e-Governance using Data Warehousing and Data Mining

2 authors:

Rakesh Kr Arora
Krishna Engineering College
22 PUBLICATIONS **70** CITATIONS

SEE PROFILE

Manoj Kumar Gupta
Trinity Institute of Innovations in Professional Studies
36 PUBLICATIONS **114** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Selection of Initial Cluster Centroids for K-means Algorithm View project

Efficient Indexing Technique to Improve the Performance of 'LIKE' Operator based Comparisons View project

e-Governance using Data Warehousing and Data Mining

Rakesh K. Arora
Krishna Engineering College
Ghaziabad, UP, India

Manoj K. Gupta
Rukmini Devi Institute of Advanced Studies
Rohini, Delhi, India

ABSTRACT
Information and Communication Technology plays dominant role in all business undertaken by the government. If all interactions with government can be done through single counter 24x7, without waiting in lines at the leisure of house than it will be highly convenient to all citizens. This will not facilitate the government to avoid corruption and reach people directly, it will additionally help government agencies in rising operational efficiencies, scale back project costs, and become citizen-friendly across a range of domains, that includes transport, municipal records, education, healthcare, ports and shipping, disaster management, crime and criminal tracking system, public distribution system and agriculture sector etc.

Keywords
e-Governance, Data Warehousing, Data Mining, G2G, G2B, G2C.

1. INTRODUCTION
The basic requirements of good governance are derived from the fact that the laws and methods are well defined, transparent and easily understandable by people. To provide such good governance in a developing country like India is a challenge itself as most of the people are not educated or are not economically strong. The challenge becomes larger in developing countries as the democratic methods are used in forming the governments. In number of cases the rules and procedures defined in the constitution themselves become obstacle in the path of governance due to absence of transparency and procedural clarities.

The solution to the above problems lies in developing a mechanism that is interactive, fast and provides a clear repository of guidelines that can be used in effective decision making for both the government and the people. *e-Governance* is the mode that has large number of advantages in implementing easy, transparent, fair and interactive solutions within minimum time frame.

2. E-GOVERNANCE
e-Governance involves collection of technology based processes that involves greater interaction between government and citizens and hereby have highly improved delivery of public services [14].

e-Governance is based on the effective utilization of information and communication technologies (ICT) with major objectives of making public representatives more transparent, accountable and effective by providing improved information and service delivery and enhanced participation of people in day to day activities [18].

'E' in e-Government stands for much more than electronic and digital world. 'E' indicates:

- Efficient – do it the right way with the goal to achieve maximum output with minimum effort and/or cost.
- Effective – do the right thing
- Empowerment – active role in governance process

- Enterprise – initiative and innovation
- Enhanced – enhanced user interface by providing 24x7x365 access to government based services
- Environment friendly – it is achieved through paperless governance.

The advancements in ICT over the years along with Internet provides effective medium to establish the communication of people with the government hereby playing major role in achieving good governance goals. The information technology is playing major role in assisting the government to provide effective governance in terms of time, cost and accessibility.

3. DATA WAREHOUSING AND DATA MINING
Data Warehouse has been defined by Inmon as *"A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process"* [11]. Data from large number of homogeneous and/or heterogeneous sources are being accumulated to form data warehouse. It provides convenient and effective platform with help of online analytical processing (OLAP) to run queries over consolidated data which is extracted from multiples data sources. A centralized repository is maintained to improve user access where large amount of data is archived for analysis purpose.

Data Mining is analysis tool used to extract knowledge from vast amount of data for effective decision making. Mathematical and statistical concepts are used to uncover patterns, trends and relationships among the huge repository of data stored in a data warehouse [3].

4. NEED OF DATA WAREHOUSING AND MINING IN E-GOVERNANCE
There are some technical issues in the implementation of e-Governance which need to be taken into consideration. Some technical issues are [24]:

- Technical infrastructure for e-governance
- Collection, handling and managing huge volume of data
- Analysis of the data for effective and correct decision making
- Online support to all departments of Government organizations
- Extraction of unknown relevant and interesting patterns (i.e. knowledge) from the huge volume of data collected
- Presentation of meaningful patterns for timely decision making process

Large amount of data is being accumulated by the governments over the years. To use such data for effective decision-making, a data warehouse need to be constructed over this enormous historical data. Number of queries that require complex analysis of data can be effectively handled by decision-makers. It also helps government in making decisions that have huge impact on citizens. The decision makers are also provided with strategic intelligence to have

better view of overall situation. This significantly assists the government in taking accurate decisions within minimum time frame without depending on their IT staff.

Data mining approach extracts new and hidden interesting patterns (i.e. knowledge) from this large volume of data sets. The e-governance administrators can use this discovered knowledge to improve the quality of service. The decision involving activity in e-governance is mainly focused on the available funds, experiences from past and ground report.

