ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645
Indian Journal of Science and Technology, *Vol 8(S4), 50-60, February 2015* DOI: 10.17485/ijst/2015/v8iS4/61494

Data Analytics for Rural Development

Anu Peisker * and Soumya Dalai

Deloitte, India; apeisker@deloitte.com, sdalai@deloitte.comm

Abstract

The impact of big data is enormous virtually in every business sector. The usage so far has been more focused on the e-commerce and marketing sectors. But the wide reach of big data can provide much more innovative, profitable and yet beneficial solutions for many perennial problems faced by different sectors. Rural India plays a vital role in economic growth of the country through agriculture, self-employment, construction, services, etc. The developments in telecommunication sector provide suitable platform for applying data analytics in rural India. The focus of the Government to build digital India through broadband highways connecting every household, village, panchayat, Government department will generate huge amount of data which can be analysed to provide solution to the never ending problems of rural India and to create smarter villages. This paper provides conceptual framework for the application of data analytics in enhancing rural development by supporting different sectors such as agriculture, banking, governance and healthcare.

Keywords: Agriculture, Applied Data Analytics, Governance, Healthcare, Rural India

1. Introduction and Motivation

The benefits of big data are not only transforming the businesses in the urban areas but also they can be equally extended to the rural parts of the country. The penetration of mobile phones and other connected devices generate huge amount of structured and un-structured data every day. Productivity in rural India is extremely low due to unscientific farming practices, fragmented land holdings, lack of agro-climatic focus for crops selection, lack of access to the right farming advice at the right time. While focus has been on improving farm production, reducing food supply-chain losses remains a relatively unaddressed problem. The rise of the middle class has also raised the expectations of citizens and businesses for higher standards of service as well as greater transparency and access to government leaders, policies, and operations. Continuing globalization is forcing Government to attract foreign investments for economic growth by improving government services to make India a desirable place for both businesses and individuals [1].The Government of India is taking steps, such as UniqueIdentity (UID) Project, rural employment poverty up liftment schemes and financial inclusion initiatives, poverty up liftment schemes and financial inclusion initiatives, to drive the overall development of the economy. The push towards policy innovation and digital India through broadband highways will drive spend on analytics, Geographical Information System (GIS), risk management software, portfolio management, portal enhancements, web content management, and customer engagement solutions. The national e-Government plan has proposed setting up of 250,000 tele-centers in rural areas to provide public access to computers and online service delivery. e.g. UNICEF's 'Digital Drum' allows populations in remote areas to seek answers on issues ranging from agriculture, health, and education to business advice and entertainment, providing a window on the interests and concerns of information seekers whose location, age and gender are generally recorded. Government is leveraging on technology to expand public and private financial networks in rural areas. Integration of information and services for citizens facilitates internal governance as well as transparency of government through external access.

Big data for development is about turning imperfect, complex, often unstructured data into actionable information. This implies leveraging advanced computational

*Author for correspondence

tools, and customer engagement solutions. The national e-Government plan has proposed setting up of 250,000 tele-centers in rural areas to provide public access to computers and online service delivery. e.g. UNICEF's 'Digital Drum' allows populations in remote areas to seek answers on issues ranging from agriculture, health, and education to business advice and entertainment, providing a window on the interests and concerns of information seekers whose location, age and gender are generally recorded. Government is leveraging on technology to expand public and private financial networks in rural areas. Integration of information and services for citizens facilitates internal governance as well as transparency of government through external access.

Big data for development is about turning imperfect, complex, often unstructured data into actionable information. This implies leveraging advanced computational tools, which have developed in other fields, to reveal trends and correlations within and across large datasets that would otherwise remain undiscovered. Big data analytics can help in achieving some of the high priority objectives of the Government from the 12th five-year plan. The efficient and meaningful use of technology and analytics in various growth drivers like education, agriculture and healthcare could help in improving productivity and employability of rural communities.

2. Big Data

Big data technologies describe a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery, and/or analysis. From a technology perspective, big data gives: the possibility of better storage (volume), the ability to process the information and make it available in real time (velocity) and the ability to deal with various kinds of data sources, including structured, semi-structured and unstructured ones (variety)[2].

Worldwide big data technology and services market is expected to grow to $16.9 billion by 2015 at a Compound Annual Growth Rate (CAGR) of 40% – about seven times the rate of the overall Information and Communication Technology (ICT) market[3].

