



POLICY IMPLEMENTATION AND ALTERNATIVE POLICY OPTIONS OF BIG DATA

ARCHIT MALHOTRA

ABSTRACT

Information has been the driving force behind improved organisation and innovations. That means, more the more we have an information, the better we can prepare ourselves to achieve the best possible results. Consequently, data-collection is an essential aspect of every organisation¹. As we become more aware of this, we are producing and collecting more data on virtually everything by incorporating technological advancements in this territory. Today, we are inundated with massive amounts of data from every aspect of our lives, including socialising, science, work, wellness, etc. Further technological advancements have allowed us to generate an ever-increasing amount of data, to the point where it has become impossible to manage with standard tools. This has resulted in the creation of the term 'Big data' to describe large, unmanageable data sets. To meet our current and future social-needs, we must develop new methods for organising this data and deriving useful information. In this review, we intend to discuss the fundamentals of one such methods which is called the big data, right from its definition to its characteristics, and further on to its policy implementation and alternative policy options².

INTRODUCTION

'Big-Data' is an umbrella term for the available complexity alongside the volume of data. It refers to datasets that are too sizable for traditional processing systems and necessitates new technologies. This refers to the quantity, velocity, variety & accuracy(veracity). This indicates that not only the data is collected quickly but it is also a great way of accessing a variety of data. Mobile device, Scanner, Sensor, Digital Camera, along with R.F.I.D³. device, have become readily available and pervasive in our daily lives, digital records of individual and public activities have increased rapidly over time⁴. The internet, via broadband and wireless networks, also enables the simple and uninterrupted transmission of machine data. Further, it has also been aided by a comparatively inexpensive information extraction from "Big-Data" along with the promise of savings-cost. High-profile examples such as "*The New York Mayor's Office of Data Analytics*" also referred as the M.O.D.A. uses data for prediction with respect to buildings at a huge risk of fire. Another instance is the "*Data-driven campaign*" held

¹ Zeng Daniel & Lusch Robert "*Big Data Analytics; Shifting of perspective from transactions to Ecosystems*" 28 pg. 2-5 (2013).

² Tsai C.W.& Lai, C.F. "*A survey as to the Big Data Analytics*" (2015).

³ Radiofrequency-Identification

⁴ Horey James & Begoli Edmon "*Designing principles for the knowledge from big data*" (2012).

during the 2012 & 2016 elections in the U.S.A. have significantly increased interest. Similar developments have also occurred in Europe, with the E.S.O.(European Statistics Office) introducing the “Big Data Group” and the United Kingdom “National Office of Statistics” launching the “Big Data Project”⁵.

DEVELOPMENT

Over the past two decades, Big Data has recently become a topic of interest due to its concealed immense potential. Diverse public and commercial sector gathers stores, and analyses large amounts of data in an effort to enhance the quality of their services. Prior to the mid-2000s, it was not a matter of significant concern, especially when compared to the ‘E-Government’. However, between Jan. 2010 and March 2016, the number of Google-searches for the word nearly multiplied by sixteen, while the number for E-government continued to drop. In March 2016, the search for e-Government was barely 3% as against the Big Data. Further, the term is distinguished not only by way of its huge volume, but also by its complexity and variety. Some machine data, such as property, tax, health records, are structured and well-defined, whereas others may be completely unstructured and have no established meaning and value. Big Data is often distinguishable by its velocity, analysis & variability. Typically, data tasks like as collection, documentation, categorization, and analysis require a significant amount of time. It is usual to see a delay of several months to a few years between an occurrence and the resulting data and analysis. In the era of Big Data, however, computer and data transmission technologies have the ability to alter the duration of a decision-making cycle by enabling real-time analysis of data to inform decisions immediately. Due to the fact that data flow can alter rapidly within minutes or hours, typical analyses that are useful for studying aggregated data over weeks, months, or years may not be sufficient for problem detection and evaluation. In addition, due to the velocity and diversity of data flow, the legitimacy and importance of data for decision making have taken on new meanings. The usage of big data is often compared to the Industrial Revolution in its magnitude. Others believe that nothing has changed except that datasets have grown larger. However, scholars expect changes in how policy-making is conducted and the manner in which it impacts citizens. While, the first group desires swifter-decisions, often supported by data having low/less chances of uncertainty. While the critical voices also review the challenges identified in the discussion of “the evidence-based policy”, in which various sources of information compete in the policymaking process and demand the comprehension capacity of decision-makers⁶. While the actual use of ‘Big-Data’ for policy-making is not novel, but the manner in which their potential or actual usage modifies the conceptual and logical arguments around decisions-making is novel.

