BIG DATA INTEGRATION ON CENTRAL AND STATE GOVERNMENT SECTORS

G.Thangaraju*

Dr.X.Agnes Kala Rani**

ABSTRACT

The term Data is nothing but a raw fact, it is growth in the computing technology as a information, database, DBMS, RDBMS, ORDBMS, DDBASES, DWHDB and the current era in the field of Big Data. The collection and processing of large amount of data is not new in the field of database, what is new is the speed at which we can now process that data. This is an upcoming of collective intelligence. Every day we create large no data's approximately around 2.5 quintillion bytes of data, in feature we are dealing with more no of huge data in different situation, and that lead us to work with term that is called "Big Data". The Big Data is defined as any data that can challenge our current technology in some manner can consider as Big Data. The Big Data are high-velocity, high-volume, and or high-variety information assets that require new forms of processing to enable enhanced decision making, insight innovation and progression optimization. In This presentation we are dealing with the challenge of Big Data in Government Sector especially an individual household or farmer to met their task to the government via Big Data Technology.

Keywords- *big data; big data in government; big data applications;*

* Research Scholar, Assistant Professor, Department of Computer Applications, Thanthai Hans Roever College, Perambalur-621212

** Research Advisor, Karpagam University, Coimbatore-641 021

1. INTRODUCTION

Big data is one of the most challenging words all around now, applications are generating massive volumes of unstructured data that describe user behavior and application performance, Today, most companies are unable to fully capitalize on this potentially valuable information due to cost and complexity, How do you capitalize on this raw data to gain better insights into your customers, enhance their user knowledge and ultimately improve prosperity? The data can come from many different sources like click stream, sensors, and also many other places, it must be processed speedily and which is too hard to be managed by all the existing tools to process those massive amount of data records. Similar way not only the giant companies deals with the huge amount of data but even the administration of any Country are canonical users of Big Data, as they also keep track of huge number of different records for their country which may include information about the people of the country, their growth and many more belongings. Here I am going to focus on some of the application of the Big Data in the field of Government and also Public sector and the society in general.

2. GOVERNMENT OF INDIA AND BIG DATA

Government of India has consists a number of departments; in some of the department can have the large data for their proper functioning and proper implementation of their policies and schemes. The GOI can handle the Big Data in different department like

1. The Ministry of Rural Development
2. The Election Commission of India
3. The Ministry of Human Resource Department
4. The Planning Commission Government of India

2.1).The Ministry of Rural Development

It can provide one of the major service to the rural peoples ie. Mahatma Gandhi National Rural Employment Grantee Scheme (MGNREGS). [1]

Figure 2.1.Home page of http://nrega.nic.in

Anyone can access the information about the MGNREGS with the Public Data Portal with respective options chooses by the user. The household of Indian nation can access their MGNREGS data from the website like no. of days worked in the current financial year , how much amount received from the bank like any queries done by the large no. of users at a time the server can response to all the users.

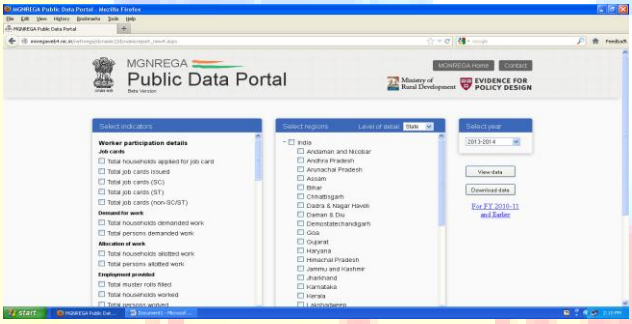

Figure 2.1.a.Home page of http://nrega.nic.in public data portal

2.2).The Election Commission of India

In this department also maintains the big data. Like the voters ID, booth name, Election manifesto of all states and all other [2]

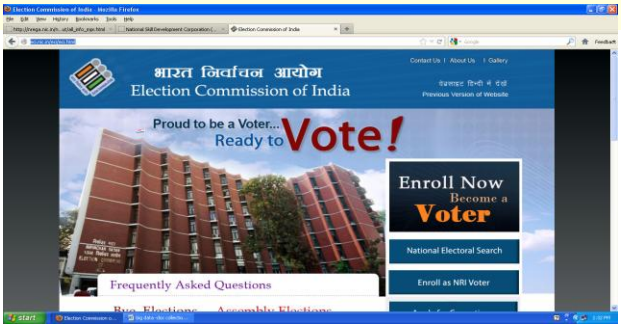

Figure 2.2.Home page of Election Commission of India

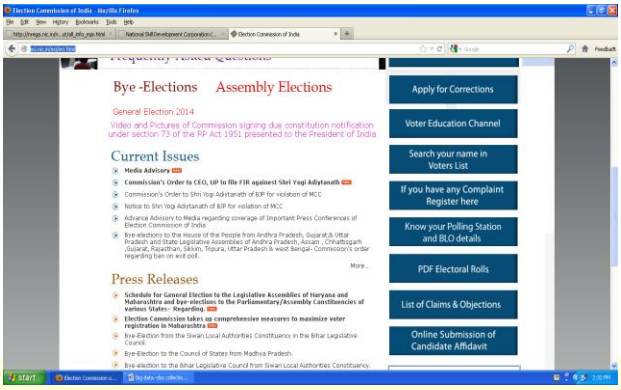

Figure 2.2.a Home page of Election Commission of India

2.3).The Ministry of Human Resource Department

In this department also play a vital role in the field of big data i.e. This department provide a technical skill to the students via the National Skill Development Corporation (NSDC) via the STAR Scheme.[3]

