

E-Governance in Higher Education: Concept and Role of Data Warehousing Techniques

Prateek Bhanti

Asst. Professor,
FASC, MITS Deemed University,
Lakshmangarh-332311, Sikar,
Rajasthan, INDIA

Urmani Kaushal

Asst. Professor,
FASC, MITS Deemed University,
Lakshmangarh-332311, Sikar,
Rajasthan, INDIA

Archana Pandey

Asst. HR Manager,
AKG Entertainment Pvt. Ltd.
Mumbai,
Maharashtra INDIA

ABSTRACT

In this paper we propose E-governance implementation for higher education system with the use of data warehousing and data mining techniques. For this we intend to frame a logical architecture design for higher education system. On the basis of this architecture we will develop the logical design of database. We aim at demonstrating the advantages and techniques of using data warehouse architecture in e-governance applications, with special reference to higher education system in India.

Keywords

E-governance, Data warehousing, Data Mining.

1. INTRODUCTION

E-government [1] refers to the use of information and communication technology to carry out government operations such as delivering government information and services. E-government is generally recognized as a means of making government more efficient while allowing it to be more responsive to customer needs [2]. E-governance can also be defined as the application of electronic means in the:-

- Monitoring the performance of various scheme started by government from time to time
- Interaction between government - citizens and government-business
- Simplification processes of the government
- Internal government operations to simplify and improve democratic, government & business aspects of governance

Application developers [3] in E-governance have to work at low level of abstraction. This means taking care of low level issues such as inter process messaging tools, integration and data modeling while defining the application logic. In solution to date, each e-governance solution has customized existing products to address an individual government agency requirement. E-governance process brings transparency in the system, so e-governance initiatives in the field of higher education can reduce the corruption up to a large extent. This type of initiative can give a better interface and opportunity to students, parents, teachers and administrative management.

1.1 Data Warehousing

Following Bill Inmon [4] a data warehouse is "a subject oriented, integrated, non volatile and time-variant collection of data in support of management's decisions". The main objective is to bring together the various data which are distributed in transactional heterogeneous databases into a single collection [5][6][7]. The centralization of voluminous amount of data allows the discovering of trends and provides hidden information that eases managers' decisions [8].

1.2 Data Mining

Data Warehousing and knowledge Management can contribute a lot to decision support system in e-Governance procedures. Data Mining is also a core step, which results in the discovery of hidden but useful information from massive database and e-governance operation which involves pragmatic use of database and the technique of Data mining and Data warehousing for taking various decisions for the benefits of the society.

2. RELATIONSHIP BETWEEN E-GOVERNANCE, DATA MINING AND DATA WAREHOUSING

The main problem for the e-governance procedures is that their operational system needs accurate repository of data. These operational systems also tend to be numerous, with overlapping and sometimes contrary definition. Until there is a repository of accurate data across the department, application of mining tool to analyze and aid in the strategic government decision is not possible. To meet these needs, there is a new breed of system that has evolved and matured in the recent years.

Following section explore the relation between e-governance, Data warehousing and Data Mining.

2.1 E-governance & Data warehousing

Data warehouse is a subject-oriented, integrated, time-variant, non-volatile collection of data, cutting across the enterprise. Until there is a repository of accurate data across the enterprise value chain, application of mining tools to analyze and aid in strategic government decisions is impossible.

At times, government departments might come across shortages of resources in one department and excess of resources in the other. This could be due to non-availability of proper data and facilities to disseminate information. Even if government departments are computerized and networked more for the purpose of Internet usage and mail transfer, the information

available in one department, which possess the data might not be utilized by other departments.

This is because the information is stored in different formats, in different platforms and in heterogeneous data base systems. A look at the information requirements at each level and the information flow across levels shows a pattern. While information that flows from top (fund sanction, allocation and disbursement details) is split up to generate information for lower levels, information that flows from grass root level (such as expenditure details, benefits details, beneficiary details etc.) is consolidated to generate information for higher levels. This pattern makes the entire vertical domain of e-Governance framework, an ideal domain for development of data warehousing and use of data mining applications [9].

2.2 E-governance & Data mining

Data mining is a broad category of applications and technologies for gathering, storing, analyzing and providing access to data to help the decision makers in making decisions. Typically, any Data Warehouse and Mining (DWM) application includes large data warehouse, decision support systems, query and reporting, On Line Analytical Process (OLAP), statistical analysis, forecasting and mining (a technology to extract unknown and hidden patterns and knowledge from within the data). DWM, therefore is well suited for e-Governance applications in the G2G (Government to Government) and G2C (Government to Citizen) environment. For effective implementation of a DWM solution, the de facto condition is a solid and reliable data warehouse on available e-Governance data from different sources [9][10].

