Big Data: Transforming Evidence based Policy Making

Authors –

Ashish Desai[1]
ashish.desai@gov.in

B. Sharlo Zack[2]
sharlo.zack@nic.in

Sambuddha Goswami[3]
sambuddha.goswami@gov.in

Abstract:

With proliferation of technology and innovation, the creation and consumption of data has been growing at an unprecedented rate. Data-driven innovation can be transformative, as it provides plethora of opportunities to improve decision-making make more informed policies and strengthen economies by reducing waste and improving services. Data can improve social and economic outcomes, if it is used systematically for generating evidence for formulating policies. So far improving governance, through the concept of e-Government and e-Governance, has been the focus of technology utilization in public administration. However, use of big data can alter the pace of development. Appropriate policies and technologies must be structured to appropriate leverage big data. All this leads to the following question: How can public sector use big data more efficiently and effectively? Using the policy cycle as a generic model this paper explores, how evidence-based policy making could be mor efficient and effective by use of Big Data. It attempts to articulate a holistic process addressing issues of technology, citizen, and users, which often are moderated by culture and governance.

1. Introduction

Applications driven by development in technology and the advancement of mobile communications has generated a "data gold mine" (Kernaghan, 2014). Telecommunications, along with advancements in cloud computing and storage techniques has boosted Big Data adoption. In line with developments scientific paradigms have also evolved through empirical science, theoretical science, and computational science. Now we are at the stage of data-intensive science - the latest approach to discover knowledge or extracting value through derivative of technology-Data (Chen et al, 2014; Chen & Zhang,2014). Applications of data has the potential to generate economic and social value far beyond those originally anticipated.

---

[1] Ashish Desai is Director at DMEO, NITI Aayog
[2] B Sharlo Zack is Young Professional at DMEO, NITI Aayog
[3] Sambuddha Goswami is a Senior Research Officer at DMEO, NITI Aayog

So far, in the Governments, improving governance, through the concept of e-Government and e-Governance, has been the focus of technology utilization in public administration. Referring to Janssen and Kuk (2016). Governments can also leverage such techniques to enhance accuracy, efficiency, speed of policy formulation and implementation, and evaluation of such interventions through the application of Big Data Analytics. Yet much of the value of data remains untapped in the public policy domain and governments are yet to capitalize on the huge potential of this technology (Gamage, 2016). We need Data-driven innovation that is transformative and has the potential not only to make more informed policies but also to provide appropriate and applicable public services for the disadvantaged class (Manzoni, 2018). All this led to the following question: What are the areas where big data and related methodologies can be applied to enhance public administration and plausible approach for execution can be explored.

Using the policy cycle as a generic model this paper explores, how evidence-based policymaking can be optimized by data. The paper further aims to build a plausible approach that could be further optimized and institutionalized to enable the use of Big Data in Public Policy Making could be supported and promoted using Big Data and its application.

2. Data Sources

Although data can be used to improve development outcomes, the challenges differ across data types based on original intent and collection methods. To conceptualize these data types and better understand those challenges, WDR 2021 sorts of data types using a two-dimensional framework shown in figure 1.0. One dimension classifies data based on original intent - public or commercial purposes. Traditional as well as new age data collected for commercial use is classified as Private Intent Data; Similarly, data collected for public intent, irrespective of method or instrument is classified as " Public Intent Data". By design, public data can be focused on purpose of its application and may be a representative of population; while private data may have a self-selection bias e.g. OTT platform data might be more incidental than representative of population.
The second dimension represents methodologies - “traditional” Vs “new” methods.

Table 1: Examples of data types based on original intent and collection methods

Data collection methods and tools	Public intent data	Private intent data
Traditional	Census, national accounts, household surveys, enterprise surveys, labor force surveys, surveys of personal finance, administrative records	Any survey conducted by private entities, including public opinion surveys deployed by private entities; administrative data from company financial accounts
New	Location data from satellite imaging, digital identification, facial recognition from public cameras, public procurement data from e-government platforms	Just-in-time digital data on individual behavior/choices from digital platforms in the private sector
Source: WDR 2021 team.		

2.1. Public Intent Data:

Public intent data are typically associated with traditional data types such as censuses and surveys, although geospatial have become more prevalent. Government data collection is for public purposes, primarily for policy making. But because the collection of public data via traditional methods tends to be relatively costly, surveys are conducted infrequently, and they often lack the granularity necessary to make meaningful inferences about subpopulations and micro segments. Meanwhile, traditional public intent data offer important advantages over new private intent data in terms of their coverage of the population—which makes them amenable to inferential analytics by researchers and government officials. Distinguishes between six types of public intent data that all serve the public good explained below.

2.1.1. Six types of public intent data

- **Administrative data**— **Administrative Data** refers to data sets collected for transactional purposes, rather than being designed from the outset as a dataset for analysis by researchers (e.g., conducting repeat social science surveys), or forming part of the carefully constructed and evaluated national statistics reporting. Records such as such as birth, marriage, and death records and data from identification systems; population, health, education, and tax records; and trade flow data e.g GSTN —are generated by a process of registration or record keeping, usually by national authorities. Administrative data also include data used by governments to run projects, programs, and services. For instance,[4]. Administrative data

[4] Mahatma Gandhi National Rural Employment Generation Programme

can often be collected unobtrusively and non-reactively, so the challenges of 'faking' survey responses is minimal.