Figure 2.3 Home page of Ministry of Human Resource Department

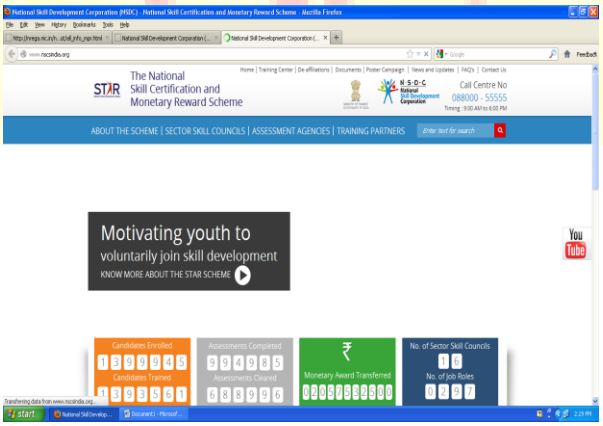

Figure 2.3.a Home page of Ministry of Human Resource Department

2.4). The Planning Commission Government of India

The Unique Identification Authority of India (UIDAI) is an agency of Government of India, which was established in the year 2009, it provides a card named as AADHAAR is delivered to all the people in the nation this process is going on till the progress is 6o % is over remaining processes is going on and monitored by the respective department. This scheme also one of the big data handling scheme. [4]

Figure 2.4. Home page for the Planning Commission Government of India (UIDAI)

The UIDAI will own and operate the unique identification Number Database. The authority is planning to give the unique identification numbers to all Indians, but won't be provided that the smart cards, that would depend on to the ministries of the country. The authority will maintain a database of residents containing the biometric and other data. Mainly the database would be containing the 12-digit unique number which is going to be issued by the Unique Identification Authority of India (UIDAI) for all the residents, all this numbers will be stored in the centralized database and will be linked to the basic demographics and biometric information photograph, ten figure prints and iris of each individual. Random numbers generated will be free from any classification based on caste, religion and geography. This database would be very large and big but this will allow those people to have identity who do not have any sort of identification [4]

AADHAAR is an ambitious government Big data project which is going to be one of the largest biometric database in the world by 2014, with the global capturing about six hundred million Indian Identities [6]. And dealing with this kind of database is not an simple task, but India is going to be one of the first countries who would be dealing with this much large amount of public biometric database.

3. BIG DATA INTEGRATION IN GOVERNMENT SECTOR

3.1.MGNREGS and Respective Bank of Household

MGNREGA has MIS Viz, NREGASoft; the architecture of the NREGASoft is equipped to capturing all the transactions in almost real time. One of the most important modules is the Electronic Fund Management System(e-FMS).

The e-FMS operates on principle of decentralized decision-making and centralized fund management. Under this System, every expenditure point is electronically linked to the central server through which the implementing agencies can generate fund transfer orders (FTO) which automatically pull the required amount from the State Fund and credit the account of the end users(wage seekers or suppliers).

e-FMS system not only promises reduction of delays in payments, corruption, parking of funds but increases accountability, transparency and optimization of resource utilization.

Steps involved:-

1. State will open/ declare a Nodal Agency Account preferably at State or at District level in a Commercial Bank (Sponsor Bank). This will be the Debtor Account. Beneficiaries to whom payments are to be made may have their account in any of the Commercial Bank/RRB, Co-operative Bank or Post Office.

2. Each Nodal Agency will register their nodal agency account in NREGASoft.

3. Each FTO from that Nodal Agency will contain that nodal agency account as Debtor Account and many credit accounts in different Commercial Bank/RRB.

4 . The FTO will be shared with Sponsor Bank having nodal agency account.

5. NREGAsoft has setup SFTP server which act as a communicating point for banks.

Fund Transfer Order after dually digitally signed will be placed on this SFTP server in different folders. Four folders will be created for each Sponsor Bank. The name and purpose of these folders are as follows:

a. **FTO-<bank-short-name> (eg. fto-sbi):**

This will contain all FTOs generated for payment with debiting account (nodal agency account) in State Bank Of India. NREGAsoft will place these FTOs and Nodal agency Sponsor bank will pull it. Each FTO will be digitally signed by two signatories of MGNREGA.

b. **FTO-ack-<bank-short-name>(eg. fto-ack-sbi)**:

This will contain all responses of fate of transactions from nodal agency Sponsor Bank. The response will be generated twice, pre-crediting and post crediting the account of worker. Sponsor bank will place these responses and NREGASoft will consume the same and reflect it in its MIS. Each response will be digitally signed by authorized person from Sponsor Bank.

c. **FTO-failed** : This will contain all FTOs which are found tampered/corrupt by Sponsor Bank. Sponsor Bank will place these FTOs and NREGASoft will consume.

d. **account statement**: This will contain account statement of the nodal agency bank account. Nodal agency account bank will place the statement and NREGASoft will take. Account statement will be digitally signed by authorized person from bank.

e. Digital Signature Certificate (DSC) management:

i. DSC of authorized signatories will be shared with bank through web services.

ii. Bank will share the authorized signatory of bank to NREGASoft through web service.

f. Response received from bank will be incorporated into NREGAsoft database at 10.00 a.m. and 5.p.m. daily.

