

REVIEW ON ROLE OF DATA WAREHOUSE IN BUSINESS INTELLIGENCE

Kanika Passi

M.Tech , Department of Computer Science & Engineering, SVIET, Banur

Abstract: *Business Intelligence refers to a set of methods and techniques that are used by organizations for tactical and strategic decision making. It leverages technologies that focus on counts, statistics and business objectives to improve business performance. Effective managerial decision-making is critically dependent upon the availability of integrated, high quality information organized that support in decision making processes. Data warehouses have emerged to meet this requirement.*

Keywords: Data Warehouse, big data, Business intelligence

I. INTRODUCTION

A subject-oriented, integrated, time-variant, non-updatable collection of data used in support of management decision-making processes.[1] Another way of saying the same thing is that a data warehouse provides a “single version of the truth” for decision making in the corporation.

Data warehouse acts as a middleware in Business application architecture, which are essential as direct accessing to operational and transactional data for decision support applications is infeasible.[2] A Data Warehouse (DW) is simply a consolidation of data from a variety of sources that is designed to support strategic and tactical decision making.[3] Its main purpose is to provide a coherent picture of the business at a point in time. By using various Data Warehousing toolsets, users can execute online queries and 'mine' their data. Many successful companies have been investing large sums of money in business intelligence and data warehousing tools and technologies.

This review is an attempt to determine the role of data warehouse in business. Various Key challenges are also discussed in this paper, which data warehouse are facing now.

II. NEED OF DATA WAREHOUSE

A data warehouse centralizes data that are scattered throughout disparate operational systems and makes them available for DS. This situation does not enable good decision making. A well- designed data warehouse adds value to data by improving their quality and consistency. A separate data warehouse eliminates much of the contention for resources that results when information applications are mixed with operational processing. When data redundancy exists in multiple databases, data quality often deteriorates. Poor business intelligence results in poor strategic and tactical decision making.

III. DATA WAREHOUSE PROCESS IN BUSINESS

Data warehousing is every business analyst's dream as all the information regarding the organization's activities gathered in one place because the purpose of a data warehouse system is to provide decision-makers the precise, integrated and timely information they need to make the right choices. But how do you make the dream a reality? A process has been described for implementing data warehouse system in Fig. 1

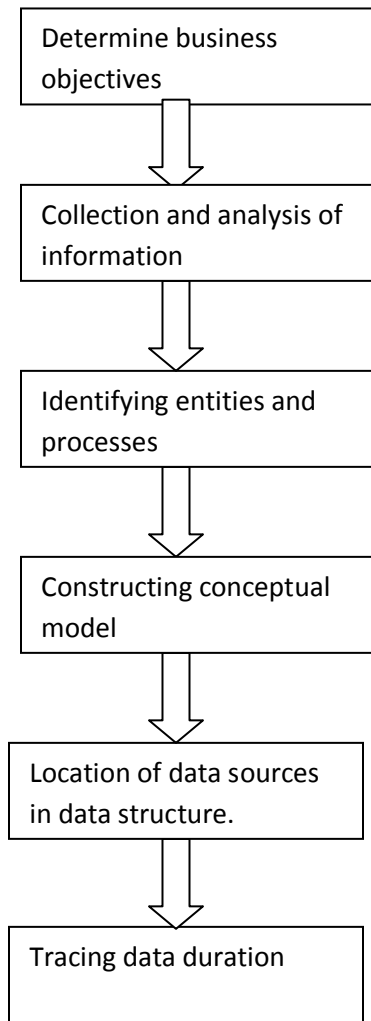


Fig. 1 Implementing Data warehouse

Steps are as follows:

- Firstly, determination of organization's objectives which includes sales, production and administration.
- Collection and analysis of the information gathered by creating summary and reports.
- Identification of various entities and processes involved and correlating with the factors affecting them.
- Construction of conceptual model by creating fact tables.
- Locating the data into a consolidated and exact data structure.
- Reconciling of the data by tracking the data.

IV. GOALS AND OBJECTIVES OF DATA WAREHOUSE.

Data warehouse helps in achieving various goals. Such as:

A. Data Integrity

The data warehouse concept seeks to integrate data across time and across different subject areas in such a way that users of the warehouse can easily obtain facts about their business.

