordinate care at home.[26] Furthermore, voice-based virtual assistants are being optimised as nurses and caregivers by companies like Careangel, for target patient populations.[27] According to the 15-country Future Health Index (FHI) 2019 report by Royal Philips, use of digital health technology

has permeated the healthcare system in India, with 76 percent healthcare professionals using digital health records in their practice.[28] As per the report, tools such as telehealth and adaptive intelligence solutions could improve access to care and improve patient satisfaction, by removing barriers to access between hospitals and patients, particularly in tier 2 and 3 cities.

Attestations to the power of data to influence social organisation demonstrate the amenability of data to model economic and social organisation. Centralising data was the centerpiece of the socialist planned economy. What matters is data's place in the society's vision for itself. The economic and social impacts will crucially depend on the identity and the role that are assigned to data in that society.

2.2 The Economic Identity of Data

The global proliferation of the Internet and smartphones has made troves of data on various consumer activities available. While online shopping has made data on consumer tastes and preferences available, social media platforms have become a reservoir of data on changing views, moods and perceptions of people. The digital economy has transformed the abundance of data into wealth; data on consumers is heavily monetised by firms that use such data to enhance consumer experience, enable personalization, and also open up the possibility of personalized pricing to maximise profits.[29]

There are good reasons for drawing parallels to oil as a commodity. There are striking commonalities between oil and data, viewed as commodities. One such feature is the structure of markets between oil oligopolies and the current giants that deal in data, Amazon, Apple, Google, Facebook and Microsoft. To the extent that accumulation of data further strengthens the position of the big data-driven firms, they represent a similarity with the oligopoly of the oil industry. However, the contention here is that there is a distinction between the economic nature of data and the economic identity associated with it. As an economic good, data is non-rivalrous. For instance, a seat is rivalrous for it cannot accommodate person A when person B is sitting on it. Data's non-rivalrous nature means that the same piece of data can be simultaneously used in multiple algorithms, applications etc. Data, as we understand it in terms of its economic identity, is also non-fungible. Unlike commodities such as oil, which are fungible and substitutable (a barrel of oil can be replaced by another barrel), a piece of data say on consumer preference cannot be perfectly substituted. Importantly, data is not naturally occurring as oil, wherein tapping and refining would make it ready for use. The act of collecting data is deliberate and driven by hypothesis and motivations that determine its use. Unlike oil it is inexhaustible.

There are currently two competing economic identities ascribed to data, (a) data as capital and (b) data as labour. No economic identity of data has any intrinsic merit over the other, and if society were to prefer one economic identity of data to the other, it would only reflect the priorities of that society.

2.2.1 Data as Capital

For many companies' data is their single biggest asset. The genesis of the view that data is a capital asset stems from a 2011 survey of 180 large public companies, the results of which found that businesses which emphasised 'data-driven decision making' performed highest in terms of output and productivity. The study thus inferred that capabilities for data driven decision making can be modeled as an intangible asset that are valued by investors and which increase output and profitability.[30] For proponents of the view,[31] data is a new form of capital, on the same level as financial

capital is in its ability to generate new digital products and services. Data also helps create more data. For example, an algorithm focused on fraud detection, generates new data, which is yet again used as input in the algorithm to improve its performance. Data collection is driven by a perpetual cycle of capital accumulation. It is not possible to build a service with data that is not captured. Datafication, like financialisation before it, is a new frontier of accumulation.[32] Data can therefore be understood as capital, but distinct from economic capital.

2.2.2 Data as Labor

Proponents of this view[33] argue that data treated as labour is a measure to reduce inequality. They posit that the larger political economy is marked by glaring inequality with the major share of national income being appropriated by owners of capital. A major source of inequality, it is argued, stems from the fact that companies that are data intensive do not pay for the data they obtain from the consumers but in turn workaround by providing consumers with free online services.