Note: Sponsor Banks must automate the above to avoid any delays in the payments. The response of the transaction must be sent latest in T+1 days without any default. Otherwise the responsibility of any further delay will be with the Banks.

6. Bank will identify a machine which will communicate with NREGAsoft SFTP server and give t h e **Real Static IP address** of that machine to NIC (nicdrd@nic.in). NIC will get the required port open for that machine and will confirm through reply all email

3.2 Integration of Election Commission of India and Revenue Department of respective state also state election commission.

The Birth and death registers are maintained by the department of revenue and department of health in district level. The automatic updation of voters list of both ECI and State Election commission is possible through the big data integration technique. Here the AADHAAR Number has play vital role for the integration of all Department.

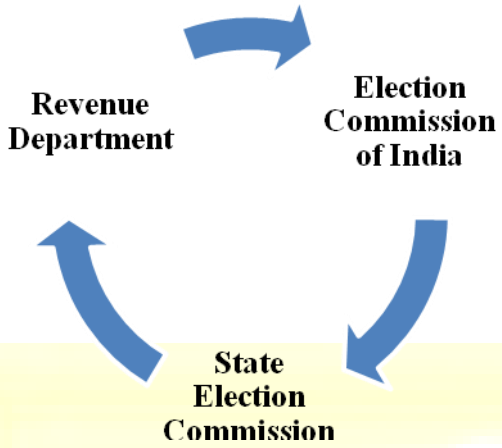

Figure 3.2 Integration of the three departments

3.3 Other Departments:-

Government of India handle their polices and process via in some other departments and schemes like [

a).Direct Benefit Transfer Schemes

b). Government Construction Projects

c). Education Department

d). Health care Department

e). Generating revenue from Government sites

The Direct Benefit Transfer Scheme This scheme is dependent on AADHAAR card that will help people to get benefit of this scheme and it will also ensure the government about the number of people who are getting benefit out of this, it includes various different schemes under it like education, scholarships etc.

4.BIG DATA INTEGRATION TOOLS

EMC consultant David Dietrich says a dirty little Secret of Big Data is that even when you're dealing with the industry's top minds, data preparation still takes up a huge amount of any Big Data project.

He lists five available technologies for smarter data preparation work:

Data Tamer, which focuses on integration and is still being developed at MIT. **Open Refine**, formerly Google Refine, which helps with clean-up.**Data Wrangler,** a cleaning and transformation tool developed by Stanford. **Reshape2** packages, which let you restructure and aggregate data. **Plyr**, which uses a split-apply-combine strategy for R. [6]

Big data has the potential to transform business, improve lives, and change our world — but your company needs to be able to unleash its potential while minimizing the risk associated with new technologies. Informatica Big Data Edition provides a safe, efficient way to integrate all types of data on Hadoop at any scale without having to learn Hadoop.[7]

5. CONCLUSION

Data Integration , Reporting, Decision making are important in computing technology and also the Big data Integration is very essential in the field of Big Data. It can be really helpful if the data or records are analyzed carefully and if we use that statistics in a right way, it can really help the developing country like India.

6.REFERENCES

[1] http://nrega.nic.in

[2] **http://eci.nic.in/eci/eci.html**

[3] http://www.nsdcindia.org/

[4] "Unique Identification Authority of India" en.wikipedia.org/wiki/Unique_Identification_Authority_of_India

[5] Applications of Big Data In Government Sector HITESH H. PARMAR 1st year M.E Student, Department of C.E, M.E.F.G.I, Rajkot, Gujarat, India. International Conference on Big Data Management & Cloud Computing, 28th July 2013, Bhubaneswar, ISBN: 978-93-83060-09-2

[6] http://www.itbusinessedge.com/blogs/integration/five-new-tools-for-smarter-big-data-integration.html

[7]http://www.informatica.com/in/products/big-data/informatica-big-data-edition/

7. AUTHOR INFORMATION

1.Mr.G.Thangaraju received his M.Phil in computer science from Mononmaniyam University, Thirunalvelli, India in 2004. Pursuing Ph.D in computer science at Karpagam University, Coimbatore, India. Currently he is working as Asst.Professor in the Department of Computer Application, Thanthai Hans Roever College, Perambalur, India. His current research interests Data warehousing, Distributed Data Base and software metrics.

2. Dr.X.Agnes Kala Rani, Research Advisor, Department of Computer Applications , Karpagam University, Coimbatore. Tamil nadu, India.