- **Censuses**, is aimed at recording information about an entire population. It can be aimed at individuals or business or be agriculture focused. For e.g., population and housing censuses provide essential information on the entire population and all dimensions required.
- **Sample surveys - drawn to be representative of the entire population, using appropriate sampling technique.**
- **Citizen-generated data** – Generated by individuals. Primarily focused in solving issues faced by citizen. This data, can provide significant value in solving citizen centric issues. Geospatial data can be enriched by combining with citizen generated data

- **Machine-generated data** automatically generated by a sensor, application, or computer process. E.g.- Air pollution monitor.
- **Geospatial -** It primarily is satellite imagery data, providing multiple layers of information based on their geographic locale. However, with progress of technology, drone data is also adding significant value.

These data types are neither exhaustive nor mutually exclusive but gives a different dimension. Every data point can then be represented with reference to time, space or person. Various data sources can be linked together through common identifier.

2.2. Private intent data

Private intent data are often, associated with digital tools and application. Though it lacks representativeness of whole population, this data has far greater frequency, granularity as compared to Public Intent data. Private intent data collected through cellphones, internet usage, etc., provides geolocation information, which the conventional methods of public intent data collected through survey just cannot provide.

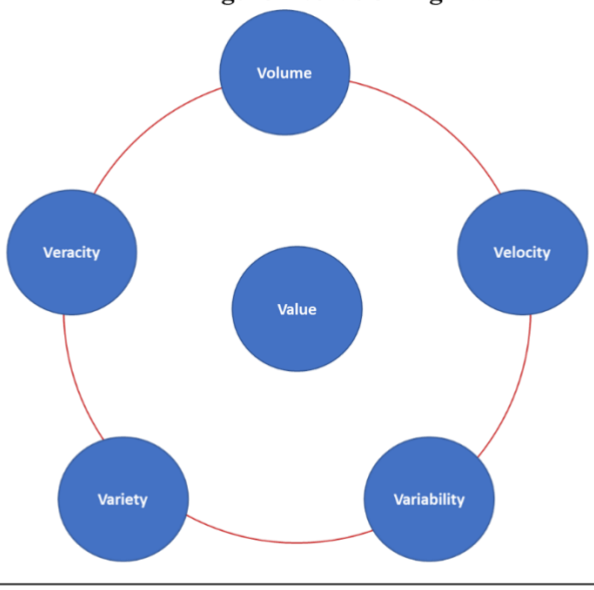
Everything a digital user does- call, text, search, leave a post, make a transaction, or accept services terms and conditions leaves a trail, which is called as a digital footprint.

The footprints can also be passively collected by tracing IP address and search history. Digital residuals have exceedingly high frequency and microgranularity, enabling to insights, inferences and understand characteristics at an individual level. Mobile network data is one of high velocity data providing information on behavior, attitude and location. Call Detail records (CDR) is intended to use for billing, it can be repurposed for behavior, attitude, and mobility patterns. Firms which rely on data, such as e-commerce or social media firms, additionally generate behavioral patterns. Combining, variables can assist in prediction of economic and social status and of behavior pattern. For example, having an iOS device consistently correlates with being in the top quartile of the income distribution in the United States. Similarly., In Germany back testing provided of e-commerce data showed time of day had a correlation with buyer's self-control and

repayment behavior. Consumers who made purchases during day were half as likely to default as compared to consumers who shopped at night.

Data residuals have big data characteristics. Various definitions of Big Data exist in literature. Initial literature speaks about "Three Vs", i.e. "Volume" representing size of data, "Velocity" representing the speed at which data accrues and "Variety" representing varying dimensions of data (Jordan, 2014; Kitchin & Lauriault, 2014, Laney 2001). Subsequently, 3 more characteristics for explaining certain features of Big Data were added i.e., "Veracity" - data quality, "Variability" - variation in data, and importantly "Value" - the potential hidden intrinsic value (IDC, 2012; White, 2012). Big data is classified in four ways: 1. Structured. 2. Unstructured 3. Semi-structured and 4. Quasi-structured .