The government institutions are now analyzing large amount of current and historical data to identify new and useful patterns from the large dataset. The area of focus includes:

1) Data warehousing,
2) On-line Analytical Processing (OLAP), and
3) Data Mining

5. E-GOVERNANCE USING DATA WAREHOUSING AND DATA MINING

Data Mining is the tool to discover previously unknown useful patterns from large heterogeneous databases. As historical data need to be accumulated from distinct sources to have better analysis and with prices of storage devices becoming drastically cheaper, the concept of data warehousing came into existence. If there is no centralized repository of accurate data, application of data mining tools is almost impossible [17][18].

There is wide disparity in allocation of resources in various government departments. The resources may be allocated additionally in one department while there may be acute shortage in other department. The reason behind this is non-availability of any facility to transfer information from one department to other. It is also possible that if various government departments are computerized, the information available in one department might not be beneficial to other departments as it may be possible that the information available is in dissimilar formats in heterogeneous database systems on diverse platforms. There are two approaches in designing Data Warehouse – Top down and bottom up approach. Information that starts from top is divided to generate information for lower levels (Top down approach), while information that begins from grass root level combined to generate information for higher levels (Bottom up approach). This technique provides an ideal domain of 'e-Governance' framework using Data Warehouse and Data Mining applications [3][25].

5.1 Mining E-Governance Data Warehouse

Data warehouse is used for collecting, storing and analyzing the data to assist the decision making process. Data mining can be applied to any kind of information repository like data warehouses, different types of database systems, World Wide Web, flat files etc. [16][17].

Therefore, data warehousing and data mining are best suited for number of applications based on e-Governance in G2B (Government to Business), G2C (Government to Citizen) and G2G (Government to Government) environment. In order to have effective implementation there should be solid Data Warehouse on data collected from heterogeneous reliable sources [5][7][15]. The subcategories of e-government are described in *Table 1*.

The various steps involved for implementing e-governance includes [2]:

Table 1: Subcategories of e-government (adapted from [23])

Parties of communication	Dominant Characteristics	Definition	Example
Government-to-Government (G2G)	Communication, coordination, standardization of information and services	E-administration	Establishing and using a common data warehouse
Government-to-Citizen (G2C)	Communication, transparency, accountability, effectiveness, efficiency, standardization of information and services, productivity	E-government	Government organization Web Sites, e-mail communication between citizens and government officials
Government-to-Business (G2B)	Communication, collaboration, commerce	E-government, E-Commerce, E-collaboration	Posting government bids on the Web, e-procurement, e-partnerships
Government-to-Civil Society Organization (G2SC)	Communication, coordination, transparency, accountability	E-governance	Electronic communication and coordination efforts after a disaster
Citizen-to-Citizen (C2C)	Communication, coordination, transparency, accountability, grassroots organization	E-governance	Electronic discussion groups on civic issues

a) **Phase I**: The e-governance is made available online i.e providing relevant information to the people (G2C & G2B). Earlier government websites were quite similar to brochure or leaflet but there is paradigm shifts as more and more information is made available on web. The major advantage is that government information is publicly accessible; processes are described and become more transparent, which improves democracy and service.

b) **Phase II**: It involves the communication between policy makers and the public (G2C & G2B). Public can get their queries solved via e-mail, use search engines, and download forms and documents. The applications can be processed online at very fast rate. Internally LANs, intranets and e-mail are used to communicate and exchange data by various government departments (G2G).

c) **Phase III**: The complexity of transactions increases in **third phase**. Complete transactions can be performed at the leisure of house. Extending/renewal of licenses, application for visa and passports filing property tax, filing income tax, and online voting are common examples. This phase handles complex queries by use of security and personalization issues. E.g. digital signatures will be mandatory to have legal transfer of services. The government has also made e-procurement compulsory for all procurements above Rs 5,000/-.

d) **Phase IV**: The major goal is to provide single counter by integrating all information systems. The employees in various government departments have to work in the coordinated manner to have cost savings, efficiency and most importantly highest customer satisfaction.

5.2 Need for Data Warehousing and Data Mining (DWDM) in e-Governance

The use of DWDM technologies will assist decision makers to reach important conclusion that can play important role in any 'e-Governance' initiative [1][9][16]. The need of DWDM in e-governance includes:

- Provision of integrated data from diverse platforms for better implementation of strategies at state or national level.

- To minimize piracy of data as storage requirement is reduced.
- To increase operational effectiveness as various employees work in coordinated manner.
- To increase transparency at highest level as relevant information will be available on web.
- To have better understanding of requirements of citizens.
- To have faster access of data for effective decision making.

