Data Warehousing and Business Analytics Implementation

S.NASIRA TABASSUM

SK.ALTHAF AHAMMED

Abstract

Data warehousing (DW) has emerged as one of the most powerful technology innovations in recent years to support organization-wide decision making and has become a key component in the information technology (IT) infrastructure Data warehousing methodologies share a common set of tasks, including business requirements analysis, data design, architectural design, implementation and deployment. This paper provides an overview of Analytics.   It explains how data from a web server's log can be harvested to generate useful and actionable business intelligence, particularly when the data is combined with existing customer and sales data in a Data Warehouse. Business Intelligence is often confused and sometimes used intermittently with data warehouse concepts. We discuss the similarities and differences between these two concepts in this paper.

Keywords: data warehouse, business intelligence

1. Introduction

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. 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. Its main purpose is to provide a coherent picture of the business at a point in time. Using various Data Warehousing toolsets, users are able to run 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. They believe that up-to-date, accurate and integrated information about their supply chain, products and customers are critical for their very survival. This website introduces some key Data Warehousing concepts and terminology. It explains Data Warehousing from a historical context and the underlying business and technology drivers that are making Data Warehouses a hot commodity.

A data warehousing (or data mart) system is the backend, or the infrastructural, component for achieving business intelligence. Business intelligence also includes the insight gained from doing data mining analysis, as well as unstructured data (thus the need for content management systems) thus providing managerial decision support for complex business questions. DW is also an enabling technology for data mining, customer-relationship management, and other business-intelligence applications. Although data warehouses have been around for quite some time, they have been plagued by high failure rates and limited spread or use. Drawing upon past research on the adoption and diffusion of innovations and on the implementation of information systems (IS), we examine the key organizational and innovation factors that influence the infusion (diffusion) of DW within organizations and also examine if more extensive infusion leads to improved organizational outcomes. Business intelligence is closely related to data warehousing. This section discusses business intelligence, as well as the relationship between business intelligence and data warehousing. As the old Chinese saying goes, "To accomplish a goal, make sure the proper tools are selected." This is especially true when the goal is to achieve business intelligence. Given the complexity of the data warehousing system and the cross-departmental implications of the project, it is easy to see why the proper selection of business intelligence software and personnel is very important. After the tools and team personnel selections are made, the data warehouse design can begin phase.

2. Business Intelligence and Data Warehouse

Business intelligence (BI) is defined as the ability for an organization to take all its capabilities and convert them into knowledge. This produces large amounts of information that can lead to the development of new opportunities. Identifying these opportunities, and implementing an effective strategy, can provide a competitive market advantage and long-term stability within the organization's industry. BI technologies provide historical, current and predictive views of business operations. Common functions of business intelligence technologies are reporting, online analytical processing   , analytics, data mining, process   mining, complex   event processing, business   performance   management , benchmarking   , text   mining   , predictive analytics and prescriptive   analytics. Business intelligence aims to support better business decision-

making. Thus a BI system can be called a decision support system(DSS). Though the term business intelligence is sometimes used as a synonym for competitive intelligence (because they both support decision making), BI uses technologies, processes, and applications to analyze mostly internal, structured data and business processes while competitive intelligence gathers, analyzes and disseminates information with a topical focus on company competitors. If understood broadly, business intelligence can include the subset of competitive intelligence. Often BI applications use data gathered from a data warehouse or a data mart. However, not all data warehouses are used for business intelligence, nor do all business intelligence applications require a data warehouse.

In order to distinguish between concepts of business intelligence and data warehouses, Forrester Research often defines business intelligence in one of two ways:

Using a broad definition: "Business Intelligence is a set of methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information used to enable more effective strategic, tactical, and operational insights and decision-making."[6] When using this definition, business intelligence also includes technologies such as data integration, data quality, data warehousing, master data management, text and content analytics, and many others that the market sometimes lumps into the information management segment. Therefore, Forrester refers to data preparation and data usage as two separate, but closely linked segments of the business intelligence architectural stack.

Forrester defines the latter, narrower business intelligence market as "referring to just the top layers of the BI architectural stack such as reporting, analytics and dashboards."[7] Thomas Davenport has argued that business intelligence should be divided into querying ,reporting, OLAP, an "alerts" tool, and business analytics. In this definition, business analytics is the subset of BI based on statistics, prediction, and optimization'[8]. Before implementing a BI solution, it is worth taking different factors into consideration before proceeding.