Figure 1 :6 Vs of Big Data



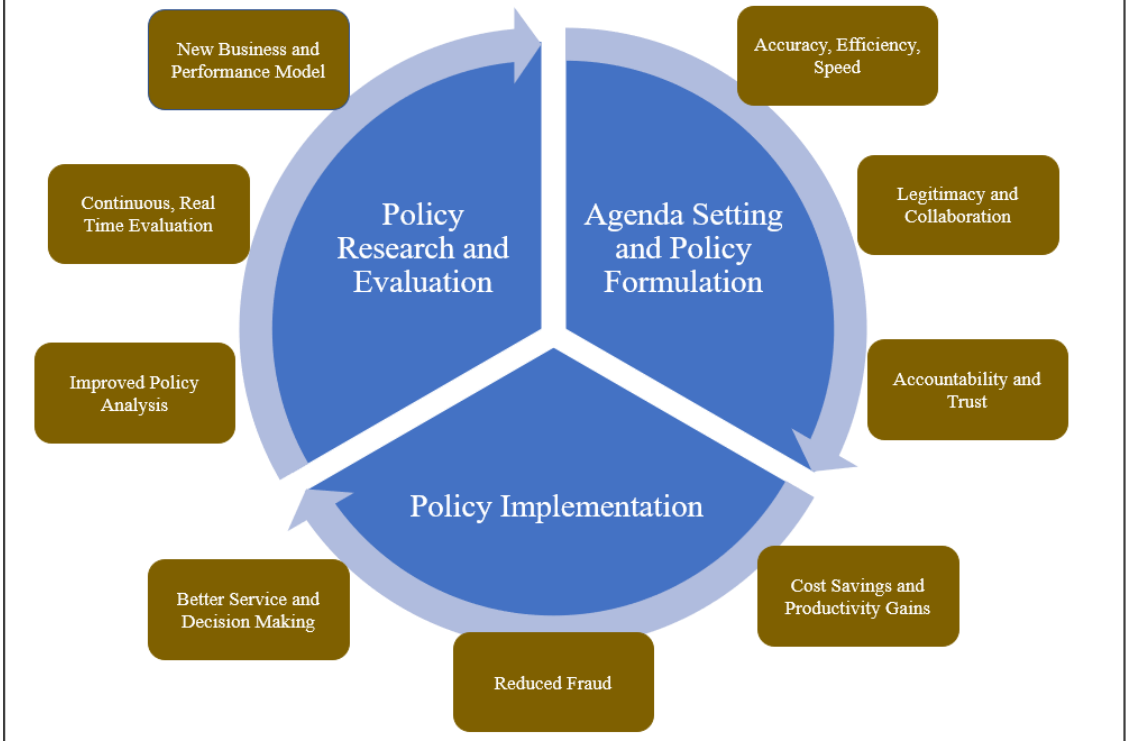
Adapted from White, 2012

3. Policy Cycle and Big Data

Public administration has ample opportunities to move beyond improvement of already existing services and Structures, via e-Government and e-Governance, towards transformation of the policy-making structure itself. Johann Höchtl, Peter Parycek & Ralph Schöllhammer (2016) describes how, through use of Big Data enabled computations, traditional "policy cycle" can be shortened by incorporating the concept of continuous evaluation, instead of introducing evaluation as a separate step at the end of the policy-making process. The policy cycle is composed of several inter-related steps- Policy Formulation, Policy Implementation, and Policy Evaluation (Wegrich and Jann 2007). Utilization of Big Data can optimize the policy cycle by bringing efficiency and

effectiveness in policy formulation, execution, and evaluation (see Figure 2). It can, also, make policies more relevant, focused, and customized- creating a new model of public policy in the digital age (Clarke and Margetts, 2014).

Figure 2 :Big Data impact on Policy Cycle framework



Source : Pencheva et al, 2020.

3.1. Agenda Setting and Policy Formulation

Big Data has the potential to increase the legitimacy of agenda-setting and policy formulation by enabling citizens and governments to engage in more meaningful dialogue and collaborate in policy design (Dsouza and Bhagwatwar, 2012). In a world of social media and the digital presence of citizens, governments can generate relevant agenda points by collecting data from social networks with high degrees of participation. This can assist policy makers in identifying citizens' policy preferences, which can then be considered by the government in setting the agenda. Big Data allows for a variety of sources to be researched and analyzed to understand the policy landscape (Loftis and Mortensen, 2018; Nowlin, 2016), identify problems, and conceptualize solutions more accurately and in greater detail (Williams, 2014). Availability of swathes of

unstructured data over the internet can be treated with big data methodologies to identify the areas that should go to the top of the list when it comes to improvements in development outcomes.

Information can support open policy discussions and debate, ushering efficiency implementation. In policy formulation, Big Data can not only help to design a policy that closely matches the preferences (Stritchet al., 2017; Taeihagh, 2017) but also in developing different scenarios and accurately predicting their possible outcomes (Cook, 2014). The high velocity of data also enables policymakers to quickly (re)act to and incorporate collective information from a variety of sources, including channels outside the formal policy realm such as social media, online consultations, and virtual town halls (Hochtl et al., 2016; Mergel, 2017).

