



# Business Intelligence: Concepts and Explanation

Pratik Mankar<sup>1</sup>, Vibhor Dharmadhikari<sup>2</sup>, Syed Mohammad Hamza<sup>3</sup>  
BE Student<sup>1, 2, 3</sup>

Department of Computer Technology  
Yeshwantrao Chavan College of Engineering, Nagpur, India

## Abstract:

For companies which maintain contact with large numbers of clients, a new data management challenge is presented: finding an effective way of integrating enterprise applications in real time. To learn from the past and forecast the future, and to achieve this companies are adopting Business Intelligence (BI). The paper explores the concepts of BI and provides a demonstration for development of an informative dashboard on World Development Indicators.

**Keywords:** business intelligence; data management; real time; enterprise; world development indicators.

## I. INTRODUCTION

Business intelligence (BI) has two basic meanings related to the use of the term intelligence. The primary term is the human intelligence capacity applied in business affairs and activities. The second relates to the intelligence as information valued for its currency and relevance. It is expert information, of organizational and individual business. In modern businesses, increasing standards, automation, and technologies have led to vast amounts of data becoming available. Data warehouse technologies have set up repositories to store this data. Improved extract, transform, load (ETL) processes have increased the speedy collecting of data. OLAP reporting technologies have allowed faster generation of real time reports which analyze the data. Business intelligence has now become the art of sifting through huge amounts of data, extracting essential information, and turning that information into knowledge upon which actions can be taken. Section I evaluates the concept Business Intelligence, its components, benefits and its emergence in the Market. Section II deals with the implementation of term Business Intelligence for analysis of World Development Indicators. Robert Stackowiak defines Business intelligence as the process of taking large amounts of data, analyzing that data, and presenting a high-level set of reports that has the essence of that data into the basis of business actions, enabling management to make fundamental daily business decisions[1].

Stackowiak view BI as way of improving business performance by providing powerful assistance for executive decision maker which would enable them to have actionable information at hand. BI is seen as technology that enables the efficiency of business operation by providing an increased value to the enterprise information. The concept of Business Intelligence (BI) is brought up by Gartner Group in 1996. It is defined as the application of a set of technologies, such as J2EE, DOTNET, Web Services, XML, data warehouse, OLAP, Data Mining, etc. to improve enterprise operation effectiveness, efficiency productivity, and support for decision to achieve competitive advantages[6]. BI includes several software for Extraction, Transformation and Loading (ETL), data warehousing, database query and reporting, multidimensional/on-line analytical processing (OLAP) data analysis, data mining and visualization[2].

### A. OLAP (Online Analytical Processing)

Online Analytical Processing or OLAP provides multidimensional, summarized views of business data and is used for reporting, analysis, modeling and planning for optimizing the business. It provides a way in which business users can slice and dice their way through data using sophisticated tools that allow for traversing of dimensions such as time or hierarchies. OLAP techniques and tools can be used to work with data warehouses or data marts designed for sophisticated enterprise systems. These systems process queries required to discover trends and analyze critical factors. Reporting software generates views of data to keep the management informed about the state of their business[5].

### B. Advanced Analytics

It is referred to as data mining, forecasting or predictive analytics, this takes advantage of statistical analysis techniques to predict or provide certainty measures on facts.

### C. Corporate Performance Management (Portals, Scorecards, Dashboards)

This general category usually provides a container for several pieces to combine into one so that the aggregate tells a story. For example, a balanced scorecard that displays states for financial metrics combined with customer and growth metrics.

### D. Real time BI

It allows for the real time distribution of metrics through email, messaging systems and/or interactive displays.

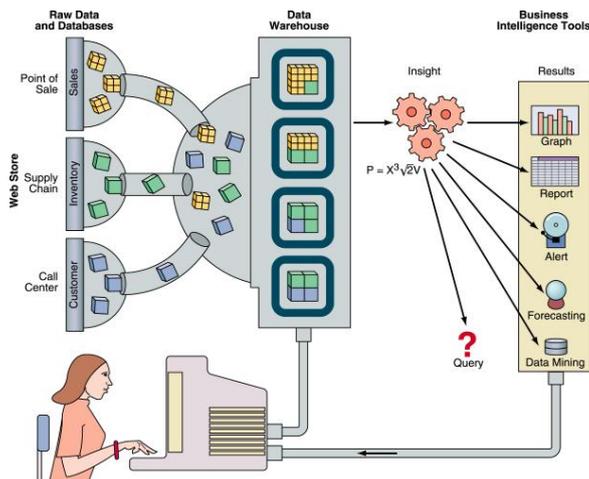
### E. Data Warehouse and Data Marts

The data warehouse is the significant component of business intelligence. The data warehouse supports the physical movement of data by handling the numerous enterprise records for integration, cleansing, aggregation and query tasks. It can also contain the operational data which can be defined as an updateable set of integrated data used for enterprise widerun time decision-making of a particular subject area. It contains live data, not snapshots, and retains minimal history. A data mart as described by Inmon, is a collection of subject areas organized for decision support based on the needs of a given department. Finance has their data mart, marketing has theirs, and sales have theirs and so on. And the data mart for marketing only faintly resembles anyone else's data mart. The

key difference is that the creation of a data mart is predicated on a specific, predefined need for a certain grouping and configuration of select data[3].