Data Mining has been successfully deployed in numerous application areas (Brachman et. al 1996) for e.g. Amazon uses it for suggesting additional books, based on customer's current choices. Within the government the focus of data mining is predominantly on the management of interaction between the government and citizen or business. In these application data mining can extract knowledge from transactional data, can lower the burden of dealing with large volume of transactions and can improve decision making abilities [11][12].

3. NEED OF DATA WAREHOUSING IN HIGHER EDUCATION

Everyone associated with higher education institution like institutional members, parents, students, employees, funding bodies and the government in other words; the stakeholders would be interested in one question about the institutions: "What is the quality of education offered in higher education institutes?" The stakeholders raise these questions with one or more interest: students for choice of institution; Parent for worth of personal investment in the education of their wards; Government accountability and policy making; funding agencies for funding decisions. To find an answer to the questions on institutional quality, these stakeholders depend on various sources of data [13].

One area of interest for regulating authorities is to control the quality and standards of higher education system. Recently public administration has become aware of the benefits of data warehousing. The above new technologies enable interactive data analysis and adhoc reporting (INMON 1996) [14].

The data warehousing can primarily provide categorical structure for static data and variable data. Any addition of variable data to certain static data would enable automatic conversion of the static data to influencing data. Data structure would further enable automatic modification of relevant database.

The database structure should have multidimensional processing pattern tailored not only to the need of the departments but also the end users i.e. the citizens.

Each university in India acts widely autonomously and competes with other universities for students. In this context the difference between time of request for funds and disbursement of funds leaves much to the forecasting and guess work. With the help of e-governance application, the process can be made automated or time bound. All details needed for the approval of funds can be fetched by a data warehouse. A data warehouse can help to provide a fair allocation of available funds.

Similarly, E- Governance system, by using centralized database of students, can help governing bodies to provide opportunity to bright students. The governing bodies can provide a common platform for best performing student and industry seeking employees to interact for jobs project research work etc. The governing bodies/ industry can very easily get the details of bright students all over in India in various fields with the help of data warehouse[14][15].

4. PROPOSED SOLUTION

For above type of applications voluminous data will be thrown that can only be handled by using data warehousing techniques. In the following section we will propose the possible implementation of the solution. The proposal is based on literature review, and on the findings of completed open source web warehouse project (Thomsen and Pederson, USA 2006).

4.1 Logical architecture of data warehouse for higher education system

A data warehouse brings together data from multiple operational (internal) and external data sources into common physically separated repository (INMON, 1996).

The main purpose of the Data Warehouse is to serve as a central reporting and data distribution environment for regulating authorities and students. The Data Warehouse acts as a hub, to facilitate the exchange of information between systems and therefore serves as the enterprise information infrastructure [16] [17].

The accessibility, transparency, efficiency and impact of e-governance service will be automatically evaluated, which will produce a large amount of data. We will propose a architectural framework of data warehouse which is expected to meet the demand of the data storage needed by e-governance application

4.1.1 The Data Layer illustrates the idea that university data should be logically visible and consolidated by data domain.

4.1.2 The Data & Services Integration Layer speaks to the fact that there should be a consistent way of interacting with university data, and a coherent strategy for sharing data across systems.

4.1.3 The Services Layer represents the separation of re-usable services from application logic. The service layer will thus consist of services with clearly defined contracts that can be used by any application. The services have initially been classified in to three major groupings: Core Services for technical and basic services, Administrative for services that provide access to administrative data or processes and Academic Services for services specific to the education and research domains.

4.1.4 The Service Integration Layer is responsible for exposing the services in the architecture in a consistent manner while enabling services to be implemented in a variety of technologies. Ideally it should also define the standard contract for a service type, therefore allowing substitution of service implementation without affecting clients of the service.

4.1.5 The Applications Layer shows groupings of applications that are built for specific purposes. By leveraging services available in the architecture, applications should generally be quicker to develop and easier to maintain.

4.1.6 The User Interfaces Layer conveys the idea that users should have a single point of access for related functions that they use. This might be implemented as one or more portals.

4.1.7 The Security Services Layer, while conceptually similar to other types of services has been shown separately because it has significant impacts at all levels within the architecture. It will be necessary to apply access and control security to data, to services, to applications and finally to user interfaces.