Therefore, data as labor is a method to tackle the widening inequality that is arising out of the dynamics of the digital economy. Data as labor involves treating data as owned by individuals who, by virtue of their activities and labor create data, and requires the owners of data to be compensated for its use. To adopt the view of data as labor is also to ensure the existence of large-scale institutions to keep in check the ability of data intensive companies and platforms to exploit their market power over data providers, and ensure a fair and vibrant market for data labor.

It is beyond doubt that data is immensely valuable to businesses and societies at large, given the wide range of welfare enhancing and profit inducing applications it empowers. The economic identities of data as capital and labor are two extreme positions within a continuum of possible identities that can be ascribed to data. The economic advantage to societies in treating data in a particular way differs with differing types of data and the different purposes and contexts these serve. There isn't a one size fits all solution.

2.2.3 Data as Public Good

A popular view on data suggests that data generated from consumers should be owned by a public repository that sells data to internet giants instead of vice versa. The public good argument based on non-rival and non-excludable principles, advocates that unprecedented volumes of data and the benefits of computing power are not for businesses alone. Data can help scientists improve research, provide citizens to access government services and build services for better governance. The challenge is to organise access to data, including the rules of engagement, ethical and legal requirements, consumer protection, etc.[34] The UK has proposed a model of Data Trusts, that provide collective power over data. Data trusts are not a legal entity or institution, but a set of relationships underpinned by a repeatable framework, compliant with parties' obligations' to share data in a fair, safe and equitable way. Besides the societal and economic benefit that can be accrued through ease in access to data through data trusts, the use of data trusts allows negotiations between data stewards and prospective data users. Therefore, it enables data sharing agreements on a case-by-case basis.

India's latest Economic Survey (July 2019) makes a persuasive argument in favour of data as a public good. It proposes several measures to improve the gathering, storing, processing and dissemination of data to improve the delivery of government services. It specifies that administrative, survey, institutional and transaction data available with the government, and the data that citizens are willing to lawfully share, to be leveraged and put to its best possible use.

Open data initiatives have also promoted the development of applications that target developmental challenges such as water management, energy consumption, tracking government expenditures, improving public education services etc.[35] A study conducted in 2013 showed that access to open data can produce US$ 3 trillion in additional value annually to the world economy.[36] Several organisations including the World Bank have started open data sharing and knowledge platforms. India's Open Data initiative is available but currently underwhelming. Efforts to work with available government data can significantly improve governance within the country.

A limiting view on data as a public good flags the point that knowledge can be made functionally excludable, where the private sector gains value from controlling it. Regimes determine the extent to which data can be made excludable, for example, in the form of taxes and patents. They conclude that while it may be possible for knowledge to be a public good, it is not possible to make the same claim for digital data. This implies, knowledge produced through digital data is inherently commercial, and operates as an interaction between individuals and firms. While data is non-rival, it is not naturally non-excludable. In cases of digital platforms such as Google, Facebook, Uber and the like, data collected on the platform is treated as excludable. These "data barons" through their data collection abilities are feeding algorithms that power their ad-targeting machines and product-recommendation engines.[37] It is argued that their dominance can effect funding to startups, thus stifling innovation.[38] In order for data to serve the public interest, it is argued further, its protection under private law should shift to data ownership, management and regulation under public law.[39] Inaccessibility to data can prove to be a major entry barrier for new businesses. Data management under public law can address the need for competition and constant innovation in any economy.[40] Data regulations are discussed in the next section of the paper.

3. ECONOMIC IMPLICATIONS OF DATA AND ITS REGULATION

Global diffusion of the internet has facilitated the rise of data flows across countries that also drive global trade and commerce.[41] Rapid digitalization of businesses and their adoption of advanced technologies like cloud computing and data analytics have transformed business operations not only for IT services companies, but also those in traditional sectors. Business activities such as monitoring production systems, managing a global workforce, optimising supply chains, etc. are largely data driven.[42] Data is becoming increasingly personal in nature, capturing identifiable aspects of individuals. This type of data is being collected and analysed to better understand preferences of customers and their willingness to pay, helping firms tailor their products accordingly.[43]

Production and processing of data, form a core business activity for many firms.[44] Business operations are becoming increasingly embedded in software that requires constant updating.[45] However, many of these activities act as inputs into other industries generating value for downstream production processes.[46] A study by Jorgenson et al (2012) points to the importance of IT capital, including data services, which can constitute a large proportion of an industry's total capital input use. The study also shows how IT-capital input was the biggest source of economic output growth in the US, after 1995 and how IT-producing and IT-using sectors contributed most to aggregate productivity growth over time.