The amount of data created every year is exponentially growing. In 2016, global IP traffic will reach 1.1 zettabytes (1021) per year, and by 2018, it will reach 1.6 zettabytes per year . Currently, 60% of data in the digital universe is attributed to mature markets such as Germany, Japan, and the United States, but by 2020, the percentage will flip, and emerging markets including Brazil, China, India, Mexico and Russia will account for the majority of data [5].Data could be coming from disparate sources like Geographical Positioning System (GPS) devices, sensors, social networking sites, company websites, traditional, legacy data sources etc. Assembling such wide and deep data at single place, integrating it and gaining meaningful understanding is a mammoth task. The impact of big data is enormous virtually in every business sector. The usage so far has been more focused on the e-commerce and marketing sectors. But the wide reach of big data can provide much more innovative, profitable and yet beneficial solutions for many perennial problems faced by different segments. e.g. The Alameda County (California) Social Services Agency deployed an information management system that combines analytics with Business Intelligence (BI) to give workers an agency wide, comprehensive view of individual cases to better track program results and meet specific targets for recipients' work participation rate and allow them the real-time visibility to monitor their progress towards those Key Performance Indicators (KPIs). Big data in agriculture has the potential to increase yield production, and as we near an era shaped around more people and less resources, this makes farming one of the most important careers in the world[6]. Many public and private companies are increasingly relying on BI and decision support tools and services to obtain better metrics from their current systems, make more timely and informed decisions, and bring about more transparency and accountability. Big data offers opportunity for improved understanding of human behaviour that can facilitate early warning, real-time awareness and feedback. Finding an implementation solution for big data is often difficult as it should combine scalability, performance, ease of use and low total cost of ownership. Challenges in the application of big data are mainly related to privacy, access and sharing of data.

3. Literature Review

The use of data has increased significantly in the last decade as the world has entered into a digital era. The booming market of mobile phones and smart devices with the capability of digital communication has significantly raised the data usage and given rise to the need for powerful analytics tools to analyse the data in a meaningful way.

Kenneth Cukier have stated that the fruits of the information society are easy to see, with a mobile in every pocket, a computer in every backpack and big-information technology systems in back offices everywhere . They also mentioned big data as a helpful connection between the citizens and the government. Big data refers to the things one can do at a large scale that cannot be done at a smaller one, to extract new insights or create new forms of value, in a way that changes markets, organizations, the relationship between citizens and government and more. Bill Franks has defined big data more precisely as big data sources are not always tightly defined up-front and typically capture everything that may be of use . This can lead to wading through messy, junk-filled data when doing analysis. Burkholder added the importance of analytics to the large unstructured data sets . Big data not only refers to very large data sets and the tools and procedures used to manipulate and analyse them, but also to a computational turn in thought and research. Frank J. Ohlhorst stated that the real trick with big data is to find the best way to deal with the varied data sources and still meet the objective of the analytical process . Danah Boyd & Kate Crawford questioned the privacy, security and ownership of big data . Bryant, Katz, Lazowska give a broad view of big data usage . According to them advances in digital sensors, communications, computation, and storage have created huge collections of data, capturing information of value to businesses, science, government, and society. Social Networks and search engines collect trillions of bytes of data every day and continually add new services such as satellite images, driving directions, and image retrieval. The societal benefits of these services are immeasurable as they have transformed the way people find and make use of information on a daily basis. The size and scope of big data is enormous and it can be beneficial for the rural population in many different sectors like Healthcare, Agriculture, Education, Banking, etc.

3.1 Healthcare

Big data analytics has already proven its success in the healthcare sector. As per Peter Groves et al, the Healthcare stakeholders now have access to promising new threads of knowledge through bigdata; so called not only for its sheer volume but for its complexity, diversity and timeliness[13]. Although in early stage, pharmaceutical-industry experts, payers and providers are beginning to analyse bigdata to collectively help the industry address problems related to variability in healthcare quality and escalating healthcare spend. Jensen et al believe that effectively integrating and efficiently analysing various forms of healthcare data over a period of time can answer many of the impending healthcare problems[14]. healthcare spend. Jensen et al believe that effectively integrating and efficiently analysing various forms of healthcare data over a period of time can answer many of the impending healthcare problems[15].

3.2 Banking

Emphasizing the importance of big data in the banking sector Halevi and Moed, stated that big data can be seen in the finance and business environment where enormous amount of stock exchange, banking, online and onsite purchasing data flows through computerized systems every day and is then captured and stored for inventory monitoring, customer and market behaviour study . Barton and Court revealed in their study, data is essential, but performance improvements and competitive advantage arise from analytics models that allow managers to predict and optimize outcomes . This implies, however, that banks, with the core capability of harnessing big data through daily transactions, customer-service records, real-time feeds, social media posts and correspondence, have a competitive advantage at deriving more insights about their business over their rivals who lacks this unique capability.

3.3 Agriculture

John-Deere is using bigdata to keep track of tractor performance by analysing data from connected devices. Lindsey Giplin argued that to feed the world's rapidly-expanding population in the coming decades, agriculture must produce more[18].

3.4 Education

Education is utilizing newer technologies and transforming the learning experience. Siemens and Long claim that big data and analytics are going to be the biggest factors in what is going to shape the future of higher education[19]. Picciano describes a case of a using analytics program in a school in Arizona to track student progress on the website of online courses they offer[20]. They track all student activity like login/logout information, number of mouse clicks and number of page views, time spent on a page, contents posted by student, etc. This gives them the idea about the performance of students and the topic or subject they are struggling to understand.