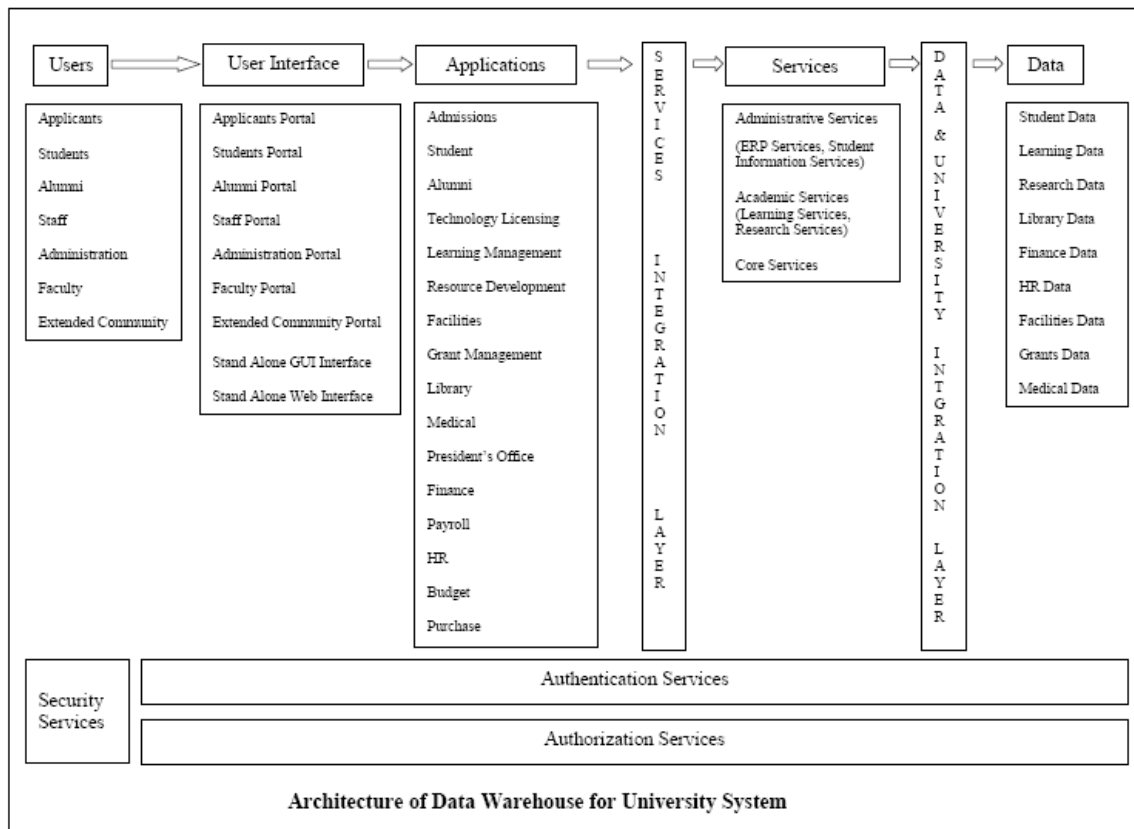


Fig. 4.1

5. LOGICAL DESIGN OF DATABASE FOR UNIVERSITY SYSTEM

The star schema makes multi-dimensional database (MDDB) functionality possible using a traditional relational database. Because relational databases are the most common data management system in organizations today, implementing

multi-dimensional views of data using a relational database are very appealing. Even if a specific MDDB solution is used, its sources likely are relational databases.

Another reason for using star schema is its ease of understanding. Fact tables in star schema are mostly in third normal form (3NF), but dimensional tables are in de-normalized

second normal form (2NF). If you want to normalize dimensional tables, they look like snowflakes (see snowflake schema) and the same problems of 3NF databases arise - you need complex queries and business users cannot easily understand the meaning of data. Although query performance may be improved by advanced DBMS technology and hardware, highly normalized tables make reporting difficult and applications complex.

The star schema architecture for university system can be illustrated in the following manner:

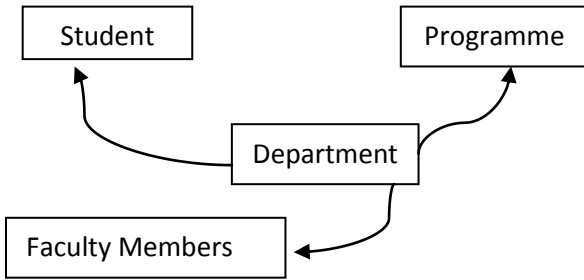


Fig. 5.1

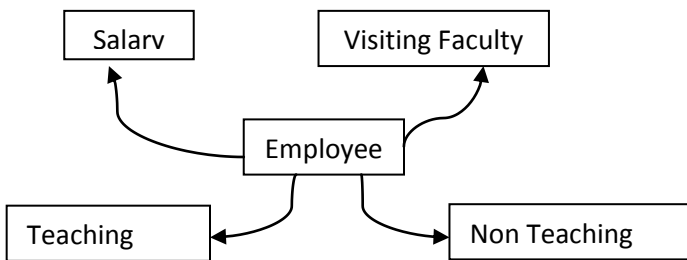


Fig. 5.2

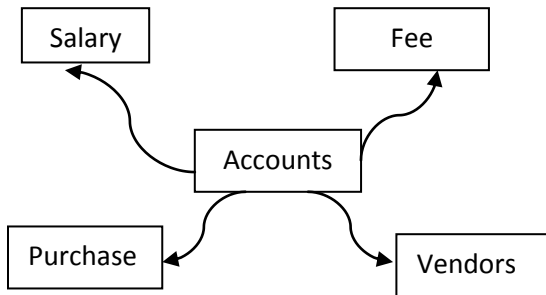


Fig. 5.3

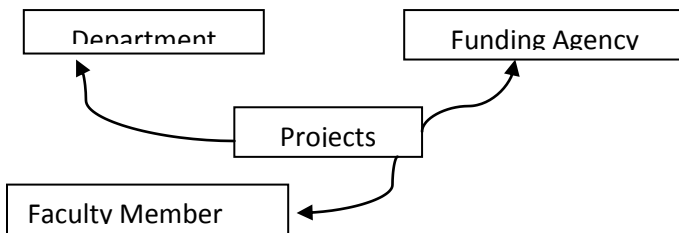


Fig. 5.4

6. CONCLUSION

Although e-governance affairs enhanced the government part work efficiency and transparency greatly, but the services were facing with the problem of accumulated massive data and thus it was far from decision and forecasting. Implementing e-

governance application in higher education will produce voluminous data storage. Records of million of people can be stored and computerized and Data Mining techniques may be helping in :

- Understanding the citizen needs in a better way
- Gain more effectiveness in operations
- Developing implementations plans on state and rational level.
- Developing extensive and effective database for the e-society
- Getting key conclusions from large amount of data that can be a critical component of any e-governance initiative
- Performance monitoring of various schemes started by government time to time. Performance can be effectively monitored by a grid of datacenters and networks which connects them up to the kiosks at village level.
- Obtaining easily decipherable and comprehensive information without the need to use sophisticated tools.
- Formulating more effective strategies and policies for citizen facilitation.

The above proposal, through an architectural framework of data warehouse demonstrates the nature and potential of e-government public service application in higher education system in India. Consequently providing better decision making and support to the government for the actual work.

7. ACKNOWLEDGMENTS

We gratefully acknowledge support from Dean and faculty members, FASC, MITS, Lakshmanarh, Sikar & AKG Entertainment Pvt. Ltd, Mumbai for the same.

8. REFERENCES

- [1] Alison Radl, Yu-Che Chen, "Computer Security in Electronic Government: A State-Local Education Information System", International Journal of Electronic Government Research, Jan-March 2005, 1(1), 79-99.
- [2] Vaibhav Panwar, "ROLE OF DATA WAREHOUSING & DATA MINING IN E-GOVERNANCE".
- [3] V.M Rao: "E-governance" : ABD Publishers, Jaipur, P-6, 2007.
- [4] W.H.. INMON, "Building the Data Warehouse", 2nd edition, John Wiley & Sons, New York(1996).
- [5] J. Widom: Research Problems in Data Warehousing. Proceedings of the 1995 International Conference on Information and Knowledge.
- [6] R. Kimball: The Data Warehouse toolkit. Wiley Computer Publishing. 1996.
- [7] S. Chaudhuri, U. Dayal: An overview of Data Warehousing and OLAP technology. ACM SIGMOD Record 26 (1) 1997.

- [8] WWW Source, Flory Andry et al, “A Design and implementation of a data warehouse for research administration universities”.(Available :[http://subs.emis.de / LNI / Proceedings / Proceedings13 / 35_aDesigandImplem.pdf](http://subs.emis.de/LNI/Proceedings/Proceedings13/35_aDesigandImplem.pdf)).
- [9] WWW Source, M.K. Sharma, “Application of Data Mining in e-governance data Warehouse”, (Available at <http://www.egovonline.net/articles-list/47-features/4168>).
- [10] Gramatikov Martin, “Data Mining techniques and the decision Making Process in the Bulgarian Public administration”.
- [11] 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.
- [12] Syvajarvi Anitt et al.,”Information management as function of Data Mining and ICT in City Government”, EGPA 2009, SGI.
- [13] Antony Stella, “Measuring Institutional Quality in higher education: Ranking, Rating and Report by Academics and Media”, University News, Vol 43, No 22, 2005, P-1.
- [14] W.H.. INMON, “Building the Data Warehouse. 2nd edition, John Wiley & Sons, New York(1996).
- [15] Syvajarvi Anitt et al.,”Information management as function of Data Mining and ICT in City Government”, EGPA 2009, SGI.
- [16] WWW Source, “Future State | Logical Architecture Vision” Available at [http:// web.mit.edu / itag/ eag-0.1/ FutureStateLogical.pdf](http://web.mit.edu/itag/eag-0.1/FutureStateLogical.pdf), Version 0.1 – August – September 2004.
- [17] Wagner Christian et. al., “Enhancing E-government in Developing Countries: Managing Knowledge though virtual community”, EJISDC (2003),14,4,1-20.