According to Kimball et al., these are the three critical areas that you need to assess within your organization before getting ready to do a BI project:[12]

1. The level of commitment and sponsorship of the project from senior management
2. The level of business need for creating a BI implementation
3. The amount and quality of business data available.

4. Business Intelligence Services

3. Related Work
Dell helps create a successful business intelligence solution.

Business needs and current state of information assets are analyzed to define a BI strategy that may include new or revised dashboards, redesigned data models and integration of source systems.

Business Intelligence Advisory Services
The Dell Business Intelligence Advisory Services include:
- Framework for establishing a business intelligence competency center (BICC)
- Plans for building out a successful business process management (BPM) program, including a collaborative approach to workshops and process flows
- Business solutions defined and aligned to corporate strategy

Business Intelligence Advisory Services helps to develop a comprehensive business intelligence (BI) strategy and roadmap that aligns with strategic enterprise goals and get the information needed to make fast, fact-based decisions.

Business Intelligence Design and Plan
After BI analysis, the next step is to design and plan the business intelligence data that is needed by the organization. The BI design and plan helps to create a detailed business and technology design that will scale to the business needs, improve data quality and turn information into actionable intelligence.

Business Intelligence Implementation
The designed BI is then implemented using the right tools that help to speed up deployment and ensure a successful adoption of business intelligence and data warehouse (BIDW) solution.

BI Management and Support
The BI needs frequent maintenance and support after it is implemented. This maximizes the value of business intelligence and data warehouse (BIDW) investment with 24x7 production support, maintenance and administration services.

The Business Intelligence Analyst should be able to think strategically about business issues and understand product & service portfolios as well as emerging technologies and their benefits to customers.

In addition to being a professional in the Intelligence, the Business Intelligence Analyst partners closely with the Business Analysts and Product Managers to support the release of new products and proof-of-concept pilots.

4. Strategy of Implementation

A. Background: This includes identifying the background of the business which includes the needs of the business clients.
B. Obtaining consistent data and creating meaningful reports. This has been a major challenge for most of the organizations as the information needs keeps changing for various reasons and very frequently, no consensus of what is the 'appropriate' data for different questions, and only few individuals have been able to access and make use of data successfully.

Decision makers have routinely depended on IT experts to compile data, sometimes waiting days or weeks for the needed answers. However, day-to-day activities demand timely and accurate information and increasing demand for more information to make business decisions which at present has been very difficult to provide on short notice.

Enormous amounts of integrated data will lead to faster and better decision making. The objectives of the Business Intelligence (BI) Implementation are to provide a new approach to data management, presentation and analytics. Implementing Business Intelligence enables to access and retrieve financial, research, personnel and miscellaneous data from one single source, give easy and secure access of relevant data to all levels of management or units, spend minimal time on data retrieval, but dedicate significant time to data analysis and decision making, make data, reports, analysis and models available to a broad user base. Business intelligence primarily focuses on processes and people. To this purpose the BI team will need to ensure data conversion and data warehousing consistent with the needs of current and future organizational goals.

Business Intelligence effort encompasses data conversion and cleansing, data warehousing and data consumption. Only "good" data will be converted and migrated into the data warehouse; older (non-cleansed) data is stored for historical purposes only.

The data warehouse allows secure storage and access to data from the new transactional systems and will allow additional storage of other applications and individual departmental needs. Rapid advances in processing speed, innovative data warehouse design and pioneering concepts of data display such as dashboards are opening up new opportunities for data management.

5. Objectives of Business Intelligence

The objective of BI is to increase its operational effectiveness. Decision makers need to have easy access to data that is timely, accurate and integrated. Although there is no absolute dollar value assessed for the BI effort, it is clear that integrated data

(1) minimizes the need for duplicate data entry and reconciliation of inconsistent information thus reducing manpower needs,

(2) drives management decisions to improve operation, and

(3) improves the intelligence

It should be noted as well that BI seeks to minimize the "cost of distrust." If the integrated data are not trusted, then users will seek the ability to "massage" that data based on the perception or the actuality of its inaccuracy, and frequently then to optimize a portion of the overall business. But this optimization of the part does not optimize the whole—if data inaccuracies are not corrected at their source (within the transactional systems and associated business processes) then the cycle will continue.

