

Big Data Challenges for e-Government Services in Nepal

Pratima Pradhan¹, Subarna Shakya²

¹Singhania University, India

²Department of Electronics and Computer Engineering, Pulchowk Campus, IOE, Tribhuvan University, Nepal

Corresponding authors: *pratima.pradhan@ntc.net.np* and *drss@ioe.edu.np*

Received: Jan 27, 2018

Revised: Feb 17, 2018

Accepted: Feb 22, 2018

Abstract: With the digital disruption governments all over the world have envisaged into big data technology with the purpose of seeking values for the benefit of governmental activities. Governments have taken up this technology as a game changer of the e-governance. Citizen centric data over the web access and activities over the smart handsets have driven large volume of data not only in structured form but largely in unstructured form. The necessity of processing and analyzing the exploding data sets has driven governments to adopt this technology. This paper studies the possibility of digitalizing Nepal e-government services with innovative problem solving measures by adapting the Big Data applications. Following the global trend of digitalization and the ICT infrastructure development, Nepal e-Government services can initiate Big Data technology to resolve many of the governmental services. This paper explores some big data applications for the e-government of Nepal and further studies some challenges that Nepal could face while adopting the Big data in the e-government services.

Key words: e-government, big data, analytics, data, ICT

1. Introduction

The changing role of government to involve citizens to resolve problems and innovate utility activities has realized to enable government to adapt to quick decision driven big data technology. The sources of data from crowd sourcing, internet of things, public private partnership programs, social network activities and sharing of data has boosted the compulsion of this big data system. Following the suit of developing and developed countries, the Nepalese government service delivery can be made efficient and transparent by introducing big data projects. However, with the growth and speed of ICT delivery, there also come a lot of challenges like security, data quality, license, and accessibility, cross boarder country business, re-usability issues, ownership conflicts, readiness to integrate and storage.

Digitalization and ICT Development: With the global ICT development, newer citizen services are being deployed to provide governmental services. The trend of ease of use of governmental operations has encouraged the use of these services. Citizens have become accustomed to using ICT in their daily life styles hence making it less hesitant to use them. Smart cities are rapidly

converting the citizens to get smarter in many ways. The increasing penetration of mobile phones with the usage of large number of smart phones tells how citizens are forced to get smarter with being mobile and socially connected. Mobile applications are finding a large innovative space to deliver newer exciting applications over the devices. Technology affordability has increased with the cost going down every year. ICT development Index (IDI) set by United Nations International Telecommunication Union (ITU) is a benchmarking standard for measuring the information society. ICT index in the year of 2016 for the top countries are Korea (8.84), Iceland (8.83) and Denmark (8.74). For countries of South Asian Association for Regional Cooperation (SAARC) the ICT index are: for Afghanistan: 1.73, India: 2.69, Pakistan: 2.35, Sri Lanka: 3.77, Bhutan: 3.74, Maldives: 5.04, Nepal: 2.5 and Bangladesh: 2.35. Mobile technology with the deployment of 3G and LTE has brought about the fast deployment of data, resulting to higher ICT penetration. Smart phones users have grown abundantly due to cost reduction and hence the explosion of usage data volume. Broad band and internet access have become cheaper with more affordability and togetherness the demand for higher consumption. This ITU report states that communication and social media are key activities for Internet users. In order to collect the missing data gaps for studying ICT deployment accurately ITU has launched project on "Big Data for Measuring the Information Society" [6]. In the global context, mobile users are growing more and more and reached 7.6 billion in Q1 of year 2017. LTE subscriptions were 2.1 billion in Q1 2017 [4]. The IOT (Internet of things) devices connecting the ICT infrastructure is exploding with the drop of device pricing. Innovative applications or use cases are increasing the popularity of such devices. This report has estimated that around 18 billion IOT devices will be seen till 2022. Connected IoT devices will be used as sensors, for cars, point of sales machines, wearable and consumer goods. Cisco has also forecasted M2M (Machine to Machine) devices will grow from 5.8 billion in 2016 to 13.7 billion by 2021 [2]. Among the IP traffic is a massive portion of video traffic and this report has stated that it will account to 82 percent by 2021. Besides online wireless technology contributing to Ip traffic, there is also the broadband technology with increasing deployment of better speed and attractive contents that contribute to larger Ip traffic. This report has further mentioned that the global average broadband speed will double from 2016 to 2021.