3.5 Governance

Even traditionally less innovative organizational bodies like government owned businesses and funds (e.g. pension fund) could benefit from the developments in big data . Lev Manovich opined that humanists should use data analysis and visualization software in their daily work, so that they can combine quantitative and qualitative approaches in all their work[22].

4. Research Methodology

4.1 Research Objective

The objective of the research is to propose a framework for the application of data analytics in enhancing rural development by supporting different sectors such as agriculture, banking, governance and healthcare. The research aims to provide a conceptual analytics platform which will help various stakeholders related to the above mentioned sectors with adequate information and forecasts; thus increase the productivity and is beneficial for socio-economic and community development.

4.2 Research Design

Initially secondary research was done to find current practices in various sectors and historical processes, workflows and trends followed. Based on this, sector specific focus groups in governance, healthcare and agriculture were made to brainstorm on the challenges faced by them and the ways to help them better meet with the desired objectives, at the same time providing incremental benefits in the future. The analysis was done upon the data collected from the focused interviews of industry experts. This qualitative research process rendered very crucial industry insights and expert opinions which were fundamental in creating a conceptual framework for overall development in rural as well as urban areas. The paper builds on some of the most recent findings in the field of data science, and findings from our own collaborative research. Following research process helped in the sequential and comprehensive study:

5. Conceptual Framework

5.1 Data Integration

Whether reforming an individual sector or undertaking a city wide operational reform, the first step is to integrate processes and data across diverse sector infrastructure. e.g. Barcelona City Council created a comprehensive data warehouse gathering data from more than 20 heterogeneous source systems supporting the automation of more than 2,500 KPIs.The 'traditional data' (official statistics, survey data, etc.) will continue to generate relevant information, but the digital data revolution presents a tremendous opportunity to gain richer, deeper insights into human experience that can complement the development indicators that are already collected.e-Governance would imply that beneficiaries should progressively benefit from a legally binding fully electronic and paperless exchange with rural development bodies and national administrations at all levels. e-Governance would also require applying the 'only once' encoding principle, allowing beneficiaries to encode data only once - those data being shared between different administrations at national and regional level - using electronic online portal functionalities. At State level, e-Governance means that digital signature should be implemented effectively by electronic portals at national and regional level and that electronic storage features are made available in national systems used for dealing with rural development projects. e.g. The 2009 US economic stimulus package, the American Recovery and Reinvestment Act (ARRA), included provisions to encourage hospitals and doctors to accelerate the migration of patient information from paper to electronic forms .The data could be sourced from various applications/systems as per the availability of data across sectors. Following are the list of top data sources:

New data sources for big data include industries that just recently began to digitize their content. In virtually all of these cases, data growth rates in the past five years have been almost infinite, since in most cases it started from zero. Industries include:

•. **Media/entertainment:** The media/entertainment industry moved to digital recording, production, and delivery in the past few years and is now collecting large amounts of rich content and user viewing behaviours.

•. **Healthcare:** The healthcare industry is quickly moving to Electronic Medical Records (EMR) and images, which it wants to use for short-term public health monitoring and long-term epidemiological research programs.

•. **Video surveillance:** Video surveillance is still transitioning from CCTV to IPTV cameras and recording systems that organizations want to analyse for behavioural patterns (security and service enhancement).

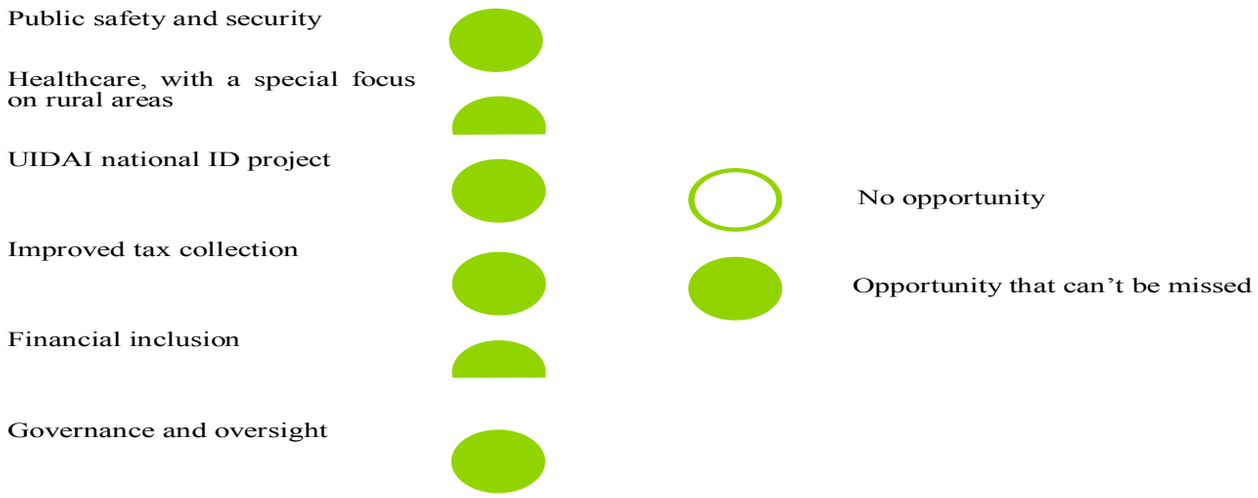

Figure 1. Opportunity to meet a spectrum of government objectives through analytics.