**F. Data Sources**

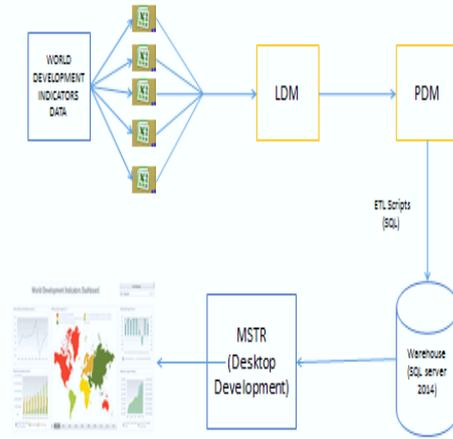
Data sources can be operational databases, historical data, external data for example, from market research companies or from the Internet, or information from the already existing data warehouse environment. The data sources can be relational databases or any other data structure that supports the line of business applications. They also can reside on many different platforms and can contain structured information, such as tables or spreadsheets, or unstructured information, such as plaintext files or pictures and other multimedia information. Businesses realize that in this very competitive, fast paced and ever-changing business environment, a key competitive quantity is how quickly they respond and adapt to change. Business intelligence enables them to use information gathered to quickly and constantly respond to changes. Fig.1 presents an understanding of BI. A BI system in other words is a combination of data warehousing and decision support systems. The figure also reveals how data from different sources can be extracted and stored to be retrieved for analysis. The basic BI functions and reports are shown in fig 1. Business Intelligence and data mining is a field that is heavily influenced by traditional statistical techniques, and most data-mining methods will reveal a strong foundation of statistical and data analysis methods. Some of the traditional datamining techniques include classification, clustering, outlier analysis, sequential patterns, time series analysis, prediction, regression, link analysis (associations), and multidimensional methods including online analytical processing (OLAP).



**Figure.1. A Basic Understanding of BI**

**II. IMPLEMENTATION**

The final result of any Business Intelligence project is to give Decision Maker a platform for doing analysis and making decisions. And dashboard is a perfect model for delivery of any BI Project. A dashboard is a user interface that, somewhat resembling an automobile's dashboard, organizes and presents information in a way that is easy to read. To develop a sense for the concepts and terms used in this paper we are taking an example for developing a dashboard related to World Development Indicators. Our aim is to create such an informative dashboard which will enable end user to understand the basics of how the World Bank collects data and analyzes it to feature its predictions and researched derivatives[4].



**Figure. 2. BI Process flow**

Fig.2 presents the process flow. From building multidimensional models to dumping the data in a database and then creating a dashboard, the process flows through various stages of development. The first step involved in this process is finding the dataset from different heterogeneous sources. For the example dataset used is the latest dataset from World Bank about the development states of various countries all around the world. The dataset is known as ‘World Development Indicators – 2015’. The data was taken from single source hence the following step was to build a multidimensional data model which include Logical data model and Physical data model. The initial dataset was inconsistent, complex and contains lots of garbage values and empty values. Hence ETL process is performed on the data. After this we have to perform data warehousing on the data. Hence the data warehouse used for this purpose was Microsoft SQL Server.

**A. Identification of KPI's**

KPI's or Key Performance Indicators are quantifiable measure used to evaluate the success of an organization, employee, etc. in meeting objectives for performance. Some of the important KPI's included are: Literacy Rate, Employment, Population, Access to Electricity, Agriculture Land, Forest Land, GDP, CO2 Emissions etc. After the identification of KPI's data needs to be pulled from data warehouse to Microstrategy. Where we need to create Microstrategy objects.

**B. Microstrategy Object's**

MicroStrategy is an enterprise business intelligence (BI) application software vendor. The MicroStrategy platform supports interactive dashboards, scorecards, highly formatted reports, ad hoc query, thresholds and alerts, and automated report distribution. Interfaces include web, desktop (for developers) and Microsoft Office integration. Some of the Microstrategy objects are:

- **Attributes:** They represent the granularity of the data in the facts table. They are generally the descriptive data from the business.
- **Facts:** They are the numeric values which can be aggregated to represent the value of some business data.
- **Hierarchies:** They represent the relationship between the various attribute values. The help in carrying out drill-up and drill-down analysis on the data.
- **Filters:** Filters usually restricts the data on the top of the report. It extract required rows from the underlying tables, it will affect the where clause in SQL.

