



Mini Review

Volume 1 Issue 2 - April 2017
DOI: 10.19080/BBOAJ.2017.01.555561

Biostat Biometrics Open Acc J

Copyright © All rights are reserved by Ajit Kumar Roy

Big Data and Data Science Initiative in India

Ajit Kumar Roy*

Central Agricultural University, India

Submission: December 29, 2016; **Published:** April 24, 2017

***Corresponding author:** Ajit Kumar Roy, College of Fisheries, Central Agricultural University, India, Email: akroy1946@yahoo.co.in

Abstract

As per Gartner, Big Data is high-volume, high-velocity and high-variety information assets that demand cost effective, innovative forms of information processing for enhanced insight and decision making. Big Data is data whose scale, diversity, and complexity require new architecture, techniques, algorithms, and analytics to manage it and extract value and hidden knowledge from both structured and unstructured data. Such large data is difficult to process using traditional database and software techniques. One of the greatest scientific challenges of the 21st century is to effectively understand and make use of the vast amount of information being produced. Data analytics will be among our most important tools in helping to understand such large-scale information.

Researchers from India are active, in the fields of astrophysics, materials science, earth and atmospheric observations, energy, computational biology, bioinformatics, cognitive science, statistics etc., which generate a lot of data. To deal with these data it requires the development of advanced algorithms, visualization techniques, data streaming methodologies and analytics. Keeping in mind the momentum that big data analytics is gaining in India, there is a need to build a sustainable eco-system that brings in a strong partnership across the industry players, government, and academia. With this objective, the Indian government has launched a Big Data Initiative. Digital India is an initiative of Government of India to integrate the government departments and the people of India. It aims at ensuring the government services are made available to citizens electronically by reducing paperwork. The NIC has created the Open Government Data portal, data.gov.in. Currently 85 government ministries, departments and agencies have contributed more than 12,000 datasets across segments such as population census, water and sanitation, health and family welfare, transportation and agriculture to data.gov.in. Present government's pet project of building 'one hundred smart cities' which has been allocated Rs 7,060 crores is another major project that would rely strongly on technology, calling for a robust cloud computing backend coupled with real-time surveillance and big data analytics technologies. The UID-Aadhar project will be the largest such citizen database on the planet. NASSCOM 10,000 Startups launched a report on "Institutionalization of Analytics in India: Big Opportunity, Big Outcome.

Introduction

Over recent times, the concepts of "big data" and "big data analytics" have become very popular. The technologies that are incorporated into big data are massive parallelism, massive data storage, distribution and computing with the help of high speed networks. Data mining and analytics are not new. According to Gartner [1] information becomes big data when the voluminous data is beyond the capacity of handling with conventional database management tools and techniques. Owing to vast amount of information that companies are producing and spreading through social networks, people are now facing the challenge of processing all this data in a short time. Webopedia defines "Big data" is a massive volume of both structured and unstructured data that is so large that it's difficult to process using traditional database and software techniques. The most comprehensive definition of the phenomenon is summarized in three points known as the "3Vs" of Big Data: Volume, Variety and Velocity.

Big data is growing so much that soon it will be reaching to the largest quantities of data termed as Yotta byte that is currently

the largest unit of measurement for digital information there by marking the end of the scale. The World Economic Forum, notes in its 2014 Global Technology Reports that data is an asset. It promises a future where decisions about business, life, and society will be taken purely based on data. Thomas Davenport and D.J. Patil brought the data scientist into the national spotlight in their October 2012 in Harvard Business Review article.

Telecommunication Usage in India

India has 900 million mobile accounts, with 600 million unique mobile subscribers. It has more than 30 million PCs. It has more than 400 million Internet users. India's number of mobile-phone subscribers topped 1 billion, becoming the only country after China to cross that milestone India has millions of 3G, 4G mobile subscribers and getting ready for 5G technologies. It has more than 23million unique users on YouTube. Average time an Indian net user spends on social media is 3 hours a day and 70% of Indian internet users watch online videos. In the process, each day, our society creates huge data. With this flood of data, the need to unlock actionable value becomes more