B. Normalization of data

Normalization includes the logical analysis of the determination of the most simplest, stable data that can be stored in database and makes warehouse more understandable.

C. Information consistency

The *metadata layer* of the data warehouse enforces information consistency by allowing data within the data warehouse to be defined in *business terms* as opposed to using database jargon.

V. DATA WAREHOUSE APPLICATION

Data warehouse are being used in different industries. The Nanhai city E-government project is based on Data warehouse technology which not only consider to provide data for high level four databases, but also solve integration, share and exchange of data in various departments of Nanhai city, especially devised application database based on data warehouse technology for better utilizing accumulated data.

An Integrated National Agriculture Resources Information System (INARIS) was developed which provide system and planned information to the scientists, decision makers in the form of OLAP System.[4]

The Atlas system includes Atlas biological Data warehouse which Acts as data infrastructure bioinformatics research, which facilitate the integration of heterogeneous biological data.[5]

VI. KEY CHALLENGES IN DATA WAREHOUSE

A. Information-driven Analysis

One of the key things to do, if you want to ensure a success, is to spend the time necessary on understanding and documenting the business and its needs. The most important activity is the modeling of the business goals and concepts.

B. Data Integration

Data integration is a key challenge in the design of a data warehouse solution. A data warehouse should present an integrated and reconciled view of organizational data to the stakeholders. The integration process should be efficient enough to manage large volumes of data and is able to eliminate any redundancies and inconsistencies while data is integrated from various systems.[6]

C. Data Warehouse Testing

Testing in data warehousing is a real challenge. A typical 20% time allocation on testing is just not enough. One of the reasons why testing is tricky is due to the reason that a top level object in data warehouse (e.g. BI reports) typically has high amount of dependency.

VII. COMPARING BING DATA AND DATA WAREHOUSE

A Big Data definition includes:

- Technology capable of holding very large amounts of data. Technology that can hold the data in inexpensive storage devices.
- Technology where the data is stored in an unstructured format.

A big data solution is a technology and data warehousing is an architecture. They are two very different things[7]. A technology is just a means to store and manage large amounts of data. A data warehouse is a way of organizing data so that there is integrity.

Table - I

Business Requirement	Data Warehouse	Big Data
Unstructured Data		Partially
Quality of Data	Yes	Yes
Analysis of prior data		Yes

VII. CONCLUSION

Data warehouse must be integrated in organization's Architecture as to really know what is going on in your corporation, you need data that is reliable, believable and accessible to everyone. A data warehouse depends on making the best, most up-to-date information available to users when they need it. There are various challenges relating to data warehouse like unable to provide timely or necessarily accurate trending analysis, Multiple data sources provide challenges in regards to data collection and reuse, Consolidated data based historical trending and forecast modeling is unattainable. Challenge relating to misuse of data is critical as data is powerful.

REFERENCES

- [1] Nectaria Tryfona, Frank Busborg, and Jens G. Borch Christiansen , "A Conceptual Model for Data Warehouse Design" , Proceedings of the 2nd ACM international workshop on Data warehousing and OLAP, New York, 1999, pp 3-8.
- [2] Robert Winter, "The Current and Future Role of Data Warehousing in Corporate Application Architecture", Proceedings of the 34th Annual Hawaii International Conference on System Sciences, Switzerland, Jan 2001, pp 13-17
- [3] Dimitri Theodoratos Timos Sellis, "Data Warehouse Configuration". Proceedings of the 23rd International Conference on Very Large Data Bases, San Francisco, 1997, pp 126- 135.
- [4] Anil Rai," Data Warehouse and its Application in agriculture", Indian Agricultural Statistics Research Institute Library Avenue, New Delhi, pp 175-183

[5] Sohrab P Shah, Yong Huang, Tao Xu, Macaire MS Yuen, John Ling and BF Francis Ouellette” Atlas – a data warehouse for integrative bioinformatics”, February 2005, pp 1-16.

[6] Salvatore T. March a, Alan R. Hevner b,” Integrated decision support systems, Journal of Decision Support Systems, April 2007, pp 1031-1043.

[7] J. P. Dijcks, "Oracle: Big data for the enterprise," Oracle White Paper, June 2013, pp 2-14.