CHARACTERISTICS

‘Big-Data’ thus, consists of voluminous data that cannot be processed by the traditional data storage. Many global organisations utilise it to process the information(data) and businesses of various organisations. Before duplication, the data-flow would exceed one hundred and fifty exa-bytes per day⁷.

⁵ De Mauro & Grimaldi M. “*What is Big Data? A Review of Key Research Topics*” (2014).

⁶ Giest S. “*Big data for policymaking: fad or fastrack?*” pg. 367–382 (2017).

⁷ Degli Esposti S. “*When Big Data meets dataveillance: The hidden side of analytics*” pg.209-225 (2014).

When it comes to explain the characteristics of Big Data; it majorly consists of the “5-V(s)”:

- Veracity
- Volume
- Value
- Variety
- Velocity

Veracity: Veracity refers to the reliability of the data. It offers numerous data filtering and translation options. Veracity is the process of efficiently handling and managing data. Big Data is also vital for corporate growth. For instance, Facebook posts include hashtags.

Volume: The term Big Data implies a “Huge quantity”. Big Data is the huge “volumes” of data generated daily from a variety of sources, such as business-processes, social media platforms, machines, networks, and human-interactions. Facebook can generate around one billion messages, the ‘like’ button is clicked around four billion times each day, and more than three hundred and fifty million new postings are submitted daily and the Big data technology is able to manage such vast quantities of data with a significant ease.

Value: One of the most essential characteristics of big data is its value. It is the data itself that is processed or stored but we store, process, and evaluate useful and trustworthy data.

Variety: Big Data can either be structured, unstructured, or semi-structured information gathered from various sources. In the past, data was collected solely via databases and spreadsheets. However, data, nowadays comes in a variety of formats, including Pdf, Email, audio files, social media postings, photos, and videos.

The data is classified as follows:

- Structured data** which includes data in a structured scheme along with all mandatory columns. It is tabular in format. The Relational Database Management System (R.D.M.S.) stores the structured data.
- Semi-structured** which the scheme is insufficiently defined. Online Transaction Processing systems often known as O.L.T.P. are designed to handle semi-structured data. It is stored in tables, or relations.
- Unstructured Data** which includes all unstructured files, audio files, log files, and image files. Some organisations though have an access to het abundant raw data, but they do not understand as to how to extract the data’s value.
- Quasi-structured Data** which includes format includes textual data with non-consistent data formats that require time and effort to format with certain tools.

Velocity: Compared to other factors, velocity plays a significant influence. Velocity determines the rate at which data is generated in real-time. It includes the connecting of incoming-data set speeds, change rate, and activity

bursts. The basic function of Big Data is to supply data on demand quickly. It thereby pertains to the rate at which data flows from sources such as business-processes, networks, social networking sites, mobile devices, sensors etc⁸.

POLICY DEVELOPMENT AND IMPLEMENTATION IN THE ERA OF BIG DATA

In order to make sure that the Government benefits from the availability of appropriate data and introduce and implement efficient policies, innovative and enhanced methods and tools are required to support and establish the policy-design and implement them so as to facilitate the perception of “*Big data for public communication*”, which would include new output methodologies⁹. In order to enable the public-administration to explore the opportunities offered by it; it is necessary to have a comprehensive understanding of the legal framework and to consider sociological, political, cultural, economic, legal, and behavioural factors¹⁰.