Literature holds evidence that in addition to those sectors that are commonly understood to be data intensive such as communication services, traditional industries also extensively use data in their business operations. Castro and McQuinn (2015)[47] argue that a vast number of 'traditional' industries from oil and gas to manufacturing and retail companies rely on cross-border data flows. They cite the example of Rio Tinto, a mining and metals company, that deploys data driven processes to coordinate operations across 40 countries. In its attempt to improve efficiency of operations, the company introduced an internal program to 'identify the size, location and quality of ore' by aggregating the data it collects in real time. The program involved collection of data from trucks and drills used in its mines all over the world. The data is then processed in its 'Processing Excellence Center' (PEC) located in Brisbane, Australia, which generates millions in savings across its global operations by rooting out logistical inefficiencies. In the case of air transport, which has significant data intensity, there are examples that corroborate this finding. Boeing is evidenced to use global data to reduce flight delays and cancellations.[48] Moreover, logistics companies use 'electronic wrappers' that track goods through ports and depots, from source to destination. This activity is supported by cross-border data flows.[49] Ultimately, consumers benefit from the higher efficiency and lower prices ensuing from an integrated international supply chain based on cross-border data flows.[50]

3.1 Data Regulations – Data Use and Abuse

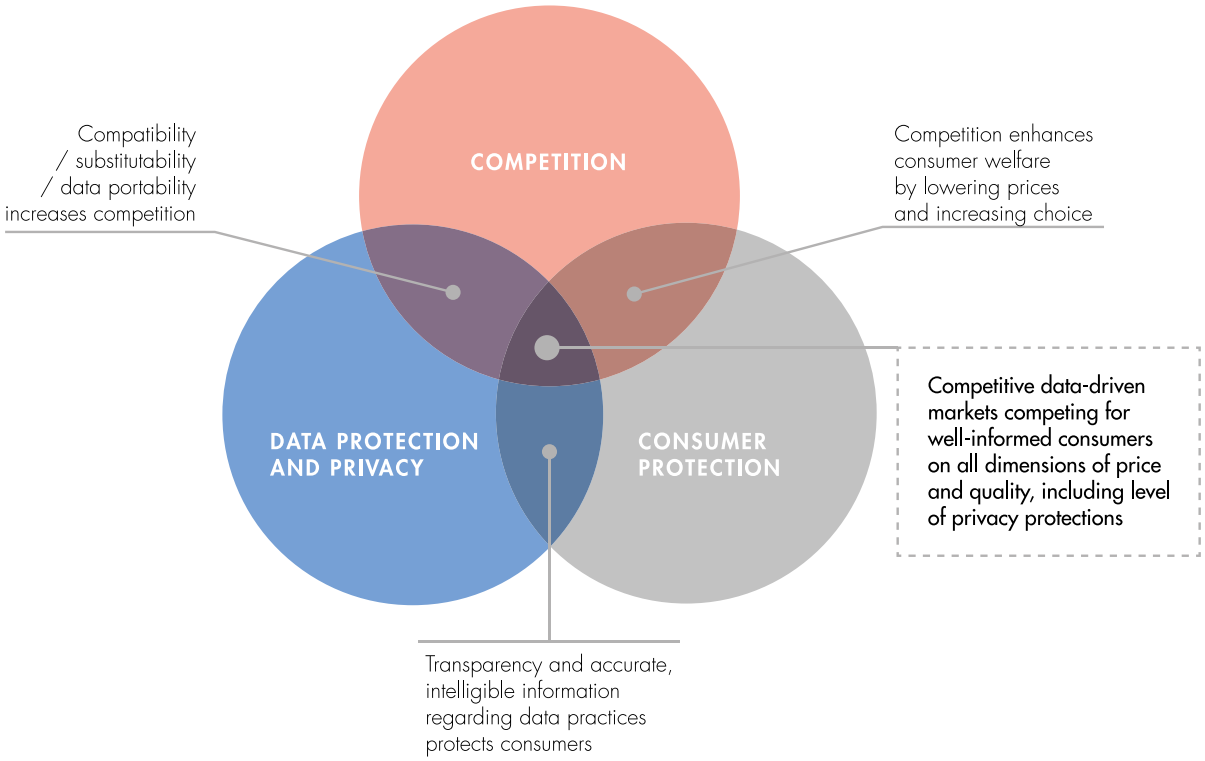
Our digital footprints are becoming increasingly permanent and indispensable to our daily existence. The value of personal data is rising with every passing day as more and more businesses fashion their strategies and develop solutions that makes use of such personal data. It is easy to imagine that a 'like' by a user on Facebook, is data that will be used by different companies to analyse user preference and target their products accordingly.[51] There is a reinforcement of network effects for firms that extensively leverage personal data. For example, users may prefer an Uber over another ride hailing service due to Uber's quicker response or due to better service quality. As more users begin using Uber, the quality of Uber's algorithm, due to greater inflow of personal data also increases the number of