- **Metrics:** Metrics are MicroStrategy objects that represent business measures and key performance indicators. From a practical perspective, metrics are the calculations performed on data stored in your database, the results of which are displayed on a report.
- **Reports:** A report is a MicroStrategy object that represents a request for a specific set of formatted data from your data source. These objects are generalized into three forms Configuration, Public and Schema objects. The definitions of all these Microstrategy objects are stored in the Metadata. MicroStrategy metadata also contains data warehouse connection information, project settings and microstrategy object definitions. Fig 3 represents the layout of the dashboard that drives three different modules that includes overview, detailed and comparator. Overview tab contains scorecard showing top countries related to specific indicators and a heat map which drives the dashboard according to the region. Brief tab is giving the detail analysis on the basis of particular indicator with respect to particular region, country and specific time frame. The brief tab is specified into four domains namely Demography, Environment, Energy, Economy. Comparator module is for comparison between any two things based on indicators, regions, countries etc. It completely depends on the choice of the user.



Figure.3. Snapshots of Dashboard

### III. CONCLUSION

The results obtained from the project will enable end user to understand the basics of how the World Bank collects data and analyzes it to feature its predictions and researched derivatives. It can also be used by govt. agencies, corporates, and ngo's to measure the growth of a country, region or the world at various levels. As the first module 'overview' includes the highest and greatest countries based on a particular indicator, it gives very unexpected results. For eg. Bahrain is the most polluting

country as per the results of 2000 and 2015. It is out of the imagination. Hence, this overview is useful for judging such countries based on some specific indicators. Overview module particularly shows the comparison of the country for year 2000 and 2015. The second module of the dashboard includes four sub-modules i.e, economy, environment, demography and energy. The visualization obtained from this module is divided into those four parts. Each part contains specific indicator that is related to the sub module. If end user wants the result of population of a particular country of a particular region in a particular year, the visualization will appear in terms of graphs and pie charts for that demand. This is applicable for all the four parts. The end result for this module helps the any type of end user mentioned above, for any type of information related to their requirement. Data is restricted only on the basis of indicators. The time slider is also applied so that if end user wants the data for particular time span only, it will be visualized in that manner. Hence, it is useful for users to analyze and understand the data very easily. It also helps in making comparison between any two countries according to a particular indicator.

### IV. FUTURE SCOPE

The power, scope, and scale of both technology and data collection has been growing exponentially. Can be used to forecast the change in trends of indicators for particular country. The potential for these dashboards is very high. This dashboard can be used for various purposes in the future. So they can bubble up points of interest and identify patterns in raw data that even the most well equipped data scientist will not be able to find manually. It's clear that there's a trend in bringing wider access to more people as well as to organizations previously too small to use legacy systems.

### V. BENEFITS OF BI

- With superior BI tools, now employees can also easily convert their business knowledge via the analytical intelligence to solve many business issues, like increase response rates from direct mail, telephone, e-mail, and Internet delivered marketing campaigns.
- With BI, firms can identify their most profitable customers and the underlying reasons for those customers' loyalty, as well as identify future customers with comparable if not greater potential.
- Analyze click-stream data to improve e-commerce strategies.
- Quickly detect warranty-reported problems to minimize the impact of product design deficiencies.
- Discover money-laundering criminal activities.
- Analyze potential growth customer profitability and reduce risk exposure through more accurate financial credit scoring of their customers.
- Determine what combinations of products and service lines customers are likely to purchase and when.
- Analyse clinical trials for experimental drugs.
- Setmoreprofitablerates for insurance premiums.

### VI. REFERENCES

[1]. Adelman Sid , Moss Larissa and Barbusinski Les. (2002) "I found several definitions of BI", DM Review. Retrieved 17 August 2002 from [http://www.dmreview.com/article\\_sub.cfm?articleId=5700](http://www.dmreview.com/article_sub.cfm?articleId=5700)

[2]. Berson Alex, Smith Stephen and Thearling Kurt. (2002) '*Building Data Mining Applications for CRM*', Tata McGraw Hill.

[3]. Curt Hall. (1999) 'Data Warehousing for Business Intelligence'. Retrieved 20 March 1999, from <http://www.cutter.com/itreports/RP68E.pdf>.

[4]. Davenport, T.H. (1993) 'Process Innovation: Reengineering Work through Information Technology', Harvard Business School Press, Boston.

[5]. Gangadharan.G.R. and Swamy, N., Sundaravalli. (2004) 'Business Intelligence Systems: Design and Implementation Strategies', Proceedings of 26<sup>th</sup> [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1372391](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1372391) International Conference on Information Technology Interfaces, Cavtat, Croatia. Retrieved 15 March, 2007 from

[6]. Goebel Michael and Le Gruenwald (1999) , 'A survey of data mining and knowledge discovery software tools', Volume 1 , Issue 1 (June 1999) Publisher ACM New York, NY, USA.