Elements of Policy making:

- “**Methodological-development**” for the use of big data in policy formulation, including an examination of the readiness of policy-making institutions and systems to analyse and absorb massive data;
- **Multidisciplinary evaluation** of political, economic, ethical, epistemo-logical and legal foundations and consequences of all the big-data activities allowing consideration of potential risks and benefits;
- **Developing scalable tools and modes** for data compilation, visualisation, analysis and incorporating pertinent open, certified and official data;
- **To understand the consequences of the increasing data-materiality** as a result of the growth of the “Internet of Things” and its ramifications for the long-term viability of the government’s effective use of big-data for improving the policy formulation;
- Create methodologies and tools for implementation and the monitoring of its effective compliance.

Such proposals should apply their methodology to policy domains that address societal issues such as the migration, environment, inequalities, radicalisation, unemployment, goods and services. The characteristics such as the Independence, coherence, consistency, quality, confidentiality, neutrality and objectivity, as well as faithful representation to important groups, must be considered when employing accessible and Big Data to expand the evidence basis for effective policymaking. Issues such as the Data protection, privacy and ethics concerns must also be addressed, along with ethical concerns around the storage, usage, and re-use of data. Whenever applicable, proposals must also assess the appropriateness of the suggested software. Further, as societal concerns

⁸ Dijcks Jose “*Oracle: Big Data for the enterprise*” (2012).

⁹ European Commission (E.U.) https://cordis.europa.eu/programme/id/H2020_CO-CREATION-06-2017 (2017).

¹⁰ Akerkar, Rajendra, Finn Rachell, Lovoll Grunde, Faravelon Aurelien, Wadhwa Kush& Grumbach Stephane “Understanding and mapping Big Data” (2015).

have become increasingly complex and interconnected, public-policy innovation and implementation using I.C.T. can enhance the efficacy, efficiency, and quality of public sector decisions. Effective and trustworthy rules must take into account the available data and the requisite proof to ensure that information is accurate and relevant. Big data has numerous opportunities, including the use of analytics to develop unique insights, the enhancement of prediction ability, and the identification of unanticipated patterns and linkages that can inform policymaking. It can assist public authorities in detecting and proving patterns of non-compliance in various areas of policy ranging from safety, health, welfare of individuals in the internal market for products, services, and persons. Efficient processing capacity and knowledge are widely employed in the retail and commercial sectors; the task, however, is to establish an effective resource to make this available to government, so enabling policy decisions to become more evidence-based and analytical. Further, open-policymaking and the incorporation of the citizens' perspective through effective engagement of key social actors - for instance, via digital sites or crowdsourcing have the much-needed potential to generate large amounts of data, which can result in more informed policy decisions. Moreover, a transparent policymaking can foster a collaborative, open, and participatory government perspective. In addition, policy-implementation can certainly benefit from efficient monitoring and enforcement technologies that are informed by data from several sources. Proposals must demonstrate the benefits that will be realised following the project phase, including in the form of enhanced public policy efficiency, increased efficiency, highly precise gains, better-consistency, and reliance on evidence leading to increased policy compliance, as well as in terms of the democracy, such as Transparency, Good-Governance, and increased trust in and perceived legitimacy of government. Increased accessibility to non-governmental entities may also be an additional effect.

A LEGAL PERSPECTIVE TO BIG DATA IN INDIA

LEGISLATIONS RELATED TO BIG DATA

With the extraordinary growth and broad acceptance of the advantages of “Big Data”, there also has been a great deal of emphasis on the obstacles and hazards connected with its collection and utilisation. If not adequately handled, the hazards could exceed its benefits and erode the widespread approval for the industry-wide implementation of Big Data. Despite this knowledge, however, countries have indeed been slow to handle it. There is no comprehensive or singular or privacy law/ data protection in the U. S. any organisation wishing to engage in the “Big Data activities” must instead comply with a variety of regulations, including sector specific privacy laws governing the data involved in their respective businesses. This is often regarded as the ‘Patchwork of Laws’, and it may be observed in multiple diverse. However, India currently lacks a comprehensive, unique, and dedicated data protection law, as well as regulations relevant to Big Data. For the purposes of Data-protection, the I.T. Act ,2000 was revised to include the “Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information), Rules, 2011”. Further, though the Indian government introduced “the Personal Data Protection Bill 2019” (P.D.P. Bill) but it has now been withdrawn. However, the Indian government has now pulled the P.D.P. Bill from the Parliament in August, 2022, while it contemplates a

