Big Data and SDGs: The State of Play in Sri Lanka and India

This work was carried out with the aid of a grant from the International Development Research Centre (IDRC), Canada.

LIRNEasia
Pro-poor. Pro-market.

Acknowledgements

LIRNEasia and the Centre for Internet & Society prepared the case studies for Sri Lanka and India respectively. Thavisha Perera-Gomez and Sriganesh Lokanathan developed the Sri Lanka case study, with support from Lasantha Fernando, whose write-up on the organization's ongoing work on disease propagation has been incorporated into the Sri Lanka case study. The authors gratefully acknowledge the time spent by Prof. Rohan Samarajiva of LIRNE*asia* in reviewing the early draft of the Sri Lanka case study. Insights provided by Ms. Champika De Silva, Deputy Director, Department of Census and Statistics, Sri Lanka were particularly useful. The Centre for Internet & Society developed the India case study. The team comprised of Sumandro Chattapadhyay, Amber Sinha, Luca Brennecke, Vidushi Marda, Leilah Elmokadem and Vanya Rakesh.

LIRNE*asia* is a pro-poor, pro-market think tank whose mission is *Catalyzing policy change through research to improve people's lives in the emerging Asia Pacific by facilitating their use of hard and soft infrastructures through the use of knowledge, information and technology*.
Contact: 12 Balcombe Place, Colombo 00800, Sri Lanka. +94 11 267 1160. info@lirneasia.net
www.lirneasia.net

The Centre for Internet and Society (CIS) is a non-profit organisation that undertakes interdisciplinary research on internet and digital technologies from policy and academic perspectives. The areas of focus include digital accessibility for persons with disabilities, access to knowledge, intellectual property rights, openness (including open data, free and open source software, open standards, open access, open educational resources, and open video), internet governance, telecommunication reform, digital privacy, and cyber-security. Contact:194, 2nd 'C' Cross, Domlur, 2nd Stage, Bengaluru, 560071, India. 080 4092 6283. sunil@cis-india.org

This work was carried out with the aid of a grant from the International Development Research Centre (IDRC), Canada.

Table of Contents

Big Data and SDGs: State of Play in Sri Lanka

Background

In September 2015, the UN General Assembly adopted 17 goals as part of a new sustainable development agenda, each with specific targets to be achieved by 2030. Critically, the agenda calls for a focus on data so as to ensure that progress (or lack thereof) is known and remedial action can be taken. Encompassing the economic, environmental and social dimensions of development, the Sustainable Development Goals (SDGs) present ambitious targets for Sri Lanka. The data requirements for the SDGs exceed those of the Millennium Development Goals in terms of the sheer number of indicators. While Sri Lanka appears to have a relatively strong statistical capacity relative to the South Asian region,[1] there will still be challenges in terms of providing accurate, timely and disaggregated data for measuring progress toward SDGs. At the ninth session of the High Level Political Forum (HLFP) in July 2016, it was noted that Sri Lanka would use a set of nationally developed indicators, standards and guidelines to report on national sustainable development.[2]

Aligning SDGs to the National Agenda

Sri Lanka has already begun numerous initiatives to better align the country's policies with the SDGs. For instance, in 2015, Sri Lanka became the first country in the Asia Pacific region to establish a Ministry for Sustainable Development.[3] The Sustainable Development Division of the Ministry of Sustainable Development and Wildlife Conservation will function as the central coordinating agency for SDGs.[4] It is also tasked with liaising with sectorial agencies to implement national policies on sustainable development, facilitate research on areas related to sustainable development.

The country has further pledged to ensure that the SDGs would be aligned to the national development agenda. Based on a preliminary mapping exercise conducted by the government, it appears that nearly three-fourths (~74%) of the targets are either fully or moderately aligned.[5] Five key departments are slated to play a key role in ensuring that the SDGs are integrated into the National Development Agenda: the Department of National Planning, Department of Census & Statistics, Department of Project Management and Monitoring, Department of National Budget and the Department of External Resources, with overall coordination provided by the Ministry of Sustainable Development and Wildlife Conservation, which would also coordinate with other stakeholders such as civil society, private sector, other government bodies, and academia among others.[6] Moreover, as the focal point for the SDGs in Sri Lanka, the Ministry of Sustainable Development and Wildlife Conservation is developing sustainable development guidelines for public sector agencies and to that end has leveraged a composite index called the Sustainability Compliance Standards Framework which is able to "is able to assess the strategic commitment to sustainable development of a project through a set of 40 performance standards and indicators on Economic, Environmental, Social Sustainability and Good Governance standards".[7]

---

[1] http://www.unescap.org/sites/default/files/SDGs%20South%20Asia%20report%202016%20rev%2014%20April%202016.pdf page 21
[2] https://sustainabledevelopment.un.org/content/documents/21634sri-lanka%20(1).pdf
[3] https://sustainabledevelopment.un.org/content/documents/21931srilanka.pdf
[4] http://msdw.gov.lk/divisions/sustainable-development-division/ ; http://msdw.gov.lk/news/national-sustainable-development-engagement-platform-was-held-on-5th-
[5] http://www.unsiap.or.jp/e-learning/1_sdg/1609_SDG_india/SRI-LANKA_SDG_Development%20Plan.pdf ; http://www.unsiap.or.jp/e-learning/1_sdg/1609_SDG_IND.html
[6] http://www.unsiap.or.jp/e-learning/1_sdg/1609_SDG_india/SRI-LANKA_SDG_Development%20Plan.pdf
[7] https://sustainabledevelopment.un.org/content/documents/21650sri-lanka3.pdf

In early 2017, the President of Sri Lanka unveiled an initiative titled, Sustainable Era to direct Sri Lanka's national sustainable path. A 31 person expert commission that includes Prof. Mohan Munasinghe and Prof. Rohan Samarajiva (of LIRNE*asia*) has been set up to draft a framework for this initiative.[8]

Mapping Institutional Inter-linkages

The government recognizes that the achievement of the SDGs would involve the participation of various actors including the government, private sector, civil society, and citizens among others. Even within government, the SDGs would require participation by 51 ministries and 425 line agencies.[9] The Ministry of Sustainable Development and Wildlife Conservation has conducted a preliminary mapping exercise to identify the linkages between various government ministries and departments to meet specific SDGs. For instance, in order to address the 7 targets of Goal 1 (No poverty), the engagement of 27 ministries and 63 agencies would be needed. Similarly, the coordination between 16 ministries and 38 agencies would be required to address the 10 targets of Goal 14 (Life below water), while SDG 6 (Water and Sanitation) would require inter-linkage between 24 ministries and 43 agencies[10]. Sri Lanka is also among the first countries to pilot the ESCAP analytical framework,[11] developed by UN ESCAP for the analysis of the inter-linkages and the interdependencies between each of the 17 SDGs with SDG 6 (Water and Sanitation) as the core.

SDG Implementation

Sri Lanka is in the process of developing a Sustainability Development Act that will provide the institutional and legal framework to support the implementation of the SDGs.[12] The government has proposed a three-phase approach for implementing the SDGs in Sri Lanka,[13] with the first phase (2016-2020) focusing on the establishment of legislative, institutional and policy frameworks, the next phase (2021-2025) focusing on infrastructure investment and the final phase (2026-2030) dealing with the establishment and implementation of relevant sustainable systems.

From a governance perspective, this includes the above-mentioned Sustainable Development Act, the formulation of a sustainable development policy and the establishment of relevant standards and guidelines. The establishment of a cabinet ministry that focused on sustainable development was also be part of this. In terms of engagement, the government proposes the establishment of platforms for engagement at both the national level and provincial level as well as the translation of national plans into provincial level sustainability plans. From an implementation perspective, the government proposes a roadmap that incorporates a strong monitoring and reporting system, national vision, as well as a mapping of the system linkages between the numerous institutions.