BI seeks fundamentally to address this cycle by providing high-quality data and tools and associated business processes, so that the BI way is clearly recognized as the superior way. Goals for BI are at two levels. Firstly, there is the goal to enable a smooth (i.e. non-interrupted) transition from the old to the new system, so as not to disrupt the current operation. Secondly, a more challenging goal is to create a data environment that will answer questions that have not been previously possible. Much of the implementation plan focuses on data migration, data warehousing and reporting.

As vast amounts of data are collected, new opportunities to analyze and model data will present themselves. The now ready availability of statistical tools to perform data mining will allow the exploration of the relationships among data in order to extract value and ultimately new insights. A plan to manage the transition and not interrupt the flow of business reporting is key to the success of the BI implementation.

6. Audience of Business Intelligence

It is important to understand the audience of the business intelligence project or data. This would help in analyzing the requirements of different business users and building an efficient database which is helpful to the actual end users.

Power-users / developers: This can be a group of technicians that produce reports that they use themselves or that they distribute to others. The project will serve this audience by providing access to a richer set of transaction data and by providing technological efficiencies that allow them to spend less time 'pulling' information and more time 'analyzing' information.

Operational Decision Makers: This is a group of managers or leaders in administrative and academic organizations. These managers, directors, business officers, departmental managers who are typically dependent upon dedicated reports today. This leads to a disparity in access to and use of information for day-to-day decision making. The project will serve this audience by providing them a set of dashboard reports that they can individually access through an easy-to-use web self-service reporting portal.

Executive Decision Makers: This group of deans, AVPs and VPs are reliant upon dedicated report writers today as well. This audience generally does not have 'direct' access to information, but their position means that they generally have programming staff who can provide information to support their decisions. Unfortunately, this information frequently needs to be 'cobbled together' from a number of different sources, lacks context (relationship to the same number last year or in another unit), and is generally provided on-demand only (i.e., no notifications or early warnings). The project will serve this audience by providing self-service dashboard reports that blend current information with contextual information. BI will also serve this audience by beginning to produce a dashboard that quickly displays some of the indicators.

There is in reality an overlap between the different audiences. Generally, BI intends to provide a richer set of information to a significantly broader audience and to shift the allocation of time away from 'pulling and maintaining' information and toward analyzing and using information.

It is also important to note that self-service reporting provided through the BI project is not just a tool technology, or project – it is a disruptive change to decision makers at all levels.

7. Access and Security of Business Intelligence

From an access and security standpoint the vision is to ensure that decision makers at all levels have access to data needed to perform their jobs. Any data directly accessible by the public should be highly controlled. General security should be provided through encryption for users and for applications that access information.

Authentication should be provided to the web application. Access to highly sensitive data such as social security numbers, credit card numbers, banking information, driver license numbers and benefit information should be granted on a 'need only' basis and available only to the authorized users.

To mitigate the risk of inappropriate use of data, the BI team advocates that a balance be struck between electronically restricting data access and training users to handle data responsibly. Besides the enormous cost, if too many technology barriers are built into the system, data users will circumvent the new data system, and proliferation of different data repositories will continue, thus defeating the purpose of BI.

8. Success Critieria of Business Intelligence

Success criteria of a business intelligence application include:
- Positive feedback from users on vastly improved access to data and information
- Perfect reproducibility and consistency of data when queried though different channels
- Improved communications and decision making at all levels.

9. Methodology of Data Warehouse

This section explains the steps to develop and deploy a data warehouse. In addition it will show differences between DW Methodology and Traditional IT Methodologies.

The Data Warehousing Methodology is organized into the different phases. Like all software implementations, Business Intelligence implementations follows all Software Development Life Cycle Management guidelines, which start with establishing scope or goals and identifying key stakeholders, users and managers to guarantee the credibility and commitment with Business Intelligence project.

Developing data warehouses is definitely different than developing other IT systems and so requires a different methodology.

Data Warehousing Methodology:
- Use of data is exploratory and less predictable
- Multidimensional Modeling
- Focus is on loading and presenting data

Traditional IT Methodology:
- Automated processes are repeated and predictable
- ERD Data Modeling
- Focus is on rapid on-line updating of data

Data warehousing is not simply creating a set of reports that are run periodically. It involves questions that may lead to initially unpredicted places. Requirements describe the needed solution in business terms. In the analysis phase detailed requirements for data warehouse are defined.