users who use Uber. Such effects not only explain the high levels of market concentration across digital platforms, but also an individual's increasing reliance on such platforms for their services. It is hard to imagine a competitive and practical alternative to Google's search engine – so entrenched is Google in our everyday lives, that 'Google it' is a valid and oft-used expression. data.

There is widespread valorisation of personal data by businesses. The individual dependencies on such businesses for important services such as communication, retail, transport, etc., raise substantive concerns regarding the privacy and autonomy of individuals to whom the data belongs. This has both unprecedented and far-reaching consequences as has been most infamously evidenced in a major data scandal that broke out in 2013 with the revelations of Edward Snowden, who revealed the extent of mass-surveillance programs run by the U.S. government on citizens and nation-states around the world. More recently, the Facebook-Cambridge Analytica scandal[52] reported how millions of US citizens' Facebook accounts were breached and their data extracted to build a powerful software program that influenced electoral choices at the U.S. ballot box, which saw the election of Donald Trump as its President in 2016. According to an estimate by IDC, between 2018 and 2022, global spending on security will grow at a compound annual rate of 9.2 percent.[53] It further estimates that, in 2019, global spending on security hardware, software and services will reach US$ 103 billion, registering a growth rate of 9.4 percent from 2018.[54]

The Australian Competition and Consumer Commission's recent Digital Platform Inquiry provides recommendations span across competition law, consumer protection, media regulation and privacy law. In their framework they explain the adverse effects of the data-driven digital economy in these contexts and also provide the intersection of issues arising from these platforms. *(Refer Figure 3.1)*. For example, enabling competition enhances welfare by expanding the choice set for consumers including access at lower prices. Government action on consumer law, competition law, privacy and data protection, will be vital in dealing with any misuse emerging from the rise of data-driven businesses.

Nation-States are devising guardrails against abuse of data in different ways. Australia has implemented localisation requirements for health records, which if personally identifiable, are prevented from being transferred outside the country. Exceptions to this provision permit transfer, processing and storage of such data outside Australia in the event that the data does not include personal information or personally identifiable information.[55] Kazakhstan's directive requires new companies using ".kz" top level domain to operate from physical servers located within the country. Malaysia requires consent before any international transfer of data and China completely restricts flow of data to the rest of the world. In April 2017, China released a draft circular which extended the scope of data localization from "critical information infrastructure" to all "network operators," including any owner or administrator of a computerized information network system. Some of the other countries that have implemented data localisation are Russia, France, Germany, Indonesia and Vietnam. The European Union's General Data Protection Regulation (GDPR), developed to protect individuals' autonomy over their data is possibly the most comprehensive data protection framework of recent times. The regulation covers both personal and sensitive personal data.[56] GDPR provides individuals with better control and easier access to their data that is held by companies.[57] It also makes companies accountable for their handling of individual's personal information. Compliance with GDPR requires companies to define data protection policies, conduct data protection impact assessments and document their data processing mechanisms.[58]