SDG Priorities for Sri Lanka

The President of Sri Lanka has stated that Sri Lanka will pay particular attention to the achievement of six goals: SDG 1 (no poverty), SDG 2 (food security), SDG 7 (energy), SDG 4 (education), SDG 10 (reducing income disparity) and SDG 11 (infrastructure and industry). Of these, particular attention will be placed on SDG 1, with the Sri

[8] http://dailynews.lk/2017/01/09/features/104139/making-development-more-sustainable
[9] https://sustainabledevelopment.un.org/content/documents/22787Uchita%20de%20Zoysa%20(01%20Dec).pdf
[10] http://www.unescap.org/sites/default/files/Session%203_%20Uchita%20de%20Zoysa_Sri%20Lanka.pdf
[11] http://www.unescap.org/sites/default/files/Integrated%20Approaches%20for%20SDG6_SDG%20Week_Concept%20Note__5%20Nov%202016_Web.pdf
[12] http://dailynews.lk/2017/01/10/local/104309/sri-lanka-sustainable-development-bill-presented-house
[13] https://sustainabledevelopment.un.org/content/documents/22787Uchita%20de%20Zoysa%20(01%20Dec).pdf

Lankan government declaring 2017 as the year of alleviating poverty. The government has also stated their goals of higher levels of benchmarking than the SDG requirement for some targets (for example, $1.25 for poverty eradication) in order to achieve national objectives for wellbeing.[14]

Data for SDG

The Department of Census and Statistics, which has been designated as the coordinating body for the compilation of SDG indicators for Sri Lanka, plays a crucial role in measuring/monitoring the SDGs.[15] While the key responsibility of integrating the SDGs to the country's National Development Framework lies with the Department of National Planning, as of September 2016, it is the Department of Census and Statistics (DCS) that is involved in developing the indicators and undertaking data collection. To that end, the DCS has formed an internal committee to evaluate the SDGs and identify data gaps. Based on initial findings from a gap analysis, as of September 2016, there was a 65% indicator gap, and only 25% of the indicators can be fully measured using available data.[16] A more in-depth exercise is ongoing. Data available from the DCS based on the most recent round of surveys is for the following targets[17]:

Table 01: Data Availability by Type of DCS Survey

Survey	Targets
Household and Income and Expenditure Survey	1.1, 1.2, 1.3, 1.4, 5b, 7.1, 10.1, 10.2
Labour Force Survey	2.1, 2.2, 2.3, 3.1, 3.7, 5b, 5.2, 5.3, 5.5, 5.6.1, 6.2, 8.3, 8.5.1, 8.5.2, 8.6, 9.2
Child Activity Survey	8.7, 8.8.1
Demographic and Health Survey	2.1.1, 2.1.2, 2.2.2, 3.1.1, 3.1.2, 3.7.1, 3.7.2, 5.2, 5.3.1, 5.6.1, 6.2.1
Annual Survey of Industries	2.3.2, 9.3.1, 9.3.2

Source: *http://www.unsiap.or.jp/e-learning/1_sdg/1609_SDG_india/SRI-LANKA_SDG_Development%20Plan.pdf*

The future plans for the DCS include developing platforms for the dissemination of SDG data, for example, an SDG webpage, and building in new components in existing surveys to capture new data.

Big Data applications in Sri Lanka

The developments within the SDG space in Sri Lanka highlight the importance of accurate and timely data to measure and monitor the SDGs. For instance, in order for the government to achieve its goal of alleviating poverty, it is imperative to have access to disaggregated poverty data that can be compared over time, and are available at relatively frequent intervals at present. However, at present, the DCS provides poverty estimation at a district level based on data collected from the Household Income and Expenditure Survey that is conducted over a period of 12 consecutive months[18] every 5 years, as per the DCS. [19] In order to obtain data at a more granular level, the World Bank and the DCS conducted two poverty mapping exercises (poverty maps for 2002 and 2012) to estimate poverty at disaggregate administrative levels using small area estimation method – a

[14] https://sustainabledevelopment.un.org/content/documents/21634sri-lanka%20(1).pdf
[15] http://www.unescap.org/sites/default/files/11_Sri_Lanka_National_Presentation-Diagnostic_Tool.pdf
[16] http://www.unsiap.or.jp/e-learning/1_sdg/1609_SDG_india/SRI-LANKA_SDG_Development%20Plan.pdf
[17] http://www.unsiap.or.jp/e-learning/1_sdg/1609_SDG_india/SRI-LANKA_SDG_Development%20Plan.pdf
[18] http://www.statistics.gov.lk/HIES/HIES2012_13FinalReport.pdf
[19] http://www.statistics.gov.lk/page.asp?page=Poverty

method that leveraged data from the population census (every 10 years)[20] and combined it with household survey data (every 5 years) to estimate household expenditure for smaller geographic areas. [21].

One avenue that the government has identified to bridge gaps in data for the SDGs is the possibility of leveraging non-traditional sources such as big data.[22] However, this is still very much in an embryonic stage and the indicators for which big data can play a role have not as yet been identified.[23] Such non- traditional sources of data have the potential to complement traditional statistics in the interval between official surveys, and develop new ways of monitoring the SDG targets. Insights derived from big data sources may help to fill the gaps relating to poverty data, rather than replace official surveys. For instance, mobile network big data can support this goal by helping to identify the poor--determine the socioeconomic status of the population, identify pockets of urban poverty, and estimate poverty rates--and by creating opportunities for greater financial inclusion among the poor. For instance, a recent study conducted in Bangladesh (Steele, et al. 2017), leveraged mobile phone data and satellite-based data to produce poverty predictions.[24] The researchers generated three maps that provided national estimates for household income, wealth index and progress out of poverty index.

There have been numerous efforts undertaken in Sri Lanka or on Sri Lanka using non-traditional data sources such as mobile network data and satellite imagery that have the potential to be leveraged to measure the progress of SDGs.

Goal 1: No Poverty

Engstrom, Hersh and Newhouse (2016) conducted a World Bank funded study in Sri Lanka using very high-resolution satellite images (VHRSI) to estimate variation in poverty across small local areas. The researchers analysed features such as density of paved and unpaved roads, building density, roof types, farmland types as shadow pixels (as a proxy for building height) as well as numerous spectral and texture features among others. The study covered 1,250 of the 13,000 Grama Niladhari divisions in Sri Lanka – the average size of the division was around 2.15 sq. km and contains around 10,000 persons.[25] The results were matched with poverty data based on the 2011 census and the researchers found that around 40%-70% of small area poverty variations could be explained using features derived from VRHSI.

Goal 3: Good Health and Wellbeing

LIRNEasia is currently collaborating with the Epidemiology Unit, Ministry of Health and University of Moratuwa to develop spatio-temporal prediction models that can forecast dengue outbreaks two to three weeks ahead of the actual incidence. It has been established that human mobility plays a significant role in the spread of dengue virus, but there had been very little success when attempting to incorporate human mobility to existing forecasting models using traditional data sources (Sarzynska, Udiani, & Zhang, 2013; Wesolowski et al., 2014)[26].