Figure 2. Research Process.

Figure 3. High level process.

Table 1. Data Sources by Industry

Healthcare	Agriculture	Governance
Electronic Health Records	Kissan SMS portal system	UID card
Vaccination records	Community Information Centres	PAN card
Clinical Information Systems	AGMARKNET	Passport
	e-Choupal	License
	Agriwatch.com	Census data
		Tax & expenditure info

•. **Transportation, logistics, retail, utilities, and telecommunications:** Sensor data is being generated at an accelerating rate from fleet GPS transceivers, RFID tag readers, smart meters, and cell phones (call and usage data records); that data is used to optimize operations and drive operational BI to realize immediate business opportunities. The complexity of handling this expanded universe of data sources is compounded by the need to link, match and transform data across business heterogeneous entities and systems, while managing scale and timeliness. Consumers are increasingly active participants in a self-service marketplace that not only records the use of affinity cards but can increasingly be combined with social networks and location-based metadata, which creates a gold mine of actionable consumer data for retailers, distributors, and manufacturers of consumer packaged goods. The key to effective data-driven decision-making is the ability to sift through large

amounts of data; and the ability to combine data from several sources to gain a more comprehensive view of the business.Seemingly unimportant data can become crucial once combined with other sources to reveal new insights.

5.2 Data Normalization

Table 2. Data Formats by Industry

Industry	Video	Image	Audio	Text/Numbers
Banking	Medium	Medium	Medium	High
Insurance	Low	Medium	Low	High
Health care	Low	High	Low	High
Government	High	Medium	High	High
Education	High	Medium	High	Medium
Agriculture	Low	Medium	Low	High

Part of the need for new technologies for big data (versus older, legacy Relational Database Management Systems) has to do with the format of the data coming in from various applications. Data quality is a big issue. A more dynamic, flexible database schema is needed to handle the structured, semi-structured, and unstructured data that comprises today's big data. These database schemas need to be able to deal with a wide range of data formats since data formats density varies significantly across industries as shown below:Big data for development generally share some or all of these features:

•. Digitally generated: data are created digitally (as opposed to being digitised manually), and can be stored using a series of ones and zeroes, and thus can be manipulated by computers.

•. Automatically collected:system automatically extracts and stores the relevant data as it is generated

•. Geographically or temporally traceable: e.g. mobile phone location data or call duration time

•. Continuously analysed: information is relevant and can be analysed in real-timeSince there is no 'perfect' data, a concept of 'fit for use' is applied keeping in mind the purpose of data. The potential of big data for development is best realised when its limitations, biases, and ultimately features, are adequately understood and taken into account when interpreting the data. Data must be able to comply with all the laws international, federal and state regulations, fiscal and monetary reporting statutes and all applicable civil rights laws, including privacy and security.

5.3 Data Storage (in the Cloud)

The cloud can offer a centralized knowledge bank which can be used to store all the industry related information. This information bank will be available to the stakeholders and other users from the sector at any place and at any time at a very reasonable cost. Vertical sectors such as financial services and retail are leading the adoption of mobile and cloud technologies in India .

Cloud-enabled device management helps in user authentication and secure file sharing and syncing of diversified application access. Real-time usage inputs data analytics to help in identifying trends, gauging utility ratios and making fact-supported decisions. These enable merchants to access all operation-related information and functionality through a Point-Of-Sale (POS) application installed on a consumer's device, such as a PC or mobile, without having to invest in specialized POS hardware, servers or storage. All the product and customer data, management, reporting, and analytics are done in the cloud. Some solutions can also work on PCs and laptops, but tablets and smartphones are the main target devices for this type of solution. e.g. Farmers in Sahara, South Africa use a cloud-based trading system that disseminates information about planting schedules, crop status, harvesting times, and market prices through mobile phones . The cloud based storage has helped health care providers with access to health record of patients remotely through a mobile device.

5.4 Metrics/Co-relation

The first step is to connect data collected from a variety of sources: network and non-network, structured and unstructured. Unique data elements (metrics) for the subject areas viz. healthcare, governance, banking, agriculture should be identified. Analysis should be done on the overall availability and quality aspects of the identified fields. A common framework for information processing is defined to discover patterns and trends in the data. A common reference guide is required that will help tag data with following parameters in Figure 4:

5.5 Analytics

Big data analytics refers to tools and methodologies that aim to transform massive quantities of raw data into 'data about the data' - for analytical purposes. Such exceedingly large data volumes cannot be analysed with ordinary assessment methods, such as sampling or simple spreadsheets. They typically rely on advanced visualization techniques and powerful algorithms. Algorithms are able to detect patterns, trends, and correlations over various time horizons in the data and can help in detecting anomalies in the form of large deviations from the expected trends or relations in the data. Visualisation tools can provide new perspectives on findings that would otherwise be difficult to realise.Biggest power of data analytics is the predictive capability that can help to determine reliable patterns and forecast what might happen in the future. e.g. When predictive analysis is applied to public health, online data has been used as part of syndromic surveillance efforts. According to the US Center for Disease Control and Prevention (CDC), mining vast quantities of health-related online data can help detect disease outbreaks 'before confirmed diagnoses or laboratory confirmation'[27].With improved real-time connectivity and data management arise the possibility to create tailored data sets, readily available for analysis. This would be the core ingredient in data-driven efficiency improvements in a number of sectors – for example, transport, logistics, energy, agriculture and environmental protection. Turning big data into actionable information requires using computational techniques to unveil trends and patterns within and between these extremely large socio-economic datasets.It is important to shift from analytics application silos to more generic, horizontal analytics environments that take in a wide array of data sources, while supporting a variety of applications and services. Cutting-edge IT components like data storage, data management and network resources can work in harmony with domain- specific analytic logic (data models, rules sets, and so on) to bring benefits in terms of agility and scalability. e.g. Indian Agricultural Statistical Research Institute has developed a Decision Support System on nutrient management in crops. Under the e-Governance programme, Soil Health Card software has been standardized and in collaboration with Indian Institute of Soil Science, Bhopal, web-based software has been developed to provide integrated nutrient management recommendations using 'Soil Test Crop Response' method for 8 states[28]. Analytics to be fully effective can be automated and run at regularly scheduled intervals, which enables immediate identification of potential high-risk exception transactions in real-time.Furthermore, informed and effective evidence-based decision-making in business and society would be facilitated by access to insights based on analysis and interpretation of more accurate and up-to-date data. Leveraging big data can reduce time lag and human inputs/errors in data collection, production and transmission.

6. User Interface

Insights can be provided through an integrated and interactive platform that implements or reuses existing APIs and IntegratedDevelopment Environments (IDEs), allowing multiple applications to be supported by the same underlying data infrastructure. Insights can be derived from a properly correlated data only. Many commercially available continuous monitoring packages include sophisticated web interfaces, e-mail notifications, workflows, remediation tracking, dashboards and/or heat maps. Following Figure 5 outlines the overall conceptual framework based on the above processes for data analytics approach towards rural development:

7. Analysis and Discussion

Currently very few cloud initiatives are seen in the Indian public sector, primarily due to data security and privacy concerns. Central Government should take initiative and push the state Governments to embrace ICT infrastructure and use private cloud model for better citizen services and improved efficiency. Some states, like Kerala, are taking the lead in providing government-to-citizen

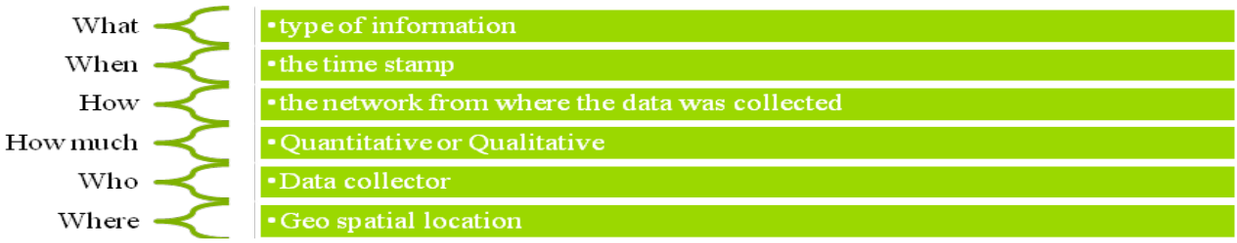

Figure 4. Parameters for tagging data.

(G2C), government-to-business (G2B), and government-to-government (G2G) services through their state data centres. UIDAI program can help in consolidation of ICT infrastructure. As the government's cloud strategies develop, open source software will play a prominent role. e.g. NIC deployed the open source-based Eucalyptus cloud software to provide e-Government to remote areas in the country. Applications such as e-Procurement, payroll, and grievance reporting will soon be hosted on Eucalyptus clouds at the national level. These technologies — e-forms, records management, e-Discovery, bar code, Enterprise Content Management (ECM), Business Process Management (BPM), and analytics — are infrastructure improvements needed to meet sectorial goals. Technical approaches to address security and privacy should include secure communication, adjusted data collection (signed nondisclosure or confidentiality agreements), secure data storage (data encryption), stringent and granular access control and audits.