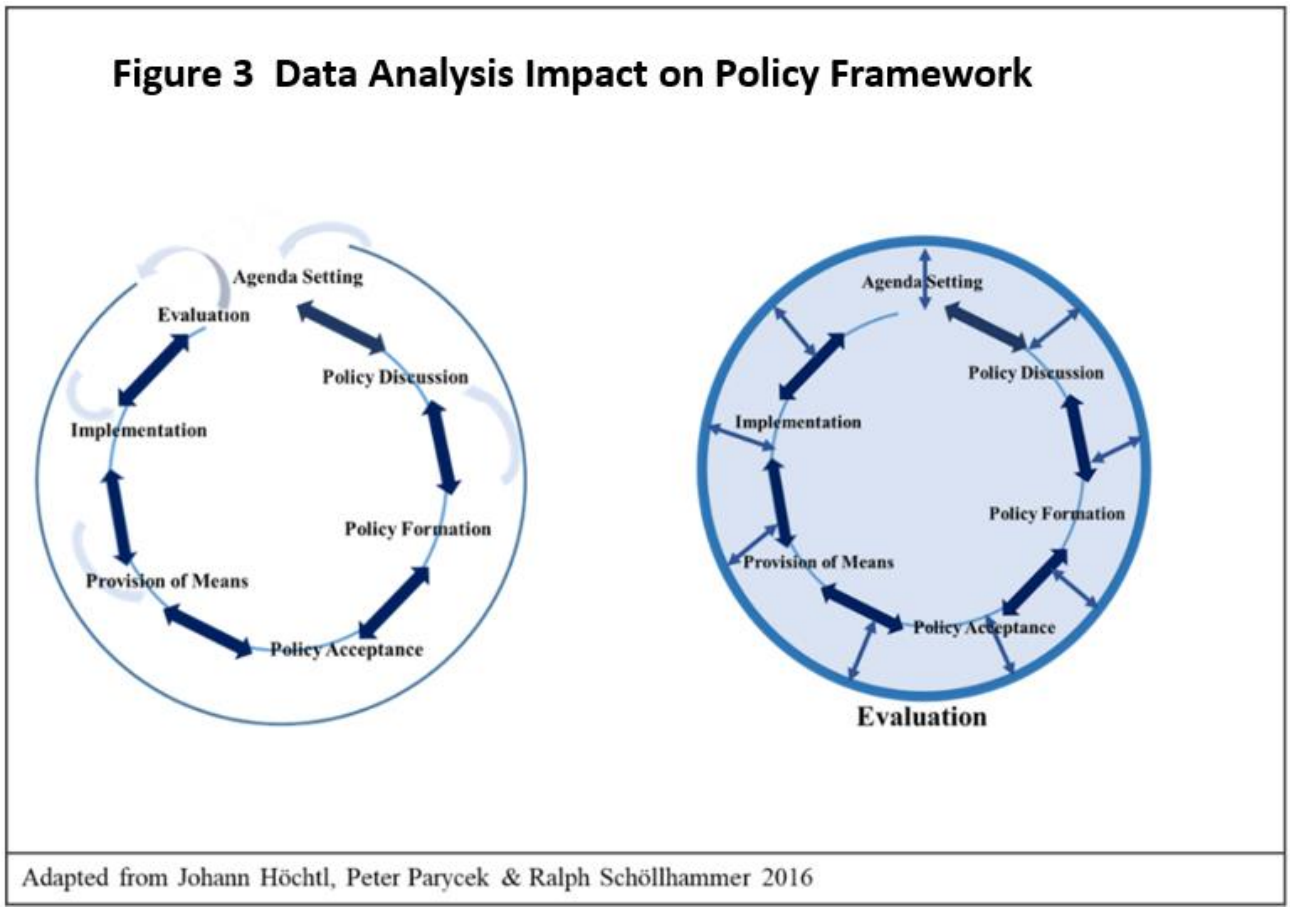
3.2. Policy Implementation

A more transparent and more performance-oriented provision of resources can be promoted by using big data-related procedures to identify patterns through analysis of data generated as part of the budgetary process. The availability of more data can facilitate a shift towards outcome-oriented budgeting and the structuring a framework that could appropriately focus and priorities resources. It could enable the identification of hidden patterns and enables us to identify pathways to research difficult to reach citizens. Big Data can lead to better supervision of the implementation process through the detection of irregularities (Maciejewski, 2017). For example, in Singapore, analytics techniques are used to analyze 12 million transactions daily to plan and better operate the transport system (Maciejewski, 2017).

3.3. Real-time time Evaluation

Evaluation can be done instantaneously or near-instantaneously at the very moment data arrives. This enables a new view on the policy cycle, namely that of continuous evaluation as opposed to evaluation happening at the end of the policy cycle in Figure 3.

Figure 3: Data Analysis Impact on Policy Framework



Big data analytics-driven policy cycle wherein evaluation does not happen at the end implementation of policy but is a continuous process of validation, assessment and recasting. This new feature is also termed as e-Policy. (Johann Höchtl, Peter Parycek & Ralph Schöllhammer, 2016). Data governance across organizational boundaries and with comparable data parameters across public organizations is imperative for leveraging data for governance. Appropriate policy and regulatory frameworks to promote cross-organizational collaboration are imperative (Penchava et al, 2020; Desouza and Jacob, 2017).