---

[20] http://www.statistics.gov.lk/PopHouSat/PDF/p7%20population%20and%20Housing%20Text-11-12-06.pdf

[21] http://www.statistics.gov.lk/page.asp?page=Poverty

[22] http://www.unsiap.or.jp/e-learning/1_sdg/1609_SDG_india/SRI-LANKA_SDG_Development%20Plan.pdf

[23] Based on call with Ms. Champika De Silva, Deputy Director, DCS

[24] https://phys.org/news/2017-02-mobile-satellite-poverty.html

[25] http://pubdocs.worldbank.org/en/594931466434554022/Poverty-in-HD-ABCDE-Presentation-v2-1-Hersh.pdf; http://pubdocs.worldbank.org/en/607414661817437796/Poverty-in-HD-draft-v2-75.pdf; https://medium.com/from-the-macroscope/mapping-poverty-from-space-with-the-world-bank-5363c1a2b5d2#.sqlwnd9fq

[26] 1. Sarzynska, M., Udiani, O., & Zhang, N. (2013). A study of gravity-linked metapopulation models for the spatial spread of dengue fever. *arXiv Preprint arXiv:1308.4589*, *2008*, 1–32. Retrieved from http://arxiv.org/abs/1308.4589

2. Wesolowski, A., Stresman, G., Eagle, N., Stevenson, J., Owaga, C., Marube, E., … Buckee, C. O. (2014). Quantifying travel behavior for infectious disease research: a comparison of data from surveys and mobile phones. *Scientific Reports*, *4*, 5678. https://doi.org/10.1038/srep05678

In the collaborative study, macro level proxy indicators for human mobility are derived by using Big Data techniques on mobile network call detail records (CDR). These derived mobility indicators are incorporated into predictive models that are based on statistical as well as machine learning techniques. Initial results show significant correlation between mobility indicators and dengue incidence and measured improvement in prediction accuracy as well.

Goal 11: Sustainable Cities

Similarly, in 2016 Facebook used a computer vision approach to develop high- resolution population maps based on satellite imagery for 20[27] countries including Sri Lanka. The maps, jointly produced with the Center for International Earth Science Information Network, of the University of Colombia provide information on population distribution at a spatial resolution of 10 meters. The population maps offer the opportunity to better understand the distribution of settlements in Sri Lanka and presents implications for the development of infrastructure and understanding exposure to hazards.[28]

Moreover, according to information on the UN big data project inventory, the World Bank has provided the government of Sri Lanka a grant to equip them with relevant technology that would enable them to update maps of Sri Lanka using geo-spatial data that is crowd sourced.[29]

Furthermore, insights derived from LIRNE*asia*'s analysis of mobile network data were used by the Western Region Megapolis Planning Project, which is a major focus of development for the Sri Lankan government. Researchers analyzed the call detail records of numerous mobile operators to infer patterns of diurnal mobility of populations. The insights derived were used as inputs to support the case for alleviating congestion in Colombo and develop peripheral regions that that saw high levels of population travel towards Colombo.

Big data for development landscape in Sri Lanka

The term big data has been gaining increasing traction in Sri Lanka—various universities offer data science as a study stream, and private sector players have also been foraying into the big data analytics space. However, in terms of the development landscape, the only player that is engaging with big data appears to be non-profit think tank LIRNE*asia*.

Since 2012, LIRNEasia has conducted research on leveraging new data sources and in particular big data (including mobile network big data) for public purposes in Sri Lanka. LIRNEasia participates in the global dialogue on big data for development with experiential evidence and are one of the few (if not the only) institutes located in the Global South that conceptualizes and implements their own policy-relevant big data analyses and research. LIRNEasia negotiated access to pseudonymized, historical call detail records from multiple mobile operators along with other data sources (satellite imagery, CCTV footage, electricity data, social media data, and official statistics) to conduct analyses on human mobility and traffic, land-use patterns, communities, high frequency and high-resolution proxy indicators of economic activity.

Conclusion

While the Sri Lankan government appears to have taken its commitment to achieve the SDGs seriously, the labyrinth of institutional inter-linkages has added a layer of complexity to this. When considering the unprecedented data requirement on countries to report timely, accurate and disaggregated data across a vast

[27] https://www.technologyreview.com/s/600852/facebooks-new-map-of-world-population-could-help-get-billions-online/
[28] https://phys.org/news/2016-02-facebook-population.html
[29] https://unstats.un.org/bigdata/inventory/?selectID=WB54

range of indicators, government departments have identified big data as a possible to tool to address data gaps in the traditional system. However, there has been very limited traction in exploring the use of big data in measuring the progress towards SDGs.

The ambiguity of the data landscape underscores the need for third party providers outside the government system to push the agenda for the use of non-traditional data sources to support the measurement of the SDGs. There is a need to raise awareness of the potential of big data for public purposes, and invest in institutional capacity building.

In summary, there is a need for supply push of big data solutions in Sri Lanka before there is demand-pull. This presents opportunity for a middleman: to gain access to (privately held) data, conduct the analyses and share insights with government institutions that may not have the relevant capacity to conduct in-house big data analyses. This can illustrated by LIRNE*asia*'s experience with the government –when LIRNE*asia* began conducting research on big data for development in 2012, there was no demand for insights. However, had they not started on the research when they did, they would not have been able to capitalize on the opportunity when the policy window opened.

Sri Lanka can also tap into the increasing global dialogue on big data in statistics, as evidenced by conferences such as the UN big data for official statistics and discussions of data innovations at the UN World Data Forum.

Big Data and SDGs: State of Play in India

Introduction

This paper seeks to reflect on literature and developments surrounding big data and development in India, specifically in context of Sustainable Development Goals (SDGs). The SDGs are an internationally adopted set of developmental targets to be achieved by 2030, with each of these 17 SDGs being comprised of various targets (169 in total).[30] Big data has the potential to aid public health, guide better decision-making, encourage innovation and enable information sharing and knowledge creation at a large scale.[31] By mapping academic literature and new developments onto each SDG, this paper aims to demonstrate the discourse surrounding how big data can be leveraged to achieve each individual goal.

In India, the Department of Science and Technology (DST) called for the preparation of a "Strategic Road Map for Big Data Analytics" in 2014. This marked the first explicit and strategic government measure in this regard. While India has not adopted a comprehensive strategy or policy to leverage big data for development purposes, it is useful to collate and analyse academic research, news items, press releases etc. To understand the present state of research around big data for development as well as emerging and ongoing initiatives and what this indicates with respect to the potential of India leveraging big data to attain the SDGs.

For the purposes of this report, we have analysed academic articles, government reports, and reports from consultancies and industry, along with news reports that provide information on the latest developments with respect to big data and development. This literature has been understood in terms of each SDG, in order to map the discourse surrounding each goal.
We initially relied on publicly available resources. However, as a note - much of the existing literature appears to be only available behind pay walled databases.

Overarching observations

Existing literature around big data and development is largely optimistic about the government's potential to harness data and solve problems by implementing practices, which are streamlined with big data analysis. While implementing big data for development, multiple scholars opine that priority should be on better delivery of services, efficient government operations, and increased collaboration between government agencies.[32] The main challenges to the usage of big data for development by government agencies are related to: privacy and security; maintaining a database that is usable, accessible, accurate, and discoverable; and technical and capacity challenges.[33] According to some scholars, big data for government can be understood as a combination

[30] Transforming our world: the 2030 Agenda for Sustainable Development, https://sustainabledevelopment.un.org/post2015/transformingourworld.

[31] Reality Mining: Using Big Data to Engineer a Better World, Nathan Eagle and Kate Greene, 2014, Massachusetts Institute of Technology.

[32] Big Data and the Opportunities and Challenges for Government Agencies, https://pdfs.semanticscholar.org/c74f/f91acfdea6724150d8c09351d40ae54270a6.pdf; Big data framework for national e-governance plan, http://ieeexplore.ieee.org/abstract/document/6756283/.