7.1 Governance

Holistic and smart Governance for better citizen services and improve efficiency

As the pressure to lower costs and use fewer resources increases, Indian government bodies are finding their operations more difficult to manage. A drive to govern smarter will drive investments in analytics. e.g. the central government's Crime and Criminal Tracking Network & Systems (CCTNS) relies heavily on analytics to provide real-time information and data on criminals to police stations across the country. Advanced e-Government solutions and one-stop-shop web-portals with sophisticated functionalities (e.g. tracking of files, online submission of applications and documents, online help desk, uploading project information) can diminish the administrative burden for businesses and individuals. Such capabilities constitute a platform on which to build and the foundational basis for process simplification at State administration level. It could also create productivity gains in states. e.g. The government aims to computerize all tax related activities for state governments to improve the tax collection base. It is important to think holistic governance rather than in silos for each department such as public safety, transportation, emergency, utilities, etc. and hence export and aggregate the data to leverage on the holistic view of the data for various systems. e.g. The National Data Sharing and Accessibility Policy passed in 2012 made interdepartmental sharing of data mandatory over government networks throughout the country. The policy will help the Indian government move from a department-centric model to a more integrated and collaborative approach over time[29]. Aggregating data into a single repository, and analysing it for patterns and anomalies, enables government to see trends that might not otherwise have been captured. Government could leverage on the data captured from citizens and use it to refine city programs and policies.

7.2 Healthcare

Provide personalized care through right intervention to the right patient at the right time

While India spends around 4.2% of its Gross Domestic Product (GDP) on healthcare, the government's contribution is roughly one-fifth, while that of households' Out-Of-Pocket (OOP) is nearly 70% of the overall healthcare expenditure . An additional challenge for healthcare is the lack of cost-of-care data as there is no uniformity of care costs.

Realizing the increasing importance of healthcare, the government has increased healthcare funding by 25% year on year in the 2012-2013 budget, with a 15% increase in allocation reserved for the National Rural Health Mission program. With the emergence of many new large and small hospitals and healthcare units in rural areas, IT hardware like storage, services, security, and networking equipment and software like digitization, clinical information systems, and hospital information systems will receive a boost from rural healthcare spending. A unified-national health analytics system would require database consolidation, data cleansing, a payment tracking system,

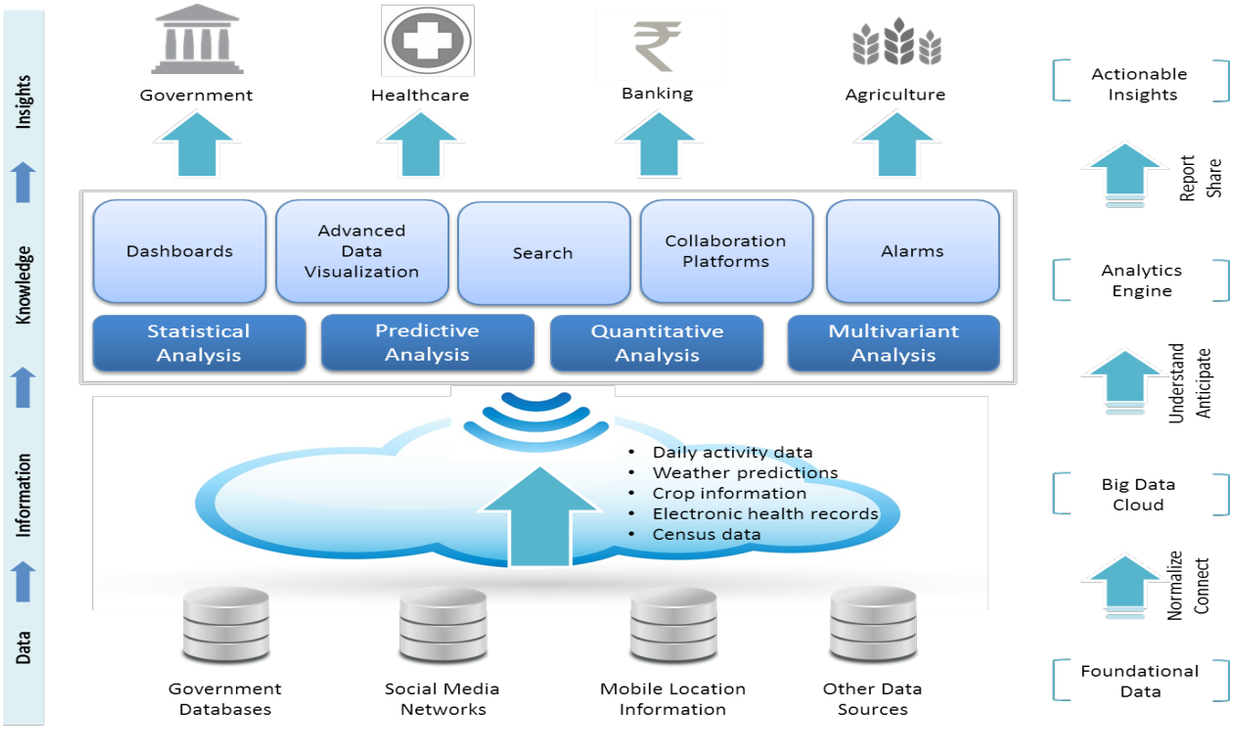

Figure 5. Conceptual framework for integrated analytics approach towards rural development.

an online reservation system for medical visits, electronic patient data and health product code standardization, such as physical exams, surgeries and drugs for reimbursement of insurance claims, standardization of payments, drug costs and reimbursements for hospital fees, etc. Content should be organized to support 'results-based' analysis.