Benefits of Big data in digitalized e-Government: The more the digitalization the more data is produced. The value of data is realized with the success of Big data technology which can process the 4Vs (Velocity, Volume, Variety and Varsity) of data characteristics. The users fuel the exponential data growth. Face book has 30 billion pieces of content shared every month [8]. The more the success of the use of the big data the more there is collaboration of services in the e-Governance. The sharing of the information in the two-way communication in the four modes G2G, G2C, G2E and G2B makes it possible only with the big data technology, hence the increase in the sources. Optimally used, the values of data can bring about transparency, organized or collaborative informative system, speed in corrective decision making and newer business opportunities. MCKinsey estimated that developed economies of Europe's government administration saved more than \$149 billion in operational efficiency improvements alone by using big data [8]. Big data can improve decision making and also organizational efficiency. The drill down real time capability is the output of the analytical systems used after mining the collaborative data. Descriptive reporting on the statistics of the history events make proof on what is going on. Predictive analytics assist to forecast and come up with statistical models used for further actions. Prescriptive modules can focus on optimization and randomized testing. These analytics can evolve the traditional services into transformational government (t-government) which is the evolutionary state of the e-government [9]. It is here discussed that the t-government can occur

through aggregation (with other cooperative government alliances and agencies), syndication (to improve service delivery), consumption (external data sources) and co-creation (innovating with external partners).

2. Big data Applications for E-government

Nepal can benefit largely by applying big data applications. Although there a wide range of big data applications Nepal government can set up in priority wise. The first big data application can be the unique citizen identification card for each resident or even each citizen. In India the "Adhaar" card [3, 5, 10] was introduced as a unique identifier for transparent citizen benefits. This card could hold the key to verification for multiple purposes. Nepal could benefit for passport, taxation, and license and citizen benefit distribution. This could be consolidated so that same information like photographs, citizenship number, address, and other personal profile could be shared form single source instead of entering the same information multiple times. Big data for taxation is leveraged by developed countries. Taxes evasion is highly known for weak tax administration. Many countries have started program for better tax collection mechanism. ICT tax management has been adopted by 159 countries of the UN member states. China is using the big data for taxation in several ways including the utilization of data with multi and cross referencing to verify taxation and tax related business processes. Brazil is another country adopting the big data systems for taxation. The U.S.A. electronically submitted taxing allows a government to study the IP addresses of the transactions in order to see if conflict resulting to fraud arises [11]. Like U.S.A. ties the social security number with tax control, Nepal can also use the unique citizen identifier to control and track the taxation. By adopting the big data technology in Nepal the Tax evasion and fraud will be minimized when consolidated into one system. Disasters are common natural phenomenon in Nepal. The country suffers from floods, earthquakes and landslides. The aftermath situation of the 2015 April, devastating earthquake in Nepal has still not been free of issues. Nepal could benefit largely with the big data application on how reliefs and benefits are distributed and to study the economic damages. Data science methods for extracting patterns from natural disaster data can provide insight for understanding the economic and human impact of natural calamities. Recent advances in cloud technologies and numerous open source tools enable this analysis with no initial infrastructure investment. This allows governments with limited resources to build systems based on data science and develop more sophisticated models utilizing higher volumes of data from a variety of sources to analyze their most challenging issues [7] The disaster relief can be faster with social media activities. The social media is integrated along with offline and online activities and included on the ground activities. Nader has emphasized the importance to involve social media in advance to the disaster, so that disaster management can be employing social media in their recovery processes and use it as a tool for recovery efforts. Resource allocation, discussions communication should be part of the social media tool for such times. Including surveys, interviews along with social media and content analysis by big data help to understand the recovery service delivery. Required skills for deeper understanding must be realized for efficient network dynamics. Location based information assist the recovery processes [1]. Even social media analysis can be used to help understand and identify the possibility of recruiting volunteers needed for disaster recovery processes [12]. Another big data application can be used in Nepal is the tourism section. Tourist data base and tracking system by mobile phones can be used to understand the whole picture of tourism industry. Hotel information, destination preferred, their social networking study, weather reports, preferred routes for travel, trekking security and facility can be enhanced with the big data analytics. Nearby Facilities and security alerts can be pushed to tourist according to their footprints collected from the mobile phones.

3. Challenges of Big data in e-governance of Nepal

The most important goal of the big data application is the attitude of the governmental bodies. The will to disrupt the existing system and digitally transform the system should come with a positive response of the question: "Will the governmental stake holders be willing to put information with transparent accuracy?" There lies an urgent need for not only the decision level to understand the service delivery by the big data system but also the understanding of the influential lower level and the IT departments. The mandatory projects must be set in Nepal's Strategic Development Goal guideline. The sustainability goal must be set prior to the project design. Responsibility must be fully taken up by the governmental houses from the start of the project design to the continual operability of the system.