[33] Big Data and the Opportunities and Challenges for Government Agencies, https://pdfs.semanticscholar.org/c74f/f91acfdea6724150d8c09351d40ae54270a6.pdf.

of four components: resource management, data organization and management, analytics and discovery, and finally decision and visualization support.[34]

Some authors have pointed out the need for corporate strategies and government policies to be well formulated before big data can be effectively utilised. In India specifically, there are a number of challenges to be overcome before big data can be used for development. Some of these are the lack of data collection and storage in India, the lack of accessibility to platforms that hold big data, and the lack of a legal framework and policies that cover issues like responsibility for collection, storage and preservation of data.[35]

Key Learnings

Through a literature survey and review of each SDG in context of governmental schemes, these are the main learnings we have had:

1. There is a lack of robust discussion around *legal* reform, with most thinkers and scholars looking towards the possible applications of big data in governance in India. The literature reviewed reflected a forward-looking, application-based analysis as opposed to an analysis based on rights and regulation.
2. The discourse around big data in development is also strikingly confined to sectors and expertise, with each discipline focusing solely on issues that crop up within their immediate scope. For example, technical scholarship largely fails to address or consider the rights framework within which big data operates and vice versa.
3. Similarly, there is very little consideration on how regulatory frameworks need to evolve in parallel with big data in development, with scholarship being based on existing frameworks but not leaning towards a possible shift or change in them.
4. While discussing big data analytics, there seems to be a dominant, default turn towards Hadoop, it will be useful to do further research on alternative analytic solutions and accessibility of the same.
5. Literature and policy are by far the most mature with respect to Health in India, with not only practical implementation but also substantial, forward-looking scholarship on frameworks, models, and possible projects. The discourse is also mature around agriculture and poverty alleviation, the latter in particular was significantly boosted by the Fintech revolution in India.
6. Much of the literature looks at convergence of data sets, structuring unstructured data, and the opportunities that will arise from such convergence. While some academics made a reference to concerns surrounding privacy and surveillance, the literature was largely not concerned with privacy implications.
7. An overarching concern through all the literature analysed is the lack of technical capacity to fully utilise the potential of big data. This is both within industry and government.
8. Initially, we tried to refer to only publicly available sources to scope literature; however, it became clear through the course of research that a large part of existing literature is only available through pay wall databases.
9. In the initial stages of research, we tried to carry out research goal-wise, however, we soon recognized that literature often touches upon multiple goals, and initiatives sometimes achieve more than a single goal. Classification was then carried out to reflect which goal and its corresponding targets are most directly concerned. The following goals did not find explicit mention in governmental policies, development schemes or academic literature. While some of them may be considered to be overlapping with other goals (for example, sustainable consumption and production patterns can be

[34] Big data framework for national e-governance plan, http://ieeexplore.ieee.org/abstract/document/6756283/.

[35] Big data: Road ahead for India, Madhukar Dayal, Sachin Garg and Rubaina Shrivastava, July - December 2014.

considered to overlap with agriculture), they have not been discussed above due to lack of emphasis and/or explicit mention. This includes Goal 5, 10, 12, 14, and 15

Goal Based Review

Goal #01: End poverty in all its forms everywhere

Practitioners in the big data and development space are hopeful of big data's potential to alleviate poverty. According to them, the collection and analysis of real time mobile phone data can highlight poverty trends, satellite imagery can provide a degree of geographic specificity to governments that will better inform decisions, and biometric data can drive efficiency in poverty alleviation programs.[36]

The push towards inclusive credit and financial access has been steadily gaining momentum since 2010. Starting with the launch of Aadhaar, the national identity number, to the development of India Stack (a set of API's that enable utilisation of digital infrastructure), to the recent demonetization policy in November 2016.

These developments have been accompanied by corresponding moves in policy and governance. In 2012, the RBI issued a directive to instruct banks to start opening Aadhaar Enabled Bank Accounts, that is, bank accounts that are linked to the UID ("Aadhaar" number) of the account holder.[37] In February 2014, the RBI suggested integration of the mobile number (of the resident / account holder) with the Aadhaar enabled Bank Accounts,[38] that would subsequently be termed 'JAM' – with J standing for 'Jan Dhan,' a scheme for no-frills accounts for rural and poorer customers, A standing for Aadhaar, and M for mobile phone number.[39] The National Payments Corporation of India has recently unveiled the Unified Payments Interface (UPI) that aims to provide a single payment routing bridge for all kinds of actors and all kinds of transactions.[40]

This emerging fintech industry in India has been projected to be a 600 billion USD industry by reports from Credit Suisse,[41] Kalaari Capital,[42] and McKinsey.[43] In February 2017, the UNESCO Chair in ICT4D held an academic-led seminar on "Affordances of Big Data for Poverty Alleviation: Evidence from India" which looked at the datafication of the public distribution system (PDS) by Aadhaar and concluded that such datafication effects the making of anti-poverty policy itself.[44]

Goal #02: End hunger, achieve food security and improved nutrition and promote sustainable agriculture

---

[36] Big data and the fight to eradicate poverty, http://insidebigdata.com/2016/05/18/big-data-and-the-fight-to-eradicate-poverty/.

[37] Financial Inclusion - Opening of Aadhaar Enabled Bank Accounts, 19 June, 2012, https://www.rbi.org.in/Scripts/NotificationUser.aspx?Id=7284&Mode=0.

[38] Report of the Technical Committee on Mobile Banking, , 7 February 2014, https://www.rbi.org.in/scripts/PublicationReportDetails.aspx?UrlPage=&ID=760.

[39] Ministry of Finance, Government of India, Economic Survey 2014-2015, Page 21-25, http://indiabudget.nic.in/es2014-15/echapter-vol1.pdf.

[40] National Payments Corporation of India, Unified Payments Interface, Background, http://www.npci.org.in/UPI_Background.aspx.

[41] India Financials Sector: Sector Review by Credit Suisse, 29 June 2016. https://www.credit-suisse.com/media/cc/docs/cn/india-digital-banking.pdf.

[42] FinTech India: Innovation for the next 400M by Balaji Srinivasa and Yash Jain, Kalaari Capital. https://www.linkedin.com/pulse/kalaari-capital-fintech-report-2016-yash-jain.

[43] How digital finance could boost growth in emerging economies - James Manyika, Susan Lund, Marc Singer, Olivia White, and Chris Berry,McKinsey and Company, September 2016. http://www.mckinsey.com/~/media/McKinsey/Global%20Themes/Employment%20and%20Growth/How%20digital%20finance%20could%20boost%20growth%20in%20emerging%20economies/MG-Digital-Finance-For-All-Full-report-September-2016.ashx.

[44] https://ict4d2004.wordpress.com/category/big-data/

The lack of reliable big data on availability of cereal, pulses, sugar etc. is considered one of the biggest threats to food security and sustainable agriculture in India.[45] At present, agricultural data is collected in a decentralised fashion, through manual collection and ad-hoc experiments to estimate crop size (such as the 500,000 crop cutting experiments that are conducted in India each year).[46] However, factors like agricultural productivity, distribution and weather advancement systems, (which would require a more formal method of data collection, storage and analysis) can help predict weather, recommend improvements to the distribution system and merge existing information to create social safety nets to combat hunger at a fraction of the cost now incurred. Agricultural productivity is also benefitted by technology such as maps, GPS enabled tractors, automated irrigation systems, satellite and drone images etc.[47] Experts have pointed out that timely access to real time data can solve various problems that presently plague India's agricultural sector. It can solve the problem of food inflation by reducing data asymmetry which results in the volatility of prices, it can reduce post harvest wastage of farm produce by monitoring quality of produce in transportation and storage, improve soil productivity and fertility by obtaining nutrient data from a combination of on field devices satellite imagery.[48]