4. Connecting Data for Development

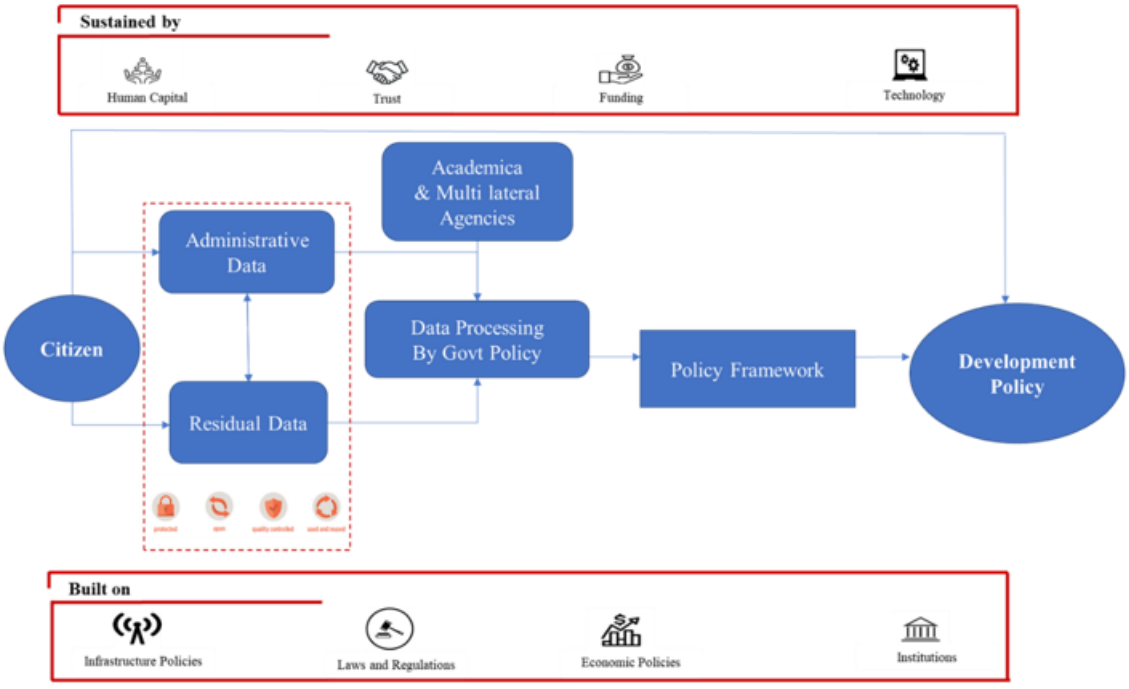
Materializing the points mentioned in the above sections are contingent upon availability of quality and granular data which is representative of the population being targeted.
Public intent and Private intent data have advantages and disadvantages and pose distinct challenges in terms of reuse and exchange. But because public intent and private intent data have inherent complementarities, they can be used jointly to bolster development. A ministry of health would be able to issue better public policy if it could connect its health data with that of other ministries such as education, labor, and planning, as well as with that of health providers, whether public or private, around the country.

For this purpose, world bank (2021) envisages an Integrated National Data System designed to realize the potential of data for development. As per World bank (2021) if the governance system is viewed as establishing the rules of the road (and the institutions governing those rules), the Integrated National Data System (INDS) can be envisioned as a network connecting all users and ensuring safe secured data exchange. An integrated national data system implies that all participants and stakeholders - government, civil society and the private sector - collaborate in a system in which data are safely produced, exchanged, and used. It does not mean that all data are stored in an integrated national database or has a centralized governance structure. It is an ecosystem, wherein the data can be accessed, shared, used, and reused for the larger public good.

The conceptual framework links data to development (Figure 4). Data is contributed by government administrative systems or private sector platforms driven by technology. Data can be processed by various participant entities, for public policy analysis, primarily Government entities which happen to be the primary producers of public intent data as well. Civil society and Individuals play a critical role as the producers and users of data that hold governments responsible for issues of public concern. Academic institutions, think tanks, and research organizations support and build programs for the analysis of data and policy suggestions. Most importantly private sector is an important contributor to data production in the national data system, whose data are subject to common standards and quality control. To manage this ecosystem, it is imperative, to set out norms and have an integrated framework of infrastructure policies, open data

laws, appropriate information legislation, and policies. Often cited is how the policy framework can address the issues and challenges of data security and privacy.

Figure 4 : Data for Development



Source : Adapted from WDR , 2021 ; Authors Input

In this context of integrated framework, addressing personal and organizational privacy is important. Personal data points might permit connecting seemingly unrelated data points to provide hidden patterns; however, it is imperative not to undermine citizen's privacy. Techniques such as anonymizing data or informed consent can address challenges posed by data privacy (Penchava et al, 2020). Another important point is data security. The high volume of data might make it susceptible to cyber attach and cybercriminals.

To balance the use of data, World Development report 2021 calls for a new social contract to enable data use to create economic and social value, ensures equitable access to that value, and fosters trust among participants. At a national level, governments should engage in dialogue with individuals, civil society, academia, and the private sector to develop rules for the safe use of data

that promotes the public good; while at the international level, we need closer international cooperation to harmonize regulations and coordinate policies (WDR , 2021).