5.3 Integration of Data Warehousing and Data Mining with e-Governance

The advantages of integrating DWDM with 'e-Governance' are:

- There is no requirement to deal with heterogeneous databases.
- Officers will be able to derive the output at multiple levels of granularity.
- There is no requirement to use complex tools to derive information from vast amount of data.
- In depth analysis of data is possible to have solution to complex queries.
- There is mass reduction of dependence on IT staff.
- Strong tool towards corruption free India.

6. ORIGINS OF E-GOVERNANCE IN INDIA

The origin of e-governance in India in mid seventies was confined mainly in the area of defense and to handle queries that involves that large amount of data related to census, elections and tax administration. The set up of National Informatics Centre (NIC) in 1976 by Government of India was major boost towards e-governance. The major push towards e-governance was initiated in 1987 with the launch of NICNET-the national satellite based computer network. The launch of District information system by NIC to computerize all district offices in India was another major step towards e-governance. During early nineties, there was significant increase in the use of IT in applications where government policies start reaching to non urban areas by having good inputs from number of NGOs and private sector.

The area of e-governance has become very wide now. The government is now implementing e-governance in every field. e-governance has now spread its wings from urban to rural areas. There is hardly any field left in which e-governance has not entered. The e-governance is playing major role in routine transactions like payment of bills and taxes, public grievance system, municipal services like maintaining records of land and property, issue of birth/death/marriage certificate, registration and attorneys of properties, traffic management, health services, disaster management, education sector, crime and criminal tracking system, public distribution systems and most importantly providing up to date information in agriculture sector.

Number of states have set up their own portals but most of these portals are incapable of providing complete solution to people by just click of the mouse. In most of cases ministries and individual departments have separate websites to provide the necessary information. This should not be the case as the user has to visit multiple websites to get relevant information. Ideally the official website should act as single window to provide necessary information and services [4][8].

The various projects initiated by various state governments towards e-governance are listed in *Table 2*:

Table 2   e-Governance Initiatives in India (Adapted from [10])

State / UT	Initiatives Taken
Tamil Nadu	Tender notices and display, Application forms related to public utility, Rasi Maiyams–Kanchipuram
Rajasthan	RajNIDHI, Lokmitra, RajSWIFT, Jan Mitra
Maharashtra	Online Complaint Management System—Mumbai, SETU
Madhya Pradesh	Computerization MP State Agricultural Marketing Board (Mandi Board), Smart Card in Transport Department, Gram Sampark, Gyandoot
Kerala	Fast, Reliable, Instant, Efficient Network for the Disbursement of Services (FRIENDS), RDNet, e-Srinkhala
Karnataka	Kaveri, Khajane, Bhoomi
Himachal Pradesh	Lok Mitra
Haryana	Nai Disha
Gujarat	G R book online, census online, Form book online, tender notice, request for Government documents online, Mahiti Shakti
Goa	Dharani Project
Delhi	Management Information System for Education, Electronic Clearance System, Computerisation of website of RCS office, Automatic Vehicle Tracking System,
Chhattisgarh	e-linking project, Treasury office, Chhattisgarh Infotech Promotion Society
Bihar	Sales Tax Administration Management Information
Andhra Pradesh	Online Transaction processing, Saukaryam, AP online—One-stop-shop on the Internet, e-Cops, FAST, MPHS, VOICE, CARD, e-Seva

Department of Electronics and Information Technology (DEITY) has initiated National e-Governance Plan (NeGP) in 2006 with the objective of improved delivery of government services to people. NeGP provides platform for various e-governance initiatives. The various central government initiatives includes e-office, unique identification project(UID), Immigration, e-procurement, e-Biz ,e-Courts ,Direct cash transfer, Aadhar Enabled Payment system(AEPS), Digital India program, MyGov citizen portal and digital cloud for every Indian.

7. CHALLENGES FOR IMPLEMENTATION OF E-GOVERNANCE IN INDIA

E-governance in India is at infant stage. However, there are limited successful and completed e-governance projects like e-Seva, CARD, etc.

Lack of insight can be attributes as major factor for failure of e-governance projects in India. Reservation and inflation can be topic of national debates but e-governance was never the issue in Indian politics. Lately the government of India has risen to occasion and started pushing the projects related to e-governance.

a) Illiteracy and limited awareness regarding positives of e-governance [6][12].

b) Casual attitude of government officers towards public. The officers cannot be punished in absence of proper guidelines.

c) Lack of electricity and internet facilities especially in rural areas to reap benefits of e-governance.

d) Huge delay in implementing e-governance projects due to technical reasons or proper support. Lack of understanding and interests among senior personnel also affect projects.

e) Diversity of country is biggest challenge. To have e-governance projects in local language is huge task to implement.

f) Absence of qualified pool of resources to manage the system is challenging task. Refusal of IT professionals to work in rural areas also affects the projects.

g) The role of public in policy making is negligible. If the opinion of people at the grassroots level is taken into account than majority of problems can be solved.