Source: Digital Platform Inquiry (2019), Australian Competition and Consumer Commission

Figure 3.1: Regulatory Concerns for Digital Platforms

In India, the discussion around a data protection framework germinated in the backdrop of the potential security challenges to Aadhaar – India's ambitious digital identity project.[59] In May 2018, personal data of millions of Indians was leaked in a data breach of the Employees' Provident Fund Organisation.[60] A report by the Ponemon Institute and IBM finds that between July 2018 and April 2019, data breaches cost organisations in India, approximately Rs. 12.8 crore on average.[61] This is significantly higher than the global average which stands at US$ 3.92 million (approximately Rs. 27.03 crore).[62] In the face of such increasing vulnerabilities, a discourse on cross-border flow of data gained significant momentum since the Justice Srikrishna committee was set up to recommend a comprehensive data protection framework for India. The Draft Personal Data Protection Bill, the RBI's guidelines on local storage of payments data and the draft e-commerce policy, have all proposed localisation of data in India as a measure to protect citizen data. However, the economic consequences of such measures are yet to be evaluated, though alternative proposals have been made by industry to minimise the scope and the unintended impact of such regulations.

The nature of data regulations adopted by countries do not reflect a clear pattern between the types of countries that did or did not (would or would not) implement localisation requirements. Countries could be categorised on the basis of economic prosperity, the depth of their digital economies, legal and regulatory capacity. Concerns around protection of data are driven by the integrity and commitment of both private entities and the state, who are its gatekeepers. The

absence of a coherent international legal regime for data protection has prompted countries to devise frameworks based on their unique cultural and legal traditions. However, global positions on data regulations are evolving and countries like Indonesia and even China, which were known for their stringent regulations of cross-border data flows have introduced amendments softening their stance. More bilateral and plurilateral negotiations between countries, developing robust data security frameworks and data adequacy requirements, might be a less disruptive way of checking the negative externalities that arise from the use of data, particularly those that are personally identifiable and sensitive in nature. A one-year review of GDPR finds that while data governance has improved since it came into effect, there is uneven implementation. According to a survey by the International Association of Privacy Professionals (IAPP), less than half of the respondents said they were fully compliant and about 20 percent believe that full compliance may not be possible.[63] The high costs of compliance (reportedly an average of USD 16 million for a Fortune 500 company) implies an upside struggle for companies to meet GDPR requirements. Some handholding from regulators will nudge companies towards compliance.

4. CONCLUSION

In describing the roots of the Industrial revolution in Western Europe, Mokyr (2005) refers to a culture of 'open science' that characterised scientific knowledge as a public good. There were no entry barriers for learning the language of the scientific method, and scientific knowledge was openly communicated so that its methods and results were easily validated and replicated. The open knowledge platform reduced the marginal cost of acquiring knowledge and triggered the industrial revolution that generated enormous economic growth. Modern iindustry is characterized by deep digitisation of scientific discovery and knowledge.[64] With continuous and fast paced technological innovations, everything 'digital' has become increasingly pervasive. The global economy today is driven by the generation, collection and analysis of data. While the new generation of digital companies are understood to be data intensive, even the traditional industries are intensifying their use of data. The value of data has been rising faster than ever driving public interest in how data must be owned, transferred, and used. Data has played a crucial role in shaping important economic and social organizations of the modern economy. The culture and social values in turn will inform how society views and treats data. The economic and social impacts of data depend crucially on the identity and the role that are assigned to data in that society. For example, there are differences in the way citizens' personal data is used by governments across the world. These are mostly grounded in their long standing cultural and legal traditions as well as the overarching principles which determine the governance of a country. In China, data on citizens and companies is harnessed not only to inform decision making but also to devise a system of incentives and punishments in order to influence social behavior, as is evidenced by its social-credit system.[65] Contrast this with the EU's guidelines on ethical applications of AI by government agencies and private bodies. They help protect the autonomy and dignity of citizens, for example, ensuring that the AI software wasn't biased by a person's race or gender, in case of say, a health diagnosis, and ensure that the option is given to a patient to have their diagnosis be explained to them by a human doctor.[66]