A social contract isn't a new idea. There are already laws that govern data creation and protection, such as the EU's GDPR, rooted in the US Fair Information Principles developed in the 1970s. But the social contract for data set out in the WDR 2021 goes further. Different data-creating and data-using parties need to collaborate and have a way to do so safely. Gathering personal data from individuals also requires their cooperation, so we need to build a trusting environment. The more people trust the system, the more likely they will share personal information, and the more useful the data become. Such a contract would enable the use and reuse of data to create economic and social value while ensuring equitable access to the value realized, as well as fostering participants' trust that they will not be harmed by data misuse. Renewed efforts are required to improve data governance domestically, as well as through closer international cooperation.

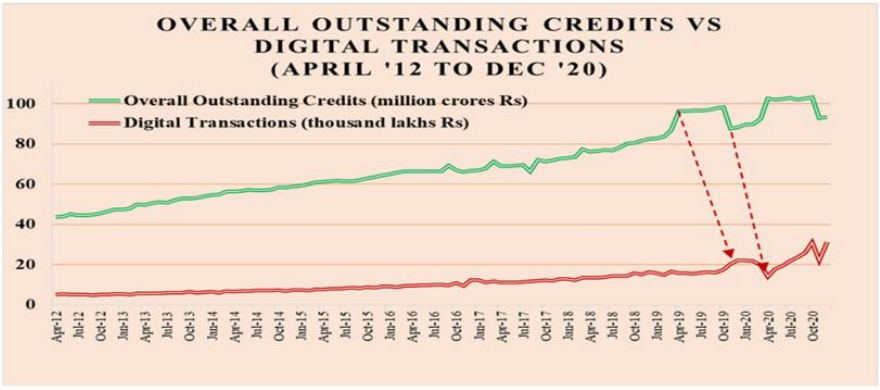
5. Illustrative India Policy Use case:

Empowerment has become a central theme of many programs and schemes across governments aimed to development. Opportunities lead to more choices. Empowerment can be said as a multi-dimensional concept and implies wherein people have freedom of expression and of exercising her choice to shape her life in all spheres of relevance. Economic freedom, freedom to transact and participate in economy. (Sen, 81; Mandal, 2013). Different studies have measured empowerment differently, one measures actively used labour force participation and financial wellbeing, (Assad et all, 2014, *Haque, Zulfiqar, 2016*). A strong correlation exists between empowerment and development of economy. Higher participation in workforce, leads to higher is the economic development (Duflo, 2012). Labour markets in developed and developing markets are strikingly different. Developed markets are more non-agricultural sector while emerging markets are more agrarian in nature. In developed economies vast majority of workers are formal wage and salary earners, while in emerging market it is of self-employment (Rustagi,2011).

There are some challenges to foster entrepreneurship in developing countries. Citizens are more constrained in economic initiatives by access to credit and infrastructure (Banerjee and Duflo

2007). Agency theory posits credit access is constrained by lack of information or rather asymmetric information availability coupled with perceived conflicting interest between financial institute and lender (Berger et al.,2001; Mason and Stark,2004). Lending based on transactional data is limited (Altman et al.,2020; Van Caneghem &Van Campenhout,2012). Even in the Indian context, it can be very seen, there is a granger causality between digital transaction credit access (Chi Value 6.77 and P value 0.034).

Figure 5 Data Analysis Impact on Policy Framework



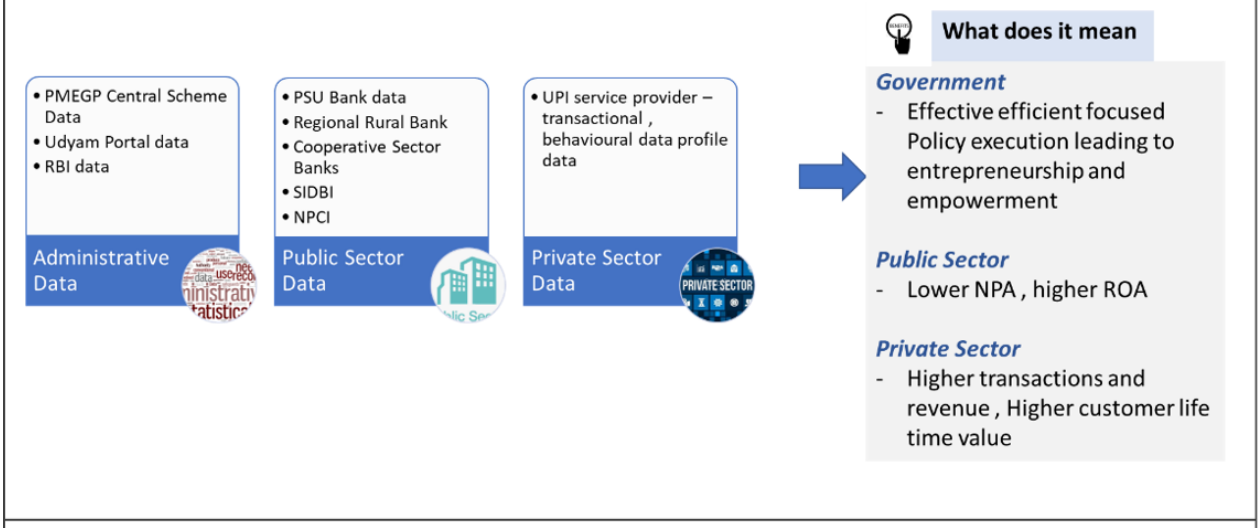
Source : RBI, NPCI data and Authors analysis

Digital payments play a significant and important role in providing avenues to transact, (Pal et all, 2020), which leads to "economic facility" required for development (Jacob 2016, Sen 2001, Pal, 2020). To have access to credit, which leads to economic wellbeing, one of the key drivers is Digital transactions, which in turn is a function of infrastructure, ecosystem, financial literacy, and digital financial literacy. Clearly digital transactions are a lead indicator of impending credit access. Digital and mobile payments is also leads to social and health empowerment, which again leads to economic empowerment.