8. CONCLUSION

An inter relationship between e-Governance and Data Warehousing & Mining is presented in this paper. The introduction of DWDM in the area of 'e-Governance' strengthens the system. Initially few sectors need to be identified to incorporate the DWDM system. Once the desired response is achieved, the same can be implemented in other fields also. Once the full system is implemented at the national level, establishment of knowledge bank can play dominant role for whole 'e-Governance' environment.

By just click of the mouse, 'e-Governance' has the potential to solve large set of problems that exist within the existing system [13]. Once it becomes a reality it will give a major boost to the rights of an ordinary citizen. The Hon'ble Prime Minister of India Shri Narendra Modi has taken major steps towards e-governance and had sanctioned Rupees One Lakh Crore in last budget towards Digital India project. The call for minimum government, maximum governance has become louder now days.

9. REFERENCES

[1] Agarwal, S., Singh, N. and Pandey, G.N. (2010), "Implementation of Data Mining and Data Warehousing In E-Governance", International Journal of Computer Applications, Volume 9– No.4, November 2010

[2] Al-Hashmi, A. and Darem, A.B., 'Understanding phases of E-Government projects', 2008 (online available at https://www.researchgate.net)

[3] Andriaans, P. and Zantinge, D., 'Data Mining', Pearson, 1996.

[4] Arora, R.K., and Gupta, M.K., 'Data Mining: Scope Out Valuable Resources From Mountains Of Information', IITM Business Review Journal, 2010.

[5] Bansal, K.L. and Sood, Satish, 'Role of Data Mining and Data Warehousing in e-governance', Intl. J. of Artificial Intelligence and Knowledge Discovery, Vol. 1 No .1, 2011

[6] Bhanti, P., Kaushal, U., and Pandey, A. (2011), "E-Governance in Higher Education: Concept and Role of Data Warehousing Techniques", International Journal of Computer Applications, Volume 18– No.1, March 2011

[7] Bhatnagar, S., E-Government: From Vision to Implementation, Sage Publications, India, 2004

[8] Chakrabarty, T., 'Towards an ideal e-governance scenario in India', 2008 (available at http://www.tcs.com/SiteCollectionDocuments/White%20 Papers/tcs_government_idealegovernanceindia.pdf)

[9] Gupta, G.K. 'Introduction to Data Mining with Case Studies', PHI, 2011.

[10] DataQuest. http://dqindia.com

[11] Inmon, W.H., 'Building the Data Warehouse', Fourth Edition, WILEY dreamtech India Pvt. Ltd., 2005.

[12] Karthika, D. and RangaRaj, R. (2013), "Survey of E-Governance based Higher Education System in India with Data Mining", International Journal of Science and Research (IJSR), Volume 2 Issue 12, December 2013

[13] Misra, D C, (2009), An E-governance Vision for India by 2020

[14] Nair, P., 'E-Governance; A Step Towards Digital Democracy' 2008 (online available at https://www.researchgate.net)

[15] Sharma, M.K., "Application of Data Mining in e-governance data Warehouse", (Available at http://www.egovonline.net/articles-list/47-features/4168).

[16] Martin, G., "Data Mining techniques and the decision Making Process in the Bulgarian Public administration".

[17] R.J. Brachman, Khabaza, T. Kloegan, W., Piatetsky-Shapiro, G. and Simoudis, E. (1996) Mining Business databases, Communication of the ACM 39,11,42-48.

[18] Syvajarvi Anitt et al., "Information management as function of Data Mining and ICT in City Government", EGPA 2009, SGI.

[19] Stuart W. Shulman, "eRulemaking: Issues in Current Research and Practice," International Journal of Public Administration Vol. 28 (2005), 621-641.

[20] The E-Government Magazine for Asia and The Middle East, "e-Gov", published in technical collaboration with Centre of Science, Development and Media Studies.

[21] United Nations E-Government Survey 2010- Leveraging e-government at a time of financial and economic crisis

[22] Vijay K.D., 'Data Mining in E-Governance Models and Applications', International journal of Business Policy and Strategic Management, 2014

[23] Yildiz, M. (2003). A general evaluation of the theory and practice of e-government (In Turkish). In M. Acar & H. Ozgur (Eds.), Cagdas Kamu Yonetimi-1 (pp. 305−328). Istanbul: Nobel Publications.

[24] Yildiz, M. (2007), "E-government research: Reviewing the literature, limitations, and ways forward", Elsevier, Government Information Quarterly 24 (2007) 646–665

[25] Zafaruddin, G.M., 'Data Warehousing and Mining: Futuristic Approach to e-Governance', IEEE International Advance Computing Conference (IACC 2009) Patiala, India, 6–7 March 2009.