Acknowledgements

References

1. ^ (Rud, Olivia (2009). Business Intelligence Success Factors: Tools for Aligning Your Business in the Global Economy. Hoboken, N.J: Wiley & Sons. ISBN 978-0-470-39240-9.)

2. ^ a b D. J. Power (10 March 2007). "A Brief History of Decision Support Systems, version 4.0". DSSResources.COM. Retrieved 10 July 2008.

3. ^ Kobielus, James (30 April 2010). "What's Not BI? Oh, Don't Get Me Started....Oops Too Late...Here Goes....". ""Business" intelligence is a non-domain-specific catchall for all the types of analytic data that can be delivered to users in reports, dashboards, and the like. When you specify the subject domain for this intelligence, then you can refer to "competitive intelligence," "market intelligence," "social intelligence," "financial intelligence," "HR intelligence," "supply chain intelligence," and the like."

4. ^ H P Luhn (1958). "A Business Intelligence System". IBM Journal 2 (4): 314. doi:10.1147/rd.24.0314.

5. ^ Power, D. J.. "A Brief History of Decision Support Systems". Retrieved 1 November 2010.

6. ^ Evelson, Boris (21 November 2008). "Topic Overview: Business Intelligence".

7. ^ Evelson, Boris (29 April 2010). "Want to know what Forrester's lead data analysts are thinking about BI and the data domain?".

8. ^ Henschen, Doug (4 January 2010). Analytics at Work: Q&A with Tom Davenport. (Interview).

9. ^ "Are You Ready for the New Business Intelligence?". Dell.com. Retrieved 2012-06-19.

10. ^ Jeanne W. Ross, Peter Weil, David C. Robertson (2006) "Enterprise Architecture As Strategy", p. 117 ISBN 1-59139-839-8.

11. ^ a b c d Swain Scheps "Business Intelligence For Dummies", 2008, ISBN 978-0-470-12726-0

12. ^ a b Watson, Hugh J.; Wixom, Barbara H. (2007). "The Current State of Business Intelligence". Computer 40 (9): 96.doi:10.1109/MC.2007.331.

13. ^ Pendse, Nigel (7 March 2008). "Consolidations in the BI industry". The OLAP Report.

14. ^ Imhoff, Claudia (4 April 2006). "Three Trends in Business Intelligence Technology".

15. ^ a b c Rao, R. (2003). "From unstructured data to actionable intelligence". IT Professional 5 (6): 29.doi:10.1109/MITP.2003.1254966.

16. ^ a b c Blumberg, R. & S. Atre (2003). "The Problem with Unstructured Data". DM Review: 42–46.

17. ^ a b Negash, S (2004). "Business Intelligence".Communications of the Association of Information Systems 13: 177–195.

18. ^ a b Inmon, B. & A. Nesavich, "Unstructured Textual Data in the Organization" from "Managing Unstructured data in the organization", Prentice Hall 2008, pp. 1–13

19. ^ Gartner Reveals Five Business Intelligence Predictions for 2009 and Beyond. gartner.com. 15 January 2009

20. ^ Campbell, Don (23 June 2009). "10 Red Hot BI Trends".Information Management.

21. ^ Wik, Philip (11 August 2011). "10 Service-Oriented Architecture and Business Intelligence". Information Management.

22. ^ Rodriguez, Carlos; Daniel, Florian; Casati, Fabio; Cappiello, Cinzia (2010). "Toward Uncertain Business Intelligence: The Case of Key Indicators". IEEE Internet Computing 14 (4): 32.doi:10.1109/MIC.2010.59.

23. ^ Rodriguez, C., Daniel, F., Casati, F. & Cappiello, C. (2009),Computing Uncertain Key Indicators from Uncertain Data, pp. 106–120 | conference = ICIQ'09 | year = 2009

[1] S.Nasira tabassum has received her Masters Degree in computer application from M.J.College of Science and Technology, Affiliated to Osmania University, Hyderabad and Pursuing M.Tech in Software Engineering from Nizam Institute of Engineering and Technology, Deshmukhi, Nalgonda Dist, Affiliated to JNTU Hyderabad, AP India.

[2] SK.Althaf ahammed has received his Masters degree in computer application from m.k.university madhurai (T.N). He is currently working as an Associate professor in Shadan Group of Colleges, Hyderabad.