“complete legal framework” to handle the online sector in order to encourage innovation in the country via the introduction of a new bill. The Joint Committee of Parliament conducted a comprehensive analysis of the P.D.P. Bill of 2019. Eighty-one changes and twelve proposals were suggested for a complete legislative structure on the digital ecosystem. Further, based on the J.C.P.’s report, a thorough legislative framework is being developed. It has therefore proposed to withdraw the bill and therefore, in the light of the committee's recommendations, the government has chosen to introduce a new bill that fits within the existing comprehensive legislative framework. This indicates our country still has a long way to go from a legislative perspective in the context of data bill¹¹.

LEGAL CASES

It was first acknowledged as a consideration with reference to the competition law in 2012 by the “Competition Commission of India” also known as the C.C.I. In the case of “*Matrimony.com Limited versus Google L.L.C.*”¹², where it was alleged that Google had abused its dominant position within the online-search advertising market through biasness and promotion of its own vertical search sites results in the manipulation of the results. In 2017, the Competition Commission examined the case of “*Vinod Kumar Gupta versus WhatsApp Inclusive*”¹³ and raised the issue of the appropriateness of the use of big data as to the competition law. However, the C.C.I. found that WhatsApp permitted its users to opt out of sharing account information with Facebook within thirty days of the new policy’s implementation. In this instance, the commission has stated that problems of privacy and privacy and did not fall under its jurisdiction and that such violations should be dealt by other authorities. The Act of 2002, which controls does not include big data and artificial intelligence (A.I.) within its scope. This is one of the greatest obstacles India confronts in relation to big data and AI. In the case of “*Rajasthan Cylinders and Containers Limited versus U.O.I.*”¹⁴; the Apex Court ruled that the C.C.I. is responsible for promoting fair and healthy market competition. Therefore, it is essential to enact laws that particularly address Data and Artificial Intelligence so that the current Commission is better suited to handle disputes concerning these topics.

ALTERNATIVE POLICY OPTIONS FOR BIG DATA

1. Data-Readiness and Digital-Era Governance

The “Data Readiness Concept” and “Digital-era Governance” (D.E.G.) are the new concepts functioning as the policy-options for this phenomenon in the public sector. It is based on the conceptual notion that “Data & Technological innovations in the sector government requires an infrastructure for deriving value from data. Further they are linked to the notion of the electronic-government technologies. It aims at the transformation

¹¹ “*Big Data, Law & Regulation of Big Data*” <https://www.lawyered.in/legal-disrupt/articles/big-data-law-regulating-big-data-sanjay-mehta/#:~:text=However%2C%20the%20Information%20Technology%20Act,the%20protection%20of%20personal%20data.>

¹² Case No. 07 & 30 of 2012.

¹³ Case No. 99 of 2016.

¹⁴ (2018) SC.C. OnLine S.C.1718.

of the government to hold it more responsive and accountable¹⁵. It is often highlighted as the macro-theory for public sector growth. This concept evaluates the public capabilities by examining the data readiness of organisations and raising considerations complementary to D.E.G. This notion majorly focuses on the alignment, competencies, and development of organisations in relation to the Big-data¹⁶.