Digital economies in advanced countries are faced with serious concerns of rising market concentration, dominance and abuse. The Digital Evolution Index, a data driven and holistic evaluation of the digital economy, uses over 100 indicators to determine the supply, demand, institutional environment, innovation and change in the digital economy for 60 countries. It captures the role of the state and the rate of digital evolution on investment, innovation and policy priorities.[67] It also measures the trustworthiness of the digital environment in each country focusing on the quality of users' experience, attitudes towards key institutions, organisations and user behaviour. India ranks poorly at 53, out of the 60 countries. The attitudinal measures of trust in an inherently networked digital economy makes it important to understand the role of data in shaping competition and innovation potential. If India gets it right, it can realise the aspiration to be a global data hub and build on its recognized software and engineering competencies to drive the digital economy.[68] The dependency on data driven services has given rise to substantive concerns regarding the privacy and autonomy of individuals to whom the data belongs. This has both unprecedented and far-reaching consequences. Regulating its use is therefore imminent. A regulatory tightrope awaits policy makers across the world that are balancing between protecting consumer interest and maximizing welfare gain on one hand and stifling innovation by over regulating on the other. The New Temple for a New India needs the foundation of a well-balanced regulation.

BIBLIOGRAPHY

Arrieta-Ibarra, Imanol, Leonard Goff, Diego Jiménez-Hernández, Jaron Lanier, and E. Glen Weyl. "Should We Treat Data as Labor? Moving beyond" Free"." In *aea Papers and Proceedings*, vol. 108, pp. 38-42. 2018.

Australian Competition and Consumer Commission, Digital Platform Inquiry (2019)

Brynjolfsson, Erik and Hitt, Lorin M. and Kim, Heekyung Hellen, Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? (April 22, 2011).

Castro, Daniel, and Alan McQuinn. "Cross-border data flows enable growth in all industries." *Information Technology and Innovation Foundation* 2 (2015): 1-21.

Chakravorti, Bhaskar, and Ravi Shankar Chaturvedi. "Digital planet 2017: how competitiveness and trust in digital economies vary across the world." *The Fletcher School, Tufts University* 70 (2017): 70.

Coase, R. "The Nature of the Firm. Economica, New Series, 4 (16), 386-405." Article first published online 19 (1937).

Cory, Nigel. Cross-border data flows: Where are the barriers, and what do they cost?. *Information Technology and Innovation Foundation*, 2017.

Furner, Jonathan. ""Data": The data." In *Information Cultures in the Digital Age*, pp. 287-306. Springer VS, Wiesbaden, 2016.

GSMA, "Cross-Border Data Flows, Realising benefits and removing barriers". (2018)

IBM, Cost of a Data Breach Report (2019)

Jensen, Robert. "The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector." *The quarterly journal of economics* 122, no. 3 (2007): 879-924.

Jorgenson, Dale W., Mun Ho, and Jon Samuels. "Information technology and US productivity growth." *Industrial productivity in Europe* (2012): 34-65.

Kathuria, R., and Mansi Kedia-Jaju. "The Impact of Internet." (2011).

Kathuria, Rajat, Mansi Kedia, Richa Sekhani, and Ujjwal Krishna. "Growth Dividends of Digital Communications: The Case for India." (2018).

Mayer-Schönberger, Viktor, and Thomas Ramge. Reinventing capitalism in the age of big data. *Basic Books*, 2018.