acute, thereby rapidly increasing demand for Big Data skills and qualified data scientists. Big Data bring new opportunities to modern society and challenges to data scientists. One of the greatest scientific challenges of the 21st century is to effectively understand and make use of the vast amount of information being produced. Keeping pace with global trend the Indian government has launched a Big Data Initiative. It aims at ensuring the government services are made available to citizens electronically by reducing paperwork. The Open Government Data Platform, data.gov.in, created by NIC for Indian government departments and ministries for easy and open access by citizens. Besides pet project of building 'one hundred smart cities' calling for a robust cloud computing backend coupled with real-time surveillance and big data analytics technologies. It is believed that 30-50 million people would be actively contributing to Mygov.in, given its current pace of growth. The UID-Aadhar project will be the largest such citizen database on the planet. Aadhar is now a day's used for the direct benefit transfer for poor people. A silent revolution happening in India. Government now is capable of analyzing the satellite picture to find the truth about the progress of projects. With the realization and usage big data analytics, the demand for data scientists is rising phenomenally. More and more business schools in India are opening programmes on analytics which will help in meeting the demand.

Big Data is supposed to be a \$25 billion industry and India has the great opportunity to take a large share from it. According to NASSCOM, the Indian market touched revenues of US \$1 billion in 2014 and this is expected to grow to US\$ 2.3 billion by the year 2018. Big data analytics accounted for some 29,000 jobs in India in 2014.

English-speaking engineering, mathematics & management graduates to equip themselves with skill to grab data science jobs as India is an attractive destination for the off-shoring of big data analytics. That is why major initiatives taken by the present government to uplift the analytics sector that is going to create huge job opportunities in data science and big data analytics [2-5].

Recent Development in India

Indian Government Initiative for Big Data

Digital India Programme: The Government of India has launched the Digital India programme with the vision to transform India into a digitally empowered society and knowledge economy. Therefore, Ministry of Communications and IT initiated to transform the entire ecosystem of public services using information technology.

What is Digital India? Government of India has taken initiative to digitize India by enabling Internet access available to over two lakh villages by 2019, promoting e-governance, e-banking and similar others to lead India into a knowledge

society. It will generate 18 lakh jobs. Prime Minister Narendra Modi's Digital India programme is like a major initiative for rural India that suffer from lack of connectivity. 'Digital India' is a central programme equip Indiato face the challenge of upgrading to a knowledge-based society. This is a vast initiative on using technology to create a participative, transparent and responsive government.

What does it aims to achieve?

The 3 major targets of the programme are:

1. To extend high-speed internet, mobile phone and bank account enabling participation in digital & financial services.
2. The programme aims to take digital literacy to the next level, and will focus on encouraging citizens to slowly move to cashless monetary transactions.
3. It aims at seamless integration across departments/ jurisdictions, and ensuring availability of facilities in real time from online and mobile platforms

The vision of Digital India is keyed on three key areas:

- a. Digital Infrastructure as a Utility to Every Citizen
- b. Governance & Services on Demand
- c. Digital Empowerment of Citizens

Six of nine Pillars of Digital India:

- a. Broadband Highways
- b. Universal Access to Phones
- c. Public Internet Access Programme
- d. e-Governance – Reforming government through Technology
- e. e-Kranti – Electronic delivery of services
- f. Early Harvest Programmes

The government has launched some initiatives in line with these objectives

- a. Digital Locker System aims to minimize the usage and sharing of e-documents online.
- b. The mobile app for My Gov would bring many features to users on a mobile phone.
- c. Swachh Bharat App to be utilized by the citizens and administrators for achieving the goals of Swachh Bharat Mission.
- d. eSign framework which would allow citizens to digitally sign a document online using Aadhar authentication.
- e. Online Registration System (ORS) has been introduced to provides important services such as online registration, payment etc.

- f. National Scholarships Portal is a one stop solution for end to end scholarship process.
- g. Digitize India Platform (DIP) would facilitate efficient delivery of services to the citizens.
- h. Bharat Net, a high-speed world's largest rural broadband connectivity project using optical fibre aims to connect almost all villages of country.

BSNL has undertaken large scale deployment of Wi-Fi hotspots throughout the country [6]. The government has been a driving force in creating platforms and large datasets which will require the acquisition and manipulation of massive amounts of data. It will provide e-governance platforms. Digital India is an initiative of Government of India to integrate the government departments and the people of India. It promises the government services to citizens electronically by reducing paperwork [6].

Governmental initiatives in big data collection and usage

People belonging to marginalized sections of society in India often do not have a valid proof of identity. Thus, they miss out on availing social benefits provided by the government. To overcome this Indian government launched a scheme to provide a unique 12-digit number ID, termed 'Aadhaar' to every resident of India. It is an identification that a person can carry for a life time to use any services like identity, bank transactions etc. Aadhaar is considered as the world's largest ID platform. Since the first set of Aadhaar numbers were issued in September 2010, 883, 644, 522 identification numbers have been issued (as of 15 July 2015; some 20 million enrolments take place per month [7]).