Prime Minister's Employment Generation Programme (PMEGP) was launched in 2008 by merging the two schemes that were in operation till 31.03.2008 namely Prime Minister's Rojgar Yojana (PMRY) and Rural Employment Generation Programme (REGP). The objective of the scheme is to generate employment opportunities in rural as well as urban areas of the country

through setting up of new self-employment ventures/projects/micro enterprises. The Scheme envisages credit linked subsidy to setup and upgrade new micro enterprises in non-farm sector. Total actual expenditure for 2020-21 was INR 1905Cr.

Figure 6 Data Use and Benefits
Prime Minister's Employment Generation Programme



Source : Author's creation

Implementation of Prime Minister's Employment Generation programme, which aids micro entrepreneurs' success; can be optimized by looking at digital payment data as well as mobility data. A confluence of administrative data, public sector data and private sector data will help us focus appropriate execution by prioritizing districts, sub- segments (segment profile) and financial institutions in that area. These exercises will not only help in accelerating growth, but also assist the financial institutions achieve appropriate level of profitability. It can provide, government as well as NGO to focus on areas of capacity building to ensure success of the intervention scheme.

6. Conclusion

There are complex issues to be considered when Big Data is studied concerning the public sector. This paper focuses on efficiently and effective use of Big Data for public policymaking. It attempts to develop a holistic and systematic approach in leveraging big data technology for evidence policy making. To have access as well as the ability to process data, it is imperative for

a national framework, which alleviates risks as well as protects citizens from infringement of personal privacy. Supporting appropriate technology infrastructure, appropriate institutional and regulatory frameworks need to be created trust in data systems.

The proposed approach can be said as a starting point to build a detailed framework, which is not just comprehensive but also flexible and agile. Future research can investigate other factors that may influence the effectiveness of big data in government and incorporate it into the framework. Capital, land, and labor are inputs to development, but data allow us to make these inputs focused and drive efficiency. However, unlike capital, land, and labor, using data once does not diminish its value, it can be repurposed creating a higher value.

"Data is the new Sun"

References

References

Altman, E., Esentato, M., & Sabato, G. (2020). Assessing the credit worthiness of Italian SMEs and mini-bond issuers. *Global Finance J*. 43, 100450.

Assad Ragui, Rana Hendy, Moundir Lassassi and Chaimaa Yassine (2014). Where Has All the Education Gone?

Analyzing Trends in Labor Force Participation in MENA. Paper prepared for the 22nd Annual Conference of the Economic Research Forum Cairo, Egypt

Banerjee, A., & Duflo, E. (2007). The Economic Lives of the Poor. Journal of Economic Perspectives. 21. 10.1257/jep.21.1.141.

Berger, A.N., Klapper, L.F., & Udell, G.F. (2001). The ability of banks to lend to informationally opaque small business. Journal of Banking & Finance 25(12): 2127–2167.

Birckland, T. (2001). An Introduction to the Policy Process: Theories, Concepts, and Models of Public Policy Making. New York, NY: ME Shape.

Chen, C. L. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information Sciences, 275 (10 August 2014): 314-347.

Chen, M., Mao, S., & Liu, Yunhao (2014). Big Data: A Survey. Mobile Networks and Applications, 19 (2): 171-209.

Clarke, A., & Margetts, H. (2014). Governments and citizens getting to know each other? Open, closed, and big data in public management reform. Policy & Internet 6(4): 393–417.

Cook, T. (2014). ‘‘Big data’’ in research on social policy. Journal of Policy Analysis and Management 33(2): 544–547.

Desouza, K., & Bhagwatwar, A. (2012). Leveraging technologies in public agencies: The case of the U.S. Census Bureau and the 2010 Census. Public Administration Review 72(4).

Desouza, K. & Jacob, B. (2017). Big data in the public sector: Lessons for practitioners and scholars. Administration & Society 49(7): 1043–1064.

Donner, J., & C.A. Tellez. (2008). Mobile banking and economic development: Linking adoption, impact, and use. Asian Journal of Communication 18 (4):318-332.

Duflo, E. (2012). "Women Empowerment and Economic Development." Journal of Economic Literature, 50 (4): 1051-79.

Dunleavy, P. (2016). “Big Data” and Policy Learning. In G. Stoker, & M. Evans (Eds.), Evidence-Based Policy Making in the Social Sciences: Methods that Matter (1 ed., pp. 143-168). (Evidence-Based Policy Making in the Social Sciences: Methods that Matter). Policy Press. https://doi.org/10.2307/j.ctt1t89d4k.15.