2. Evidentiary-policymaking

They relate to the notion of “Evidence-Based Policymaking”. It majorly focuses upon the receptivity of the policy formation “cycle” to such input, and prevention of a disagreement regarding how and where evidentiary contributions might add value to the process. The integration of this phenomena into an existing institutional context and the ability of individuals or government bodies to locate and utilise “the data-based information” are two issues presented in this context. Both are connected in the sense that limited government capability can lead to the participation of more players, which ultimately increases institutional complexity. When government have poor levels of analytical capacity, they run the risk of implementing “scientific knowledge ineffectively into decision-making” process, as described by the term ‘political analytical capacity’. This typically leads in the additional stakeholders with the necessary capabilities to extract pertinent information from the supplied resource¹⁷.

3. Data-culture among organisations

The “Data culture” in the Public Organisations is the knowledge that big data is not just an Information Technology issue, but also an issue that requires attention from organization-structures and skills. Specifically, it emphasises the significance of civil servants and politicians understanding how to locate, evaluate, and exploit big data, as well as the institutional structure to enable this, such as training or the exchange of data between the government agencies. This is referenced in passing in the vast majority of contemporary writing on public-policy and Big Data¹⁸. Similarly, the D.E.G. framework also embraces the notion that the literacy levels associated with new technology in government are commonly insufficient, and thus, the respective employees must be trained in skillsets¹⁹.

CONSTRAINTS

Big Data is not a magic cure, despite some promises to the contrary. This indicates that like every coin, Big Data has its other side as well which is often regarded as the constraints of Big Data. Big Data solutions are not applicable or suitable for all security issues, and like all other tools, they have both intrinsic strengths and weaknesses. Even in this digital age, the correct data is not always accessible. While at times the data just, doesn’t exist, while other times there are issues with data retention time, and at times the different data

¹⁵ Holmes Bev & Best Allan “*Systems knowledge, thinking, & action ;Towards better models & methods*” 6(2), pg. 145–159 (2010).

¹⁶ Bertot, J. & Choi, H. “*Big data, open gov. & e-gov.*” 19 pg.5–16 (2014).

¹⁷ Crouse, K., Gitomer D., Joyce J. “*A Review of the D.C. Impact*” (2013).

¹⁸ Howlett M., & Perl, A., Ramesh, M. “*Studying Public Cycles, Policy Policy & Policy Subsystems*” Oxford University Press (2009).

¹⁹ Yiu, C. “*The big data opportunity; Making government faster, smarter & more personal*” (2012).

platforms reveal to be incompatible with one another. The quality of data is also not guaranteed, as it may be out of date, distorted, biased, or even falsified. These vulnerabilities can make their way into Big Data analysis, undermining efforts to increase security via Big Data analyses.

CONCLUSION

In a report from “The White House” (2014), President Barack Obama emphasised that the Big Data often leads to ‘vexing issues’ which implies that these technologies have the capabilities to cause societal harms beyond damage to our privacy, for instance the discrimination against individuals and groups, while at the same time highlighting the massive opportunities they tend to offer to with the aims of improving the public services, as well as directing the economic growth and improvement of the safety and health of our communities. Further, a study of more than one hundred and eighty pieces of primarily recent literature and pieces of actual practical evidence reaffirms the fact that the “Big Data” poses both opportunities and threats for our future development as an unprecedented volume of cost-effective data can be utilised to guide decision-making in critical areas of growth, such as security, health care, economic productivity, disaster- and resource management, among others. The continuous transition from the “Information Age” to the “Knowledge Age” appears to be the next logical step in the extraction of practical knowledge from huge quantities of available digital information. Thus, it can conclude that the “Big Data” is a technological advancement, which implies that the spread of technological breakthroughs is never instantaneous and uniform, but invariably results in divisions for the duration of the diffusion process. Thus, this review argues that this technology is undergoing an unequal diffusion process. This inexorably produces a new aspect of the disparity; a split in the ability to bring the analytical processing of data at the foreground of very well decision-making, and hence a divide in data-based knowledge. These development obstacles contribute towards the inherent dangers of the “Big Data paradigm”, which also includes issues over the “States and Corporation control”, as well as the blind confidence in faulty algorithms. This demonstrates that the introduction of the “Big Data” is not a fix. This transformation must be well accompanied and guided by proactive policy alternatives and targeted development projects.