There has been a steady development of technology that aids in achieving this goal. For example, Cropin Technology, established in August 2010, uses big-data analytics to provide insight at the level of individual farms, clusters of farms, districts, states, or even the whole of India.[49] There has also been a conscious effort to build capacity within relevant stakeholders. ICAR-National Academy of Agricultural Research Management (NAARM), Hyderabad conducted a training programme on "Big Data Analytics in Agriculture" which was supported by the Big Data Initiatives Division, Department of Science and technology. The major objectives of the training programme were to sensitize the use of emerging big data techniques and tools in agriculture and allied sectors agricultural research. Participants in the workshop included the scientists, faculty members and research managers from the institutes of the ICAR, State Agricultural Universities (SAUs), Krishi Vigyan Kendras (KVKs), other academic and research institutions, who have advanced knowledge of statistics, computers, remote sensing and GIS.[50]

Public - Private partnerships have also come into existence. In 2016, Microsoft, in partnership with the International Crops Research Institute for Semi-Arid Tropics (ICRISAT) and Andhra Pradesh Government, tested a sowing application for farmers in Andhra Pradesh, and demonstrated a 30 per cent higher average in yield per hectare.[51]

Goal #03: Ensure healthy lives and promote well-being for all at all ages

The big data and governance discourse is arguably most mature around issues of health and medical care. According to a report by the Government of Gujarat, telemedicine is expected to be $18.7 million in 2017. According to the report, the increasing budget allocation to the Ministry of Health and Family Welfare signals a corresponding increase in big data analytics.[52] The Government of India also has a telemedicine initiative in

[45] http://www.huffingtonpost.in/nidhi-nath-srinivas/digital-farm-india-how-bi_b_8278454.html.
[46] http://dsbb.imf.org/pages/SDDS/DQAFBase.aspx?ctycode=IND&catcode=NAG00.
[47] https://vibrantgujarat.com/writereaddata/images/pdf/project-profiles/IT-Services-in-Big-Data-analytics-and-its-application-in-Government.pdf.
[48] https://www.entrepreneur.com/article/283050.
[49] http://fortune.com/2014/05/30/cropping-up-on-every-farm-big-data-technology/.
[50] http://www.biovoicenews.com/india-slowly-catches-up-on-big-data-analytics-in-agriculture/.
[51] http://india.smartcitiescouncil.com/article/microsoft%E2%80%99s-intelligent-cloud-boosts-crop-yield-30-andhra.
[52] https://vibrantgujarat.com/writereaddata/images/pdf/project-profiles/IT-Services-in-Big-Data-analytics-and-its-application-in-Government.pdf.

collaboration with Apollo Hospitals, which enables people in rural areas to consult doctors online and also order generic drugs.[53]

Integration of patient data, data on the effects of drugs, medical data, research and development data and financial records can help in identifying patterns, enhanced healthcare, and more proactive treatment. Solutions of this nature will meet some important requirements: reliability and scalability in storage and processing infrastructure, search engine capabilities for retrieving posts with high availability, scalable real time store for retrieving statistics with high availability.[54] The Institute of Genomics and Integrative Biology in the Council of Scientific and Industrial Research (CSIR), is experimenting with the possibility of collecting large amounts of data by coordinating data collection efforts between hospitals and stand alone clinics. A clinic at Lakhimpur Kheri in Uttar Pradesh is one among several such centres in the country, the others being in Haryana and Hyderabad, which are equipped with equipment that analyses blood samples and heartbeat rhythms. The data is streamed into centralized servers in Delhi.[55]

Programming frameworks that support the storing and processing of large scale data sets, like Hadoop, have made data storage cheap and accessible. Scholars note that much of health data is unstructured, and Hadoop is the nucleus proposal for organizing Big Data to give solutions to the difficulty of making big data valuable for analytics reasons. The biggest challenge in healthcare is the fact that data in existing systems can't be related to each other.[56]

Various types of medical and healthcare data can be collected from a variety of sources. Some examples are clinical data including unstructured documents, prescriptions and images, day to day research publications and medical references, huge amounts of genomic data for analyzing behavior. Arogya Mobile Health Pvt. Ltd. has conducted preliminary trials in Uttarakhand to collect and scan assorted health parameters in the hope that they will throw up patterns. It essentially aims to marry the ubiquity of mobile phones with burgeoning concerns over lifestyle diseases, primary school-educated health workers will collect weight, temperature, blood pressure and electrocardiogram readings from roughly 50,000 villagers in 50 villages. The Bluetooth-enabled devices that they use will send the information to the cloud and a pre-programmed algorithm will instantly determine whether someone needs to go see a doctor.[57] KGB or Kooda, Gandagi, Badboo (garbage, dirt, bad smell) is an app developed by a team of researchers at the affordable healthcare division in the Public Health Foundation of India (PHFI). The app uses social networking platforms to address issues of sanitation and hygiene in urban areas.[58] In order to fix rural health care, some authors propose the following measures to bridge the gap between quality and affordability in government hospitals:

a. e-Health file - creation of such a file for each patient with all the information about the patient being easily retrievable and accurate.
b. Telemedicine is also prescribed as a solution to rural exclusion from healthcare facilities and resources.
c. e-Prescribe - doctors to have an electronic prescription system.

[53] https://vibrantgujarat.com/writereaddata/images/pdf/project-profiles/IT-Services-in-Big-Data-analytics-and-its-application-in-Government.pdf.
[54] Leveraging Big Data Analytics and Hadoop in Developing India's Healthcare Services, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.684.7086&rep=rep1&type=pdf.
[55] http://www.livemint.com/Companies/zomT9nDjvanuuUxE79anOL/Big-data-revolution-in-healthcare-sector.html.
[56] Leveraging Big Data Analytics and Hadoop in Developing India's Healthcare Services, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.684.7086&rep=rep1&type=pdf.
[57] http://www.livemint.com/Companies/zomT9nDjvanuuUxE79anOL/Big-data-revolution-in-healthcare-sector.html.
[58] http://www.livemint.com/Companies/zomT9nDjvanuuUxE79anOL/Big-data-revolution-in-healthcare-sector.html.

d. Electronic medical records[59]

By making use of the ever increasing amount of data, we can find new insights. This means not just mining patient records, medical images, diagnostic reports etc. but also continuous analysis of data streams. All three parties - payer, provider and pharmaceutical company - need to work collaboratively to share data/insight.[60]

Goal #04: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all

Literature surrounding big data and education largely focuses on the potential of big data to strengthen capabilities within educational institutions. According to some authors, adapting to analytics has the potential to make educational institutions more intentional and intelligent with data and evidence, thus leading to a network of relations between knowledge and skills.[61] Big data is also considered a "strategic tool" by some authors, as it enables decision making based on data that can boost the quality of education, enhance transparency within institutions and also equips institutions to understand their students better by collecting more information about them.[62] For example, since 2006, Akshara Foundation has been collecting primary data about resources, facilities, and records from 40,000 schools in the state of Karnataka, with the goal of improving the math and reading skills of students at those schools.[63]

Other initiatives include "Swayam", a Government of India MOOC platform that aims to provide quality and cost effective education to 1- million students. [64] Further, in order to mitigate negative effects of discrepancies in the teacher - student ratio, there has been a push from some state governments, like Gujarat, to embrace the potential online education, and to use services such as EkStep which looks at improving the quality of teaching, and Web Real Time Communication (WebRTC) which provides peer to peer video based sessions amongst students.[65]

Goal #06: Ensure availability and sustainable management of water and sanitation for all

The Bangalore Water Supply and Sewerage Board (BWSSB) is using Big Data and predictive analytics technology from IBM to create systems for monitoring water distribution systems. In partnership with IBM, the BWSSB has built an operational dashboard, which serves as a "command center" for managing the city's water supply networks.[66] Elsewhere, the Kerala Water Authority (KWA), Government of Kerala, India is using IBM's Analytics and Mobility solutions to analyze, monitor and manage water distribution in the city of Thiruvananthapuram.[67]