Gamage, P. (2016). New development: Leveraging ‘big data’ analytics in the public sector. Public Money & Management 36(5): 385–390.

Haque, A., & Zulfiqar, M. (2016). Women's Economic empowerment through Financial Literacy, Financial Attitude and Financial Wellbeing. International Journal of Business and Social Science, 7(3), pp. 78-88.

Höchtl, J., Parycek, P., & Schöllhammer, R. (2016). Big data in the policy cycle: Policy decision making in the digital era. Journal of Organizational Computing and Electronic Commerce 26(1–2): 147–169.

Huang, T., Lan, L., Fang, X., An, P., Min, J., & Wang, F. (2015). Promises and Challenges of Big Data Computing in Health Sciences. Special Issue on Computation, Business, and Health Science, 2 (1): 2-11.

IDC (2012). The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East.

Janssen, M., & Kuk, G. (2016). The challenges and limits of big data algorithms in technocratic governance. Government Information Quarterly, 33(3): 371–377.

Jordan, S. R. (2014). Beneficence and the Expert Bureaucracy. Public Integrity, 16 (4): 375-394.

Kernaghan, K. (2014). Digital dilemmas: Values, ethics and information technology. Canadian Public Administration, 57 (2): 295-317.

Kitchin, R., & Lauriault, T. P. (2014). Small data in the era of big data. GeoJournal, 80 (4): 1-13.

Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity and Variety. META Group Research Note, 6: 70.

Loftis, M. & Mortensen, P. (2018). Collaborating with the machines: A hybrid method for classifying policy documents. The Policy Studies Journal. DOI: 10.1111/psj.12245.

Maciejewski, M. (2017). To do more, better, faster and more cheaply: Using big data in public administration. International Review of Administrative Sciences 83(1): 120–135.

Mandal K C. (2013). Concept and Types of Women Empowerment. international Forum of Teaching and Studies Vol. 9 No. 2: 17-30

Manzoni, J. (2018). Civil service transformation speech. London. 24 January 2018. Available at: www.gov.uk/government/speeches/civil-service-transformation-speech (accessed 2 February 2022).

Mason, C., & Stark, M. (2004). What do investors look for in a business plan? A comparison of the investment criteria of bankers, venture capitalists and business angels. International Small Business Journal, 22(3), 227– 248. https://doi.org/10.1177/0266242604042377

Mergel, I. (2017). Building holistic evidence for social media impact. Public Administration Review 77(4): 489–495.

Nowlin, M. (2016). Modeling issue definitions using quantitative text analysis. The Policy Studies Journal 44(3): 309–331.

Pal, A., De, R., Herath, T., & Rao, H. R. (2019). A review of contextual factors affecting mobile payment adoption and use. Journal of Banking and Financial Technology, 3(1), 43–57.

Pal, A., De, R., Herath, T. (2020). The Role of Mobile Payment Technology in Sustainable and Human-Centric Development: Evidence from the Post-Demonetization Period in India. Information Systems Frontiers (2020) 22:607–631

Pencheva, I., Esteve, M., & Mikhaylov, S. (2020). Big Data and AI – A transformational shift for government: So, what next for research?. Public Policy and Administration 35(1) 24–44.

Sen, A. K. (1981). Poverty and Famines: An Essay on Entitlement and Deprivation. Oxford: Clarendon.

Sen, A. (2001). Development as Freedom (New Ed edition ed.). OUP Oxford, Oxford.

Siegler, M. G. (2010). 'Eric Schmidt: Every 2 Days We Create As Much Information As We Did Up To 2003.' TechCrunch. http://tcrn.ch/1h55bnj (Accessed 2 February 2022).

Smith, G., Noble, M., Antilla C., Gill, L., Zaida, A., Wright, G., Dibben C., & Barnes, H. (2004). The Value of Linked Administrative Records for Longitudinal Analysis. (London: Economic and Social Research Council). Report to the ESRC National Longitudinal Strategy Committee.

Stritch, J., Pedersen, M., & Taggart, G. (2017). The opportunities and limitations of using Mechanical Turk (MTURK) in public administration and management scholarship. International Public Management Journal 20(3): 489–511.

Taeihagh, A. (2017). Crowdsourcing: A new tool for policy-making? Policy Sciences 50(4): 629–647.

Van Caneghem, T., Van Campenhout, G. (2012). Quantity and quality of information and SME financial structure. Small Bus. Econ. 39, 341–358.

Wegrich, K., & Jann, W. (2006). Chapter 4: Theories of the policy cycle. In: Fischer F, Miller GJ and Sidney MJ (eds) Handbook of Public Policy Analysis. Theory, Politics, and Methods. Boca Raton: CRC Press, pp. 43–62.

White, M. (2012). Digital Workplaces Vision and Reality. Business Information Review, 29 (4): 205-214.

Williamson, A. (2014). Big data and the implications for government. Legal Information Management 14(4): 253–257.

World Bank. (2021). World Development Report 2021: Data for Better Lives. Washington, DC: World Bank.