Yet to be digitalized: Historical files are very necessary to be included in the analysis. In Nepal, much volume of data stays back in documentation paper form. To be digitalized is a great challenge. This may lack the attitude as well as incur large cost regarding to digitalize and systematize the data. Staffs that are used to paper works may not be able to adapt a new way of working. This has to be under the process of an organization digital transformation procedures. Another challenge is the continual effort and monitoring of the system of digitalization of the documentations. If not strongly monitored the procedure will be kept aside after some time and the before process of paper based work will be taken up again. The mandatory process of digitalization must be the only method of work flow.

Collaboration of data: Governmental data is collective and have values of being cross correlated with several governmental data. Such collaboration may be a first step of discussion which may even distract the project of the big data implementations. Collaboration of data is not just at the start of the project. The requirement of data being collaborated should be continuous during the operational stages. Newer and newer data sources will be a need requirement in the later days also and must be taken up to avoid misunderstanding with various governmental organizations. The initial collaboration must be thought at the beginning of planning of big data project for the e-governance. Many governmental departments will not be sure or even reluctant to integrate the data. The bureaucrats have to be educated on the project benefit and their continual supports. Sources of data must be started small. Willingness to collaborate timely between departments of governmental sectors must be agreed from the very beginning of project design. Sharing of corrective and clean data must be pictured well.

Clean Data: The success of ingestion of data and analytics lie highly on the intense of the data being clean. Data obtained from various sources may be replicated or even mismatched. The analysis and the reports will waver from actual meaning and actual interpretation if the attributes are misunderstood. Cleaning of data need skillful insights and must be dealt in caution.

Trust: Trust in the system lies in what values are produced to the users. Trust is also dependent upon the latest updates, clean and reliable data and useful data. The availability of the system must be very high. Trust and frequency of use depends largely on the ease of use. The extent of the trust to use depends on the data reliability and the usefulness. Data errors like inconsistency, duplication or missing will create mistrust in the user. If citizen do not understand how to use, it becomes a challenge in the sustainability of the system. Usage can be increased by applying technology in smart phones by innovative device applications. Trust of data is also related to personal information being watched. To increase the usability the feedback mechanism from the users must be part of the monitoring and operation of the system. Newer requirement can attract the citizens from

using it. Hence, surveying and understanding must be a continual process for sustainability of the e-government big data application. It becomes a challenging part in the context when Nepali language is used in the input variables.

Ownership: 'Who owns the data', can be an issue right from the source of collection. The data can be handed or shared over for re-usability within the government departments and also with other business houses. As the big data application increases, more and more data values will be sought with more integration and sharing of data. Ownership issues must be taken up with great sensitivity. The graveness of the issue will be noticed when the privacy of the data is taken up. Some departments may take this lightly when sharing the data and cause exposure of privacy data. Ownership and how to use data must be clearly stated in the sharing of data licensing documents. Controlling of re-use of data can be difficult without strong monitoring from the point of collection. The point of the collection must be careful when handing down the data to other parties.

Responsibility: Recruitment and retaining of technical staffs would also fall into the higher priorities. The ownership and responsibility of technology must be outlined for long time service delivery. There will be requirement of up-gradation or even replacement of technology modules as time goes by. Here, it becomes important to design the project, keeping in mind of the long period sustainability of the service delivery. Technology handling departments must be set up at the beginning of requirement analysis and continued throughout with local technical experts and knowledge pass-over environment. Hence, the ownership of system and responsibility must be clearly outlined with the setup of the systems data security, analytics and monitoring mechanisms. Standardization of services with monitoring Service level agreement (SLA) and Key performance Indicators (KPIs) must be kept at the start of the draft of the plan.

Privacy: Usage of personal data is a sensitive issue in many countries. However, personal information is often needed to understand the business values. Like location record and transaction details of users are a must to analyze the behavior context of the end users. Heat maps are based on usage patterns of individuals. Businesses are predicated and solutions are prescribed on the pattern analysis of individual records. So footprint of the movement of users and their usages are thus known to the analytic system and their experts. When, data are re-used, these data with private information are handed down to the next business party for their business benefits. The private information may include phone numbers, age, gender, addresses, email addresses, their website surfing habits, their bank information and other behavior patterns. The sensitivity of privacy depends on the countries' way of seeing their transparency in the dealings of e-governance activities.

Cost: The hesitance of transforming the e-governance with big data application comes with the big question of financing the project. Definitely to implement some system the cost is always a first issue. Big data projects are not expected to be a 'big bang' success projects. The goals of the big data can be started with keeping it small. The projects must be designed keeping in mind of the challenges of the big data in e-government. Looking into the global project benefit it is highlighted that cost for big data project in terms of hardware, software, licensing and operational is lesser than conventional data analytics. Since, hardware is highly scalable for big data system, it is essential to start a project in small scale and expanded slowly.