BPM suites are an important tool to coordinate processes that span clinical and administrative systems. Focus is required on portals to provide access to the information silos (e.g. EMR, ECM, and Digital Imaging and Communications in Medicine [DICOM]) that collectively comprise the patient record.

7.3 Agriculture

To make the right decisions at right time and to bring out best possible solutions in the field of agriculture

There will be 9 billion people that need to be fed on this planet by 2050 . To feed the world's rapidly-expanding population in the coming decades, agriculture must produce more. According to the Economic Survey of India, in 2013-14, agriculture industry contributed 15.2% to the GDP. The agriculture sector grew by 4.7% in the year 2013-14. Indian agriculture sector accounts for 54.6% of employment in the country.The 1940s to the late 1960s spanned the Green Revolution, a time when research and technological initiatives spurred the growth of agricultural production worldwide, particularly in developing nations. Some of the technologies included advanced irrigation systems, pesticides, synthetic nitrogen fertilizer, and improved crop varieties that were modified to become hybrid species so farmers could produce higher yields. The agricultural world has evolved to rely so much on seed sellers, weed-killer sellers, fertilizer sellers, and local tractor dealers.Non-availability of timely and relevant content, non-integration of services, poor advisory services, lack of localization, and in particular non-availability of agricultural information kiosks/ knowledge centres at the grass root level are some of the obstacles in the path of rural communities. Agriculture is at the heart of the social development of India as it provides livelihood to majority of the population. ICT tools should be utilized for accelerating the growth of agricultural sector which will in turn boost the economic growth of the country. e.g. in the farming, Wireless Sensor Network can be used in monitoring crops, soil moisture, nutrient content and environmental management, security and safety. GPS is a technique for remote sensing using satellite technologies, GIS, agronomy and soil sciences which are used to increase the agricultural output. This technology is useful for plotting source of produce and to map farms[32]. e.g.

China United Property Insurance signed a partnership with the National Engineering Research Center for Information Technology in Agriculture to establish an agriculture insurance technology lab to collect information through GIS technology and remote sensors to improve data for agriculture insurance[33]. The power of farming data is extraordinary. If someone knows the data of an operation, they also know when and where the crops are, how much yield, how much it costs, and the farm's profits. Tracking data on a new level would be revolutionary. A Monsanto estimate stated that planting advice could lead to a $20 billion a year increase in worldwide crop production . With the help of mobile phone applications farmers can instantly contribute to the knowledge database by uploading the crop and soil related related pictures, videos and any other important information. Farmers are not experts in forecasting parameters such as weather conditions exactly when required, sale's price for their crop produce, unexpected expenditures etc. Therefore they may be in a losing position if they cannot manage the future problems e.g. John Deere, DuPont, and Dow Chemical have joined forces to use data relayed from tractors to make recommendations to farmers on planting, seeds, fertilizers and other inputs.Precision technology is a farming management concept that measures and responds to field variability for crops, often using satellites and GPS tracking systems. A survey of soybean farmers in 2012 showed a rapid payback using these technologies – an average 15% savings on crop inputs such as seed, fertilizer, and chemicals . Precision farming along with predictive analytics could help in early warning techniques, new varieties of pest control to optimize production and regulations for quality control. Farmers can get most up-to-date farming and propagation techniques, pest control knowledge, and can also track the whole process from production, distribution to consumption. They can also leverage the systematic methods of information collection, supply chain logistics, market forecasting and business decision-making.

8. Recommendations

The success of big data for rural development depends on the participation and support of both, public and private bodies. Support in terms of finance, standards development, data sharing and access, analytical tools, and technology. Development of proposed framework in practice will demand collaboration between various government, enterprises and academic institutes to work together for the benefit of rural communities. •

Skilled solution architects with knowledge of both industry and analytics can bring added value and insight.

- Develop solutions that can provide real-time analysis of customer data streams in order to improve customer experience metrics proactively.
- Partner with niche global technology providers for analytics and Machine-to-Machine (M2M) communications to enable analytics integration with IT areas such as BI, data centre infrastructure management and building management systems.
- Enable farmers to use Interactive Voice Response (IVR) Systems, Call Centres regularly in order to solve their problems. Providing farmers easy procedures to use Data Registration (for inputs, costs, and results), Early Warning, and Decision Support System facilities through computers for good farm management .
- Different administrations should make more and more use of e-Government practices and the administrative processes should be redesigned, made more efficient, which will drastically reduce the frequency with which beneficiaries have to submit the same information to different authorities and follow up the status.

9. Limitations of Research

The use of big data is at a nascent stage in India; especially since the conceptual framework is for the rural development; it was not possible to collect primary data for research.

10. Scope of Future Research

The proposed framework is at a very high level to provide the basis for further sector specific research in advanced analytics. e.g. Identifying patients at risk of readmission can guide efficient resource utilization and can potentially save millions of healthcare dollars each year. Effectively making predictions from complex hospitalization data will require the development of novel advanced analytical models.