Goal #07: Ensure access to affordable, reliable, sustainable and modern energy for all

[59] Role of big data analytics in rural health care https://pdfs.semanticscholar.org/3f9e/9d8cb8a695558390dac8fa2269248eb99aa4.pdf.
[60] Role of big data analytics in rural health care https://pdfs.semanticscholar.org/3f9e/9d8cb8a695558390dac8fa2269248eb99aa4.pdf.
[61] Significance of Big Data and Analytics in Higher Education B. Tulasi
[62] The Development of Big Data and Knowledge Management for Higher Education in India, http://www.ijmrbs.com/ijmrbsadmin/upload/IJMRBS_588af035ec2a7.pdf.
[63] https://www.datanami.com/2015/03/13/how-big-data-is-helping-indian-schoolchildren/.
[64] https://vibrantgujarat.com/writereaddata/images/pdf/project-profiles/IT-Services-in-Big-Data-analytics-and-its-application-in-Government.pdf.
[65] https://vibrantgujarat.com/writereaddata/images/pdf/project-profiles/IT-Services-in-Big-Data-analytics-and-its-application-in-Government.pdf.
[66] http://cio.economictimes.indiatimes.com/news/case-studies/bengaluru-uses-big-data-analytics-to-check-unaccounted-water-supply/30732789.
[67] http://www.dqindia.com/8-innovative-examples-of-big-data-usage-in-india/5/.

In a push towards energy efficiency, the Government of India has expressed its intent towards being "fully smart" by 2021 - 2022, with a view to install a smart meter in every home.[68] The India Smart Grid Task Force (ISGTF) and the India Smart Grid Forum (ISGF) were set up in 2010 to advise the Ministry of Power on policies, programmes and developments with respect to smart grids and also develop a smart grid roadmap for India.[69] In August 2013, the Government of India notified its "Smart Grid Vision and Roadmap for India" which seeks to "enable the development and deployment of an Indian Smart Grid model", and also set the stage for the launch of the National Smart Grid Mission (NSGM) which would have "its own resources, authority, functional & financial autonomy to plan and monitor implementation of the policies and programmes prescribed in the roadmap".[70]

There have been parallel government initiatives that look to utilise the potential of big data for development in this regard. In 2013, the Ministry of Power commissioned 14 Smart Grid pilot projects, where the Government would cover 50% of project cost.[71] The first such trial was conducted in Puducherry,[72] with the state Government recently signed a 46 crore smart meter project with Chinese firm DongFan electronics.[73] Ericsson has also signed a deal with the government of Assam in early 2016 to install 15,000 smart meters over the next three years.[74] A model Smart Grid Regulation was developed by the Forum of Regulators in 2015, [75] and the Government also launched the Ujwal Discom Assurance Yojana (UDAY) to increase efficiency of electricity distribution, where one of the commitments is to make installation of smart meters compulsory for consumers using more than 200 units of electricity a month.

Goal #08: Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all

CreditMantri is an online credit "trainer" that helps users improve their loan ratings.[76] Ahmedabad-based startup, LendingKart will use social media to evaluate borrower creditworthiness as part of the collaboration with Singapore-based startup Lenddo where these companies will explore alternative credit scoring solutions based on non-financial data sources and expect to speedily improve financial inclusion in the Indian market.[77] Neogrowth, a similar fintech loan provider, designs repayment plans for working capital loans based on the business cycles of small businesses, as evidenced by their digital transaction history.[78] FintechLabs uses data innovatively to provide credit assessment as well as credit underwriting – contributing in a hybrid lending model, where various NBFCs and banks contribute some portion of capital to be lent to customers based on an algorithmic assessment of their creditworthiness. It has also come up with a solution that facilitates the

[68] We should look at a smart meter in every home,: Piyush Goyal, The Hindu BusinessLine, http://www.thehindubusinessline.com/specials/clean-tech/we-should-look-at-a-smart-meter-in-every-home-piyush-goyal/article8245384.ece, Feb 16, 2016.
[69] http://www.indiasmartgrid.org/about-us.php.
[70] http://www.cyantechnology.com/wp-content/uploads/2014/04/Smart-Grid-Vision-and-Roadmap-for-India.pdf, Page 3.
[71] https://www.metering.com/14-smart-grid-pilot-projects-shortlisted-in-india/
[72] https://www.metering.com/india-puducherry-smart-grid-pilot-to-be-expanded-to-87000-meters/
[73] http://www.deccanchronicle.com/nation/politics/070716/puducherry-government-in-tie-up-with-dongfan-electronics-for-smart-grid-project.html
[74] http://www.business-standard.com/article/companies/ericsson-to-deploy-15-000-smart-meters-in-assam-116041300896_1.html.
[75] http://www.indiasmartgrid.org/reports/FOR%20-%20Smart%20grid%20regulations.pdf.
[76] https://www.bbva.com/en/news/economy/financial-and-commercial-services/fintech/7-keys-bring-fintech-financial-inclusion-together/.
[77] http://economictimes.indiatimes.com/articleshow/49829470.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst.
[78] http://www.developmentoutlook.org/2016/10/the-digital-age-of-financial-services.html.

conversion of financial data stored in physical form, in bank passbooks, into digital formats, making credit assessment easier and faster and making credit itself accessible in remote areas.[79]

Goal #09: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation
One of the five targets under this goal looks to build infrastructure that support economic development, with a focus on equitable and affordable access for all.[80] The Government of India launched its "100 Smart Cities Mission" in June 2015 in order to use the potential of technology to improve the quality of life in India through urbanisation and local area development.[81] There have been initial inroads into the realisation of this goal. For instance, Intelligent Transport System projects have been initiated in cities like Mumbai, Bangalore, Ahmedabad, Pune, Mysore.[82]

In context of private sector involvement, Microsoft has offered cloud services for smart cities,[83] Bajaj Electricals has partnered with Cisco to jointly bid and participate in building smart cities.[84] IBM is a major player in the smart cities mission,[85] Ericsson has a dedicated division to focus on three industries – utility, transport and public safety and smart city projects,[86] and firms like Cisco and HP are working with local bodies to implement technology in the core areas of the Smart Cities programme, i.e. in the areas of mobility, e-governance, water, waste management, education, healthcare, smart energy etc.[87]

The Ministry of Urban Development (MoUD) developed guidelines for these smart cities, that contemplate collaborations with governments, multilateral and bilateral institutions, and private sector players who can provide technical assistance and support.[88] The Swedish government is India's knowledge partner in this exercise, and will also help Maharashtra develop smart cities .[89] The governments of Malaysia,[90] the US[91], UAE,[92]

[79] http://www.developmentoutlook.org/2016/10/the-digital-age-of-financial-services.html.
[80] https://sustainabledevelopment.un.org/post2015/transformingourworld.
[81] Smart Cities in India : an Overview, http://cis-india.org/internet-governance/blog/smart-cities-in-india-an-overview.
[82] http://pearl.niua.org/sites/default/files/books/GP-IN1_UT.pdf
[83] http://www.hindustantimes.com/tech/nadella-announces-partnerships-with-indian-start-ups-snapdeal/story-TUDyqqiJjVC6UlVAMdnzhJ.html
[84] http://articles.economictimes.indiatimes.com/2015-01-18/news/58200774_1_smart-cities-project-major-cisco-narendra-modi-government
[85] http://www.business-standard.com/article/companies/global-tech-giants-gear-up-for-digital-india-road-map-114111701120_1.html
[86] http://economictimes.indiatimes.com/tech/hardware/ericsson-india-sets-up-new-division-to-focus-on-smart-cities-space/articleshow/47538127.cms
[87] http://indianexpress.com/article/india/india-others/new-age-townships-tech-tonic-shift-for-smarter-cities/
[88] Mission Statement and Guidelines, http://smartcities.gov.in/writereaddata/smartcityguidelines.pdf
[89] http://www.business-standard.com/article/pti-stories/sweden-offers-tie-up-to-maharashtra-for-smart-development-115101700684_1.html
[90] http://economictimes.indiatimes.com/news/economy/infrastructure/malaysia-proposes-to-invest-30-billion-in-urban-development-and-housing-projects-in-india/articleshow/49383787.cms
[91] http://indianexpress.com/article/india/india-others/india-us-to-set-up-task-for-developing-3-smart-cities/
[92] http://www.thenational.ae/business/technology/indias-smart-city-push-gets-boost-from-the-uae

and Spain[93] have expressed interest in this field, whereas Israel,[94] the UK,[95] France[96], and Germany[97] have outlined specific tasks for collaboration.