It is also the largest biometric programme in the world, as biometric data of each person is captured for all practical purposes. This unique ID is now used for Gas subsidies all over the country. Digi Locker provides a personal storage space in the cloud to Indian citizens. Organizations that are registered with Digi Locker can push electronic copies of documents and certificates (e.g. driving license, Voter ID, School certificates) directly into citizens' lockers. Citizens can also upload scanned copies of their legacy documents in their accounts signed using the eSign facility provided in Digi Locker. A citizen can provide that without having to submit paper copies. Since its soft launch on 10 February 2015 (the official launch took place on 1st July 2015), over 825,000 users have registered to use the digital lockers, with over 1.5 million documents being uploaded.

Demonetisation: In an address to the nation at midnight on November 8, India's Prime Minister, Narendra Modi, declared that the two highest denomination currency notes—the 500 rupees note and the 1,000 rupees note—won't remain legal tender. These notes were banned to prevent the illegal financial transactions like Fake currency, Black money, Corruption and Money laundering which has been in India for many years. The move sucked out about 86 percent of the Rs. 15.44 lakh

crores cash in circulation. This resulted in a cash crunch. The demonetisation initiative brought out a fair share of criticism and acclamation. Many welcome the government's decisions to demonetize the currency notes in circulation in the economy as bold and revolutionary as in a single master stroke, to combat parallel economy, counterfeit currency in circulation and terror financing. On the other hand, critics argue that it has struck a body blow on economic activity in India. The outcome of the whole exercise cannot be determined now, some experts say, pointing out that demonetisation is still a work in progress. Post demonetization cash position was getting better day by day and the rush at ATMs had come down significantly. Furthermore, Income tax raids and cash seizures are a common feature today. The Enforcement Directorate raided several forex establishments making back dated entries and money laundering using backdated accounting was carried out by co-operative banks, jewelers, sellers of iPhones, and several other businesses. While the pros and cons of the measure continue to be debated, the consensus appears to be that people at large has supported the move. Demonetisation may wipe out the present stock of black currency but problem is to unearth assets such as gold & Jewelry and real estate investments. Radical action, to move toward a less-cash economy, might have been better since we know that high-value banknotes in many countries are primarily used for criminal purposes. Today is the era of big data analytics utilizing Open Source Real-Time Big Data Analytics tools a lot of insight can be revealed to check black money [8].

Cashless Economy: To combat black money, corruption, counterfeit currency and terror financing, government took decision to ban Rs 500 & 1000 notes followed by putting a cap on withdrawals from banks and ATMs, Prime Minister Narendra Modi initiated a notion of a cashless economy. A cultural-economic revolution in the making. Cash is more expensive for exchequer, because of tax evasion, corruption and the need to keep recirculation old, spoilt, currency, and enabling transfers. At the same time, digital is very expensive for citizens. of data. A digital economy is an economy which is tracked in real time. Each transaction is mapped facilitating cleaner economy resulting in an improved climate for foreign investment and boost economic growth. Focusing on Digital Economy for speed, accountability, and transparency Government has pushed the digital theme in every area of the budget. India declared the three-point agenda for Budget-2017 aiming to Transform, Energize and Clean India. India has promoted BHIM app, including cash back scheme for merchants; Aadhaar Pay to be launched for people who don't have mobile phones.

Nobel Prize-winning economist Joseph Stiglitz said at the World Economic Forum meeting in Davos, Switzerland, that the benefits of demonetization outweigh the cost. 'Cashless economy is a boon, Muhammad Yunus, founder of Grameen Bank and Nobel Peace Prize winner commented. Cashless transactions will 'curb generation of black money; reduce real estate

prices; saving that RBI spent Rs 27 billion on just the activity of currency issuance and management in Financial year 2015; pave way for universal availability of banking services to all as no physical infrastructure is needed other than digital to check soiled notes or counterfeit currency; reduced costs of operating ATMs;. Availability of internet connection and financial literacy may pose as a problem.

The Budget announcements by finance minister Arun Jaitley clearly reflect Modi's confidence that the aam aadmi would support the "bigger moral purpose" behind demonetization. applying Analytics 3.0, and Big Data Analysis Platforms & Tools like (Hadoop, Map Reduce, Grid Gain and HPCC many hidden fact can be unearthed facilitating cost reduction and faster, appropriate policy decisions to smoothen the goal of cashless economy [9].