Ministry of Electronics and Information Technology, 2019, "India's Trillion-Dollar Digital Opportunity", Government of India

Posner, Eric A., and E. Glen Weyl. "Property is only another name for monopoly." *Journal of Legal Analysis* 9, no. 1 (2017): 51-123.

Singh, Kanhaiya, and M. R. Saluja. "Input–Output Table for India 2013–2014: Based on the New Series of National Accounts Statistics and Supply and the Use Table." *Margin: The Journal of Applied Economic Research* 12, no. 2 (2018): 197-223.

Stodden, Victoria. "Enabling reproducibility in big data research: Balancing confidentiality and scientific transparency." *Privacy, Big Data, and the Public Good: Frameworks for Engagement* 1 (2014): 112-135.

The Centre for Internet & Society, AI and the Manufacturing and Services Industry in India (2018)

Van der Marel, Erik, Matthias Bauer, Hosuk Lee-Makiyama, and Bert Verschelde. "A methodology to estimate the costs of data regulations." *International Economics* 146 (2016): 12-39

Woodward, James F. "Data and phenomena: a restatement and defense." *Synthese* 182, no. 1 (2011): 165-179.

ENDNOTES

1 Clive Humby, 2006

2 https://breachlevelindex.com/ Accessed on July 31st, 2019

3 Cost of a Data Breach Report (2019), IBM Available at https://www.ibm.com/security/data-breach?cm_mc_uid=96100784365115645554923&cm_mc_sid_50200000=63591181564555492400

4 https://kr-asia.com/indonesian-it-minister-says-to-rethink-strict-data-localization-laws

5 Ministry of Electronics and Information Technology, 2019, "India's Trillion-Dollar Digital Opportunity", Government of India

6 https://www.livemint.com/industry/telecom/internet-users-exceed-500-million-rural-india-driving-growth-report-1552300847307.html

7 See: Stodden, (2014)w

8 See: Cory (2017)ww

9 Furner J. (2016) "Data": The data. In: Kelly M., Bielby J. (eds) Information Cultures in the Digital Age. Springer VS, Wiesbaden

10 Ibid

11 AI and the Manufacturing and Services Industry in India, The Centre for Internet & Society, 2018

12 https://www.expresshealthcare.in/features/big-data-analytics-and-indian-healthcare/162330/

13 https://www.analyticsindiamag.com/top-6-indian-agritech-startups-that-are-revolutionising-agriculture/

14 https://www.analyticsindiamag.com/how-big-data-is-the-game-changer-for-indian-government-in-e-governance/

15 Coase, R.H., 1937, "The Nature of the Firm", Economica, Vol. 4, Issue 16

16 Benkler, Y., 2002, "Coase's Penguin, or Linux and the Nature of the Firm", The Yale Law Journal Vol. 112, No.3 pp 369-446

17 Kathuria, R., and Mansi Kedia-Jaju. "The Impact of Internet." (2011).

18 Kathuria, Rajat, Mansi Kedia, Richa Sekhani, and Ujjwal Krishna. "Growth Dividends of Digital Communications: The Case for India." (2018).

19 Mayer-Schönberger, V., Ramge, T., 2018, "Reinventing Capitalism in the Age of Big Data", Basic Books

20 Jensen, Robert. "The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector." The quarterly journal of economics 122, no. 3 (2007): 879-924.

21 https://www.livemint.com/Industry/Eskj6NQIwsuSfsgjURbVZO/Mobile-internet-has-changed-the-landscape-of-matchmaking-in.html

22 *Ibid*

23 https://timesofindia.indiatimes.com/business/international-business/the-web-of-matchmaking/articleshow/67585171.cms

24 https://economictimes.indiatimes.com/small-biz/startups/features/the-millennials-of-india-are-turning-to-algorithms-for-love/articleshow/67666187.cms?from=mdr

25 https://www.forbes.com/sites/shourjyasanyal/2018/10/31/how-is-ai-revolutionizing-elderly-care/#53a64201e07d