Goal #11: Make cities and human settlements inclusive, safe, resilient and sustainable

Towards this goal, some scholars have attempted to understand how data reveals relationships between citizens and the state in order to appreciate the possibilities of deployment of data in urbanisation. They observe that there is an increased availability of data sets with some degree of spatial information, new types of data are being generated from technologies that are already in use, and there is a marked increase in access and control of these data sets by skilled professionals in the IT, research, and commercial sectors.[98] The important task here, is to make priority based political choices intelligently by using available data. Other authors, like Rohan Samarajiva, believe that citizens, as the ultimate beneficiaries of urban development, should be considered the primary sensors while allocating resources. This, he argues, recognises the inherent complexity of the system and supports incremental changes in urban development.[99]

While discussing data mining techniques that can be used to combat natural disasters, some thinkers believe that reasons for poor disaster management procedures followed in India are inadequate early warning systems, poor preparation before the disaster occurs, inadequate and slow relief operation, lack of proper administration, slow process of rehabilitation and reconstruction, poor management of finances for relief work and lack of effective help to victims. These authors argue that data mining and analytical techniques can help with the three disaster response tasks: 1) prediction, 2) detection and 3) disaster management strategy. Data mining can help with detecting the natural disaster, which assists in preparation for the proper management strategies. Although you cannot predict a natural disaster during data, there are insights that data can offer - such as vulnerable areas, susceptible times etc. [100]

Other scholars believe that leveraging the power of social media can create a massive repository containing data that would not be available through traditional sources of information. For example, some scholars have proposed a transfer learning method for urban waterlogging disaster analysis, which provides the basis for traffic management agencies to generate proactive traffic operation strategies in order to alleviate congestion. They suggest transferring waterlogging by using the copious amounts of information from social media and satellite data to improve urban waterlogging analysis. By using a multiview discriminant transfer learning method between cities for urban waterlogging disaster analysis, this method could potentially address the gap between data availability in large cities and the same in small towns. It intends to transfer knowledge from large cities to small ones for urban waterlogging disaster analysis.[101]

---

[93] http://indianexpress.com/article/india/india-others/smart-cities-projectfrom-france-to-us-a-rush-to-offer-assistance-funds/

[94] http://timesofindia.indiatimes.com/india/Maharashtra-CM-and-Tel-Aviv-mayor-agree-to-cooperate-on-smart-cities/articleshow/47071659.cms

[95] http://www.dnaindia.com/money/report-uk-joins-hand-with-govt-to-develop-3-smart-cities-in-india-2147386

[96] http://timesofindia.indiatimes.com/smart-cities/Chandigarh-Puducherry-Nagpur-on-Frances-smart-city-radar/articleshow/50707788.cms

[97] http://economictimes.indiatimes.com/news/economy/infrastructure/germany-to-develop-kochi-coimbatore-bhubaneswar-as-smart-cities/articleshow/51295816.cms

[98] Data, Urbanisation and the City, Anant Mariganti and Partha Mukhopadhay, EPW, Vol 1, No 22, May 30, 2015.

[99] Big Data to Improve Urban Planning, Rohan Samarajiva, EPW, Vol 1, No 22, May 30, 2015.

[100] A review on application of data mining techniques to combat natural disasters, Saptarsi Goswami, Sanjay Chakraborty, Sanhita Ghosh, Amlan Chakrabarti, Basabi Chakraborty, Ain Shams Engineering Journal, January 2016.

[101] Social Media meets Big Urban Data: A Case Study of Urban Waterlogging Analysis, Ningyu Zhang, Huajun Chen, Jiaoyan Chen, and Xi Chen, Hindawi Publishing Corporation Computational Intelligence and Neuroscience Volume 2016, Article ID 3264587.

There have been several initiatives from both the government as well as the private sector to achieve these goals. SocialCops, a data and analytics startup, directly crowdsources data regarding public infrastructure and access to services via mobiles from citizens. Tracking parameters such as teacher attendance in public schools, quality of public infrastructure, adherence and access to medical care on a continuous basis can now be done easily by collaborating with partner non-profit organizations.[102] Chennai Corporation was the first in India to adopt a dashboard to manage civic issues and track the progress of maintenance work in the neighbourhood with portals that are proactive in disclosure of public data. The Chennai Data Portal and Dashboard project (CDP) is part of the e-governance project.[103]
Also, in a bid to build high calibre data for the Indian cities and realizing the importance of data to serve as a catalyst to create truly smart, sustainable, resilient, inclusive and prosperous cities of tomorrow, City Data for India initiative was announced in January 2017 by Tata Trusts in partnership with the World Council on City Data (WCCD).[104] In India, Cisco is working with state governments and its partners to digitally transform 14 cities and plans to connect 100 cities as part of the Digital India initiative.[105]

Goal #13: Take urgent action to combat climate change and its impacts
While we did not find literature that reflected on big data's potential with respect to climate change, there have been some initiatives. The Surat Municipal Corporation (SMC) has set up a system to deal with climate change, named Urban Service monitoring system or UrSMS. The system put in place in 2012 helps gather disease data from almost all government and private hospitals in real time. The data is analyzed by SMC to ensure immediate interventions. UrSMS has been shortlisted for the Big Data Climate Challenge competition that aims to show how projects working with big data can push action against climate change.[106] Bharat Light and Power (BLP) is also exploring ways in which cloud computing and big data could enable the more efficient operation of clean energy plants.[107] Responding to India's sustainable energy challenges, BLP has entered into a 10 year strategic plan with IBM, under which they can use its SoftLayer Cloud capabilities to monitor the vast amounts of data being collected in order to create more efficient power plants.[108]

Goal #16: Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels.
Currently, the PMO is using big data to process citizens' ideas and sentiments through mygov.in.[109] Cloudera offers public sector entities a secure and cost efficient place to store and analyse their data, empowering them to derive new insights and correlation while extending the value of existing investments. Big Data is also used by the Department of Human Services to improve service delivery, including creating personalised services and detecting fraud and compliance issues.[110] It is also possible to use big data analytics for combating tax evasion and increases the government's revenue base. This, according to some thinkers, improvement of the

[102] https://yourstory.com/2014/07/socialcops-funding/.
[103] http://www.thehindu.com/features/homes-and-gardens/now-info-just-a-click-away/article6966275.ece.
[104] http://www.newkerala.com/news/fullnews-223106.html.
[105] http://tech.economictimes.indiatimes.com/news/corporate/cisco-committed-to-help-build-100-smart-cities-in-india-chuck-robbins/55202167.
[106] http://timesofindia.indiatimes.com/city/surat/UrSMS-shortlisted-for-Big-Data-Climate-Challenge-contest/articleshow/42968420.cms
[107] https://www.weforum.org/agenda/2014/05/big-data-will-clean-indias-energy-act/.
[108] https://www-03.ibm.com/press/us/en/pressrelease/42524.wss.
[109] http://www.greatlearning.in/blog/pmo-india-big-data/.
[110] Big Data Analytics in Government and Organizations, https://www.ijarcsse.com/docs/papers/Volume_5/5_May2015/V5I5-0418.pdf.