A detailed study would be required to explore the execution techniques in order to increase the effectiveness of analysis.

11. References

1. Belissent J, Mines C, Darashkevich Y. Smart City Leaders Need Better Governance Tools. Forrester; 2011 May.
2. Big data analytics. Ericsson White paper; 2013 Aug.
3. IDC Big Data and Business Analytics Conference 2013; 2013 Nov; Abu Dhabi, United Arab Emirates. Available from: http://idc-cema.com/eng/events/54217-idc-big-data-and-business-analytics-conference-2013/7-overview
4. Cisco VNI. The Zettabyte Era - Trends and Analysis. 2014 Jun. Available from: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI_Hyperconnectivity_WP.html
5. EMC Corporation. Digital Universe Invaded By Sensors. 2014 Apr. Available from: http://www.emc.com/about/news/press/2014/20140409-01.htm
6. Tech Republic. 2014 May. Availaible from: http://www.techrepublic.com/resource-library/downloads/how-big-data-is-changing-farming-pdf-download/
7. Mayer-Schonberger V, Cukier K. Big Data: A revolution that will transform how we live, work, and think. Hodder and Stoughton; 2013.
8. Franks B. Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with advanced analytics. Wiley; 2012.
9. Burkholder L. Philosophy and the Computer. Boulder, San Francisco, and Oxford: Westview Press; 1992.
10. Ohlhorst FJ. Big Data Analytics: Turning Big Data into Big Money. Wiley; 2012.
11. Boyd D, Crawford K. Critical Questions for Big Data. Taylor Francis Online; 2012 May 10.
12. Bryant RE, Katz RH, Lazowska ED. Big-Data Computing: Creating Revolutionary Breakthroughs in Commerce, Science and Society. 2008 Dec.
13. Groves P, et al. The big-data revolution in health care: accelerating value and innovation; 2013 Jan.
14. Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. Nat Rev Genet. 2012; 13:395-405.
15. Bollier D. The Promise and Peril of Big Data. Big Data and Health Care. 2013. p. 25.
16. Halevi G, Moed H. Evolution of big data as a research and scientific topic. Research Trends. 2012; (30):1-34.
17. Barton D, Court D. Making Advanced Analytics Work for You. Harv Bus Rev; 2012.
18. Gilpin L. How big data is going to help feed nine billion people by 2050. Tech Republic; 2014.
19. Siemens G, Long P. Penetrating the fog: analytics in learning and education. EDUCAUSE Review. 2011; 46(5):30–2.
20. Picciano AG. The evolution of big data and learning analytics in american higher education. J Async Learn Network. 2012; 16(3):9–20.
21. Retirement Agencies: Get More Than a Toe in the Big Data Pool. Accenture; 2014.
22. Manovich L. Cultural Analytics: Visualizing Cultural Patterns in the Era of More Media. Manovich; 2009.
23. Le Clair C. Electronic Medical Records Need More To Support Meaningful Use. Forrester; 2010 May.
24. Big Data for Development: Challenges & Opportunities. Global Pulse; 2012 May.
25. Emerging market analysis: disruptive opportunities flourish in india as cloud and mobility converge. Gartner; 2014 Jul.
26. CSI Communications. Applications of Cloud Computing for Agriculture Sector. 2013 Nov.
27. Centre for disease control and prevention. Syndromic Surveillance (SS), 2013 May. Available from: http://www.cdc.gov/ehrmeaningfuluse/Syndromic.html
28. Desai C. Tools for Crop Model: Review and Scope. 2013 Nov. Available from: http://www.csi-india.org/c/document_ library/get_file?uuid=c72f771c-334c-43dd-8542-bf3edb62b491&groupId=10157
29. What India's 12th National Five-Year Plan Means for ICT Vendors. Forrester Research. 2012 May.
30. Rising medical bills cripple aam admi. Times of India News. 2012 Mar. Available from: http://timesofindia.indiatimes.com/india/Rising-medical-bills-cripple-aam-admi/articleshow/12283337.cms
31. AgMRC Renewable Energy & Climate Change Newsletter, More on Feeding 9 Billion People by 2050. 2012 Jan. Available from: http://www.agmrc.org/renewable_energy/renewable_energy/more-on-feeding-nine-billion-people-by-2050/
32. CSI Communications. ICT in Agriculture – Indispensable. 2013 Nov.
33. Gartner. Emerging Market Analysis: IT. China; 2014 and Beyond, April 2014.
34. Bunge J. Big Data Comes to the Farm, Sowing Mistrust. Wall St J. 2014 Feb. Available from: http://online.wsj.com/news/articles/SB10001424052702304450904579369283869192124
35. Precision agriculture: higher profit, lower cost. Precision Ag Institute; 2012 Nov. Available from: http://www.precisionag.com/institute/precision-agriculture-higher-profit-lower-cost/
36. CSI Communications. Prospective Usage of ICT by Farmers for Agriculture. 2013 Nov.