Big data application in healthcare-indian scenario

According to Prof Raghupathi, "Healthcare data is rarely standardized, often fragmented, or generated in legacy IT systems with incompatible formats." This is one of the biggest challenges in India. Now a day's huge data are collected by hospitals, primary care providers, researchers, health insurers, and state and central governments. Each of these acts preventing data transparency across the healthcare system," says Ramaswamy. Another challenge would be veracity of data. "Different types of data from different systems, adherence to standard formats, inter-operability issues and homogeneity would also pose a great challenge," says Ramaswamy. In addition to aggregating a massive amount of data, there's the challenge of maintaining patient privacy. Healthcare providers can analyze patient history data, real-time data from monitors, clinical factors, lifestyle choices and social determinants to provide a holistic view of the patient and develop the most effective care plans in addition to identify, predict and healthcare plans [10].

Urban Sciences, Big Data and India's Smart City Initiative. Cities are repositioning themselves to play a pivotal role in the development of humanity; though because of rapid population growth and nonstop urban expansion; cities are stressed with a variety of challenges related to urban life such as urban planning and management, environmental management, urban safety, resource mobilization and utilization, urban health, energy efficiency, traffic management, social activities, recreation and entertainment. Letting down to cope with any of the aforesaid challenges might be a threat to the city's prosperity and quality of life affecting its residents adversely hence 'Smart City' concept has been materialized as a problem-solving technological instrument to real urban world problems. Smart cities are also recognized as lively cities which can respond to resident's basic needs and aspirations in 'real time's.

Future of Big Data in India

Presently India is having a large pool of English-speaking engineering, mathematics & management graduates. They are

equipping themselves with skill to grab emerging data science jobs. India is an attractive destination for the off-shoring of big data analytics. Indian entrepreneurs are in the process to exploit as well as major initiatives taken by the present government to uplift the analytics sector that is going to create huge job opportunities in data science and big data analytics. Besides of widespread and constant use of telecommunications devices huge data are continuously generated from Global Positioning System (GPS) devices, automated teller machines, scanning devices, sensors, mobile phones, satellites and social media. With respect to Job creation and opportunities of data scientists in India and abroad suggested that the Analytics and Big Data sector has seen a consistent growth in the last five years and is expected to further grow at a CAGR of 33.2% and 26.4% respectively, almost six to eight times that of the overall IT sector. According to NASSCOM, the Indian market is expected to grow to US \$2.3 billion by the year 2018. Big data analytics accounted for some 29,000 jobs in India in 2014 [11].

Experts find potentials in Big Data Analytics and concerned about problems, too. Along with potentials of big data analysis, some fear that work with gigantic stores of information could lead to privacy abuses and mistaken forecasts. But while leading technologists and researchers around the world look forward to the positive impact of Big Data, many argue about potential drawbacks. Big Data is leading to knowledge revolution that's spreading, almost invisibly, through various activities of everyday life.

Risks and Challenges

Big data's promise has been met with warnings about its cons also. Perhaps the most severe risks and most urgent avenues for research and debate are to individual rights, privacy, identity, and security [12].

References

1. <http://www.gartner.com/technology/home.jsp>
2. <https://data.gov.in/>
3. Dave Evans (2012) How the Internet of Everything Will Change the World for the Better #IoE [Info graphic], CISCO BLOGS.
4. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012-2017, at 1 (2013), available at
5. Omer Tene, Jules Polonetsky (2013) Big Data for All: Privacy and User Control in the Age of Analytics, 11 NW J. TECH. & INTELL. PROP. 239: 270-726.
6. <https://fusion.werindia.com/hot-from-the-oven/digital-india-integrates-citizens-records>
7. <http://www.linkaadharcard.com/basic-question-about-digital-locker/>
8. Roy Ajit (2017) Facts and Figures of Demonetization in India - Views, Reactions and Impact: Impact of demonetisation Kindle Edition
9. Roy Ajit (2017) Cashless Economy in India- Present Scenario, Potential, Prospect and Challenges (Demonetisation in India Book 3)
10. Raghupathi W, Raghupathi V (2014) Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2: 3.

11. <http://www.nasscom.in/nasscom-big-data-summit-redefining-analytics-landscape-india> 12. Jules Polonetsky, Omer Tene (2013) Privacy and Big Data.



This work is licensed under Creative Commons Attribution 4.0 License

**Your next submission with Juniper Publishers
will reach you the below assets**

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats
(Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission
<https://juniperpublishers.com/online-submission.php>