government tax net by convergence of data.[111] Janalakshmi Financial Services uses Big Data to get insights into the behavior of the unbanked masses. Using these insights, the firm designs a customized loan. The microfinance firm will also use geo-spatial insights with credit-decisions process to help it improve distribution and collection models.[112]

Goal #17: Strengthen the means of implementation and revitalize the global partnership for sustainable development

As companies and governments work together to analyse unstructured data from non traditional sources that require more than simple queries, the non linear rise in computing becomes important to consider. Big data can enable the government to do existing things more cheaply, do existing things better, and new things that we don't do yet. Public sector data sources can potentially increase efficiency, cost effectiveness, and control over fraud. Data driven decision making has been the backbone of government economic policy. [113]

Big Data in India: Government Initiatives and Legal Framework

Introduction

In 2014, the Department of Science and Technology (henceforth, DST) of the Government of India engaged the Consultancy Development Centre to prepare a "Strategic Roadmap for Big Data Analytics" in India. While the Government of India has been exploring the use of big data analytics since at least 2009, this report by the DST was the first exercise to generate a strategic vision on use of big data in India, both in the public and the private sectors, and to build an enabling configuration of initiatives and frameworks. The report, however, was not made public. Although India has seen a growing number of big data events and start-ups in recent years, the active participation of the government remains largely absent. In this brief case study we document the initiatives taken by the Government of India to support adoption of and innovation with big data technologies in public and private sectors, and the components in the existing legal framework for governance of information technologies in India that are most relevant for big data.

The two key findings are the following:

1. The Government of India is yet to produce a comprehensive strategy vision on how public and private sector entities may adopt and innovate with big data technologies, and systemic challenges thereof, although a few but large-scale big data analytics projects have already been undertaken by government agencies, mostly in partnership with IT companies.
2. No general legal framework exists in India either to govern use and implementation of big data technologies for either public or private services, and the only legal safeguard for users and citizens are provided by a disparate and uneven landscape of sector-specific laws.

Government Initiatives

In a call for proposals for projects to be supported under its Big Data Initiative, the DST names the following major fields of application for big data:

[111] http://journals.sagepub.com/doi/full/10.1177/0256090915575450. Also see: http://economictimes.indiatimes.com/tech/ites/government-plans-to-use-big-data-analytics-for-tax-collections-infosys/articleshow/25222293.cms?intenttarget=no.
[112] http://www.dqindia.com/8-innovative-examples-of-big-data-usage-in-india/6/.
[113] Big Data: Prospects and Challenges http://journals.sagepub.com/doi/full/10.1177/0256090915575450.

- "Complex Big Data Applications in Science, Engineering, Medicine, Healthcare, Finance, Business, Law, and Education
- Indian Traditional Knowledge
- Transportation
- Retailing, social media and Telecommunication
- Big Data Analytics in Small Business Enterprises (SMEs)
- Big Data Analytics in Central and State Governments, Public Sector and Society in General
- Real-life Case Studies of Value Creation through Big Data Analytics
- Big Data as a Service
- Big Data Industry deployments & Standards and Experiences of Big Data Govt Deployments/ Projects."

As a key component of this Big Data Initiative, DST proposed to set up a series of Centres of Excellence on big data technology and data science research across the country. The first of such Centres was declared to be established in Hyderabad in partnership with NASSCOM.
The Department of Biotechnology (henceforth, DBT) is one of the rare government agencies in India that has actively initiated a sector-specific discussion about the possibilities and implication of big data. At the Indian Science Congress 2016, the DBT organised a series of lecture and discussion on the challenges and opportunities of using big data for biotechnology research in India. The speakers discussed biological data analytics in agriculture, machine learning in oncology research, and big data and personalised medicine.

The other key government agency to take leadership in this topic has been, expectedly, the Ministry of Electronics and Information Technology. Last year, it organised a workshop on "Leveraging Big Data Analytics for Bringing Social Change". The workshop focused on identifying "areas where [big data] can be applied in Government domain … [and also to] give direction to industry and academia to focus on making scientific advances and research efforts in areas which need immediate attention for making data analytics acceptable in Government." No post-workshop report has been available yet. The Convergence Commission Division of the Ministry has also undertaken a few big data projects - such as to provide information and decision support system for intensive care units and for crowd control and management at large-scale public events - but no details about such projects are available in public.

As for public-private partnerships involving big data are concerned, several state governments have engaged global IT companies such as Microsoft (using machine learning for ex-ante identification of potential school dropouts in Andhra Pradesh) and PWC (using sentiment analysis to evaluate the mood and responses of the commentators on the citizen engagement platform called MyGov). Most recently, Google and Reliance, respectively, have initiated two exemplary projects on massive-scale business process digitisation and communication data mining projects.

Legal Framework
While the scope of application of big data is immense across sectors like medicine, credit, healthcare, law enforcement, justice system and public policy, there exist no general regulatory framework to govern its practice and implications.

India has a limited data protection law under the Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules 2011 formed under section 43A of the Information Technology Act, 2000. While this allows for challenging of the indiscriminate collection of personal data, it fails to take into account the changing nature of data collection and processing. The present definition of personal data hinges on the factor of identification (data that is capable of identifying a person). Yet this definition does not encompass information that is associated to an already identified individual - such as habits, location, or

activity. The definition of personal data also addresses only the identification of 'such person' and does not address data that is related to a particular person but that also reveals identifying information about another person - either directly - or when combined with other data points. A growing area of research globally is the social consequences of big data with a particular focus on its tendency to replicate or amplify existing and structural inequalities. India has some anti-discrimination laws in the form of constitutional protections that prohibit discrimination on the grounds of certain identified attributes. However, these laws are inadequate to deal with the subliminal issues of bias and discrimination that can be caused due to big data.
The Consumer Protection Bill, 2015, tabled in the Parliament towards the end of the monsoon session has introduced an expansive definition of the term "unfair trade practices." The definition as per the Bill includes the disclosure "to any other person any personal information given in confidence by the consumer." Although this provision could be used in the future to address consumer privacy and data sharing concerns, its applicability and operationalisation are unclarified. The conversation on use of competition or antitrust laws to govern big data is still at an early stage in India, and how the role of data as market power is included in Indian jurisprudence remains to be seen.

The use of big data across different citizen facing sectors is translating in the possible application of sector specific laws on big data practices. For instance, use of big data in law enforcement practices such as predictive policing and social media monitoring could potentially raise questions about inferences drawn from big data analysis could reconfigure evidence law and existing criminal law standards. The use of big data analysis in medicine and healthcare practices is on the rise, and we are already seeing legal proposals such as the draft Electronic Data Records standards in order to both enable and govern collection of medical data.

Last but not the least, the use of big data in the emerging Fintech sector is leading to an expansive discussion on how financial regulators like the Reserve Bank of India should view new business models driven by big data. A recent report published by the Ministry of Finance proposes an independent regulator for the digital payments sector, conceptualised as a separated from banking services related to borrowing and lending. In our comments to this report we have highlighted that "the decision by the Government of India to withdraw the legal tender character of the old high denomination banknotes (that is, Rs. 500 Rs. 1,000 notes), declared on November 08, 2016, have generated unprecedented data about the user base and transaction patterns of digital payments systems in India, when pushed to its extreme use due to the circumstances." We eagerly await further communication from the government regarding the way forward for governance of financial big data in India, as well as use of the same for data-driven regulation and policy-making.