Data Scientists: Data scientist is the key role to the success of the big data applications. Knowledge of system is essential right from the beginning of the planning of the system. Data experts must be able to sort out what is essential and what is not. The values can be analyzed only with expert view and deep data knowledge. Misinterpretation of the data can mislead and even make the projects not

useful. The sustainability and the growth of the governmental service delivery would be dependent to a large extent on the success of involvement of local experts. Data scientist must be created and trained continuously.

Open data: Entering a new era of sharing data Open data platform is become a key source of governmental information system world wise. The broad sense of allowing to sharing data is the concept of open data. Open data is allowed for access with the openness of being used for further analytics by all interested. Information is kept in open platform so as to benefit from sharing and collaborating for further analysis. Researchers well use this information for developing innovative applications and usage. Third party business houses can well utilize the open data information for enhancing their business. This open data is sometimes discussed about as the twenty-first-century asset for also the small and medium-sized enterprises. The willingness to share data from various governmental sectors enriches the collaborative sharing of data.

4. Discussion for Big Data for Nepal E-Government

Although most of the countries have seen the benefit of big data, Nepal government has not explored the government efficiency improvement with the adoption of this technology. The technology is still not well known among the government bureaucracy. The attitude of the government administration must be forced to be changed by first training the bureaucrats and their concerned IT departments. Nepal's Strategic Development Goal (SDG) documentation must outline the big data compulsion projects. Training must be added in local administration and technical institutes. Consultant hiring from reputed global organization can play a key role in turning the traditional working scenario. National fiscal year budget reading must include the major projects. Trust of the users must be gained by announcing and prioritizing the projects. Data utilization program and open data sharing policies must be studied and put into the SDG outlines. The Open Governmental Data (OGD) platforms must be immediately opened with '<https://data.gov.np>' and encouraged for reuse. Integration and collaboration must be clearly guided at least within the inter department administrations. Reuse of data is for non repeated and reliable data collection. Hence, monitoring of the regular updates must be taken care by the government departments. Web user interfaces must be user friendly and must show activities to gain more trust among the users. Users must be allowed to update their profile especially their personal information like phone numbers, addresses, job description and education. Telecommunication sector also sits on top of a gold mine of data. Nepal Telecommunication Authority (NTA) must make guidelines for sharing of telecommunication data for reuse not only for business houses but for government crowd sourcing programs.

5. Conclusion

Nepal is transforming governmental systems into digital systems. The government of Nepal has encouraged the growth of ICT index by incorporating some ICT infrastructure projects since the last fiscal year. Broadband backbone projects and reach of rural areas are being prioritized along with the hype of mobile penetration with speed focused data services. Nepal has to benefit and seek the data values. Big data technologies can be adopted to solve many of the issues faced in the e-governmental sectors. With the adoption of big data it is believed that trust that is lacking in the political Nepal scenario will be recovered and the total economy will be boosted by the benefits of big data in e-government.

References

- [1] Afzalan N, Evans-Cowley JS and Barijough MM (2017), From Big to Little Data for Natural Disaster Recovery: How Online and On-the-ground Activities are Connected ? Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2526177.
- [2] Cisco (2017), Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021.
- [3] Dayal M and Singh N (2016), An Anatomization of Aadhaar Card data set – A big data challenge, *Procedia Computer Science*, 85 (2016) 733 – 739.
- [4] Ericsson (2017), Ericsson Mobility Report.
- [5] Graymatter (2017), Big Data architecture delivered successfully, Available at https://pentaho.graymatter.co.in/wp-content/themes/graymatter/pdf/UIDAI_A%20case%20study.pdf
- [6] ITU (2016), Measuring the Information Society Report.
- [7] Joseph K and Cand KS (2014), Predicting Impact of Natural Calamities in Era of Big Data and Data Science, BYU ScholarsArchive, *9th International Congress on Environmental Modeling and Software*. 15.
- [8] Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C and Byers AH (2011), McKinsey & Company, Big data: The next frontier for innovation, competition and productivity.
- [9] Rhoda CJ and Norman AJ (2013), Big Data and Transformational Government, *Published by the IEEE Computer Society*.
- [10] Simmhan Y, Shukla A and Verma A (2015), Benchmarking Fast-Data Platforms for the Aadhaar Biometric Database, Available at <https://www.researchgate.net/publication/282906171>
- [11] Tomar L, Guicheney W, Guicheney H and Zimani T (2016), Big Data on the public Sector, Inter_Americian Developmnet Bank.
- [12] Velev D and Zlateva PV (Dec 2016), An analysis of the relation between natural disasters and Big Data, Available at <https://www.researchgate.net/publication/312147428>