26 *Ibid*

27 *Ibid*

28 https://www.livemint.com/news/india/digital-health-technology-can-revolutionise-healthcare-in-india-report-1566457450376.html

29 https://review.chicagobooth.edu/marketing/2018/article/are-you-ready-personalized-pricing

30 Brynjolfsson, Erik and Hitt, Lorin M. and Kim, Heekyung Hellen, Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? (April 22, 2011). Available at SSRN: https://ssrn.com/abstract=1819486 or http://dx.doi.org/10.2139/ssrn.1819486

31 Most comprehensively formulated in https://www.technologyreview.com/s/601081/the-rise-of-data-capital/

[32] Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*.

[33] Ibarra, I. A., Goff, L., Hernandez, D.J., Lanier, J., Weyl, G.E., 2018, "Should We Treat Data as Labor? Moving Beyond "Free"", The American Economic Review: Papers and Proceedings, Vol 1, No.1.

[34] Privacy, Big Data and the Public Good, Frameworks for Engagement, Edited by Lane, Stodden, Bender and Nissenbaum, Cambridge University Press

[35] See: http://opendatatoolkit.worldbank.org/en/essentials.html

[36] McKinsey Global Institute (2013)

[37] https://www.technologyreview.com/s/611425/its-time-to-rein-in-the-data-barons/

[38] *Ibid*

[39] *Ibid*

[40] *Ibid*

[41] Cory, Nigel. Cross-border data flows: Where are the barriers, and what do they cost?. Information Technology and Innovation Foundation, 2017.

[42] *Ibid*

[43] *Ibid*

[44] Van der Marel, Erik, Matthias Bauer, Hosuk Lee-Makiyama, and Bert Verschelde. "A methodology to estimate the costs of data regulations." *International Economics* 146 (2016): 12-39.

[45] *Ibid*

[46] Van der Marel, Erik, Matthias Bauer, Hosuk Lee-Makiyama, and Bert Verschelde. "A methodology to estimate the costs of data regulations." *International Economics* 146 (2016): 12-39.

[47] Castro, D., McQuinn, A., 2015, "Cross-Border Data flows Enable Growth in All Industries", Information Technology and Innovation Foundation, February.

[48] http://www.ejinsight.com/20190425-how-businesses-benefit-from-cross-border-data-flows/

[49] "Cross-Border Data Flows, Realising benefits and removing barriers". GSMA (2018)

[50] *Ibid*

[51] https://review.chicagobooth.edu/marketing/2018/article/are-you-ready-personalized-pricing

[52] https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

[53] https://www.zdnet.com/article/global-security-spending-to-top-103-billion-in-2019-says-idc/

[54] *Ibid*

[55] *Ibid*

[56] https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

[57] *Ibid*

[58] *Ibid*

[59] https://timesofindia.indiatimes.com/blogs/toi-edit-page/privacy-law-needs-public-debate-data-protection-law-must-guard-against-internal-as-well-as-foreign-threats-to-citizen-privacy/

[60] https://qz.com/india/1325647/data-breaches-cost-indian-companies-millions-of-dollars-says-ibm-study/

[61] https://www.business-standard.com/article/pti-stories/organisations-in-india-lost-rs-12-8-cr-on-average-to-data-breaches-in-jul-18-apr-19-ibm-119072300710_1.html

[62] *Ibid*

[63] https://iapp.org/about/approaching-one-year-gdpr-anniversary-iapp-reports-estimated-500000-organizations-registered-dpos-in-europe/

[64] See: Stodden, (2014)

[65] https://www.technologyreview.com/s/611815/who-needs-democracy-when-you-have-data/

[66] https://www.theverge.com/2019/4/8/18300149/eu-artificial-intelligence-ai-ethical-guidelines-recommendations

[67] Chakravorti, Bhaskar, and Ravi Shankar Chaturvedi. "Digital planet 2017: how competitiveness and trust in digital economies vary across the world." The Fletcher School, Tufts University 70 (2017): 70.

[68] *Ibid*