ISSN (Print) : 2320 – 9798
ISSN (Online): 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
*Vol. 1, Issue 4, June 2013*

# BUSINESS INTELLIGENCE IN A NUTSHELL

Navita Kumari

Research Scholar, Guru Kashi University, Bathinda, India.

**ABSTRACT:** Business Intelligence is gaining popularity in organizations. Business Intelligence systems are developed to help the organizations understand their customers, operations and performance. BI systems act as measurements units for a business or an organization. Poor intelligence leads to poor decision making. BI has problems like data quality, which an organization has to overcome, in order to improve its decision making. A serious effort has been made to understand the cause of poor quality business analytics over the past several years. Most organizations and analysts agree that the main reason of erroneous reporting is the operational data which is used for generating these analytical reports as the data is filled with errors, duplications and inconsistencies. This research paper explores the concepts of BI, data quality characteristics and data quality issues, types of BI tools and delivery mechanisms and the key features of BI architecture.

**Keywords:** Business Intelligence (BI), Data Quality (DQ), Decision Support System (DSS), Enterprise Data Warehouse (EDW), Extract Transform Load (ETL), Key Performance Indicator (KPI), Online Analytical Processing (OLAP).

## I. INTRODUCTION

Business intelligence (BI) is defined as the ability for an organization to take all its processes and capabilities and then convert these into knowledge, ultimately getting right information for the right people, at the right time, through the right channel. This produces large amounts of information which can lead to the development of new opportunities for the organization. When these opportunities have been identified and a strategy has been effectively implemented, BI can provide an organization with a competitive advantage in the market, and stability in the long run (within its industry) [17]. 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) [5].BI tools are a type of application software designed to retrieve, analyze and report data. The tools generally read data that have been previously stored, often, though not necessarily, in a data warehouse. BI has experienced high growth and BI technologies have gained lot of popularity. According to Hancock & Toren, 2006, BI is "a set of concepts, methods, and technologies for turning separated data in an organization into useful information in order to improve business performance"[15]. In a Business Intelligence environment, data from various sources are extracted, transformed and loaded (ETL) into an Enterprise Data Warehouse (EDW) and, from EDW, they are used for generation of reports across the organization.BI process and its various stages are shown in Fig 1. Data quality plays an important and critical role in BI success, since poor data quality can affect business decisions at all levels of the organization and such poor quality data also hinders the growth of organization.

The fact is that the business runs on data; it acts as fuel for the corporate industry engine. A company cannot understand its customers, suppliers, competitors or its own people, processes, and company performance without good data. Therefore, business and IT should work together to ensure high-quality data [11].

## II. REVIEW OF LITERATURE

In a 1958 article, IBM researcher Hans Peter Luhn used the term business intelligence. He defined intelligence as: "the ability to apprehend the interrelationships of presented facts in such a way as to guide action towards a desired goal. [16]"

Business intelligence as it is understood today is said to have evolved from the decision support systems which began in the 1960s and developed throughout the mid-1980s. DSS originated in the computer-aided models created to assist with decision making and planning. From DSS, data warehouses, Executive Information Systems, OLAP and business intelligence came into focus beginning in the late 80s.

ISSN (Print) : 2320 – 9798
ISSN (Online): 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
*Vol. 1, Issue 4, June 2013*

Howard Dresner (1989) proposed "business intelligence" as an umbrella term to describe "concepts and methods to improve business decision making by using fact-based computerized support systems.[9]" It was not until the late 1990s that this usage was widespread.

Stackowiak et al. (2007) defined Business intelligence as the process of taking large amounts of data, analyzing that data, and presenting a high-level set of reports that condense the essence of that data into the basis of business actions, enabling management to make fundamental daily business decisions[20]. Cui et al. (2007) view BI as way and method of improving business performance by providing powerful assists for executive decision maker to enable them to have actionable information at hand[8]. BI tools are seen as technology that enables the efficiency of business operation by providing an increased value to the enterprise information and hence the way this information is utilized.

Zeng et al. (2006) define BI as "The process of collection, treatment and diffusion of information that has an objective, the reduction of uncertainty in the making of all strategic decisions." Experts describe Business intelligence as a "business management term used to describe applications and technologies which are used to gather, provide access to analyze data and information about an enterprise, in order to help them make better informed business decisions."[24]

Tvrdikova (2007) describes the basic characteristic for BI tool is that it is ability to collect data from heterogeneous source, to possess advance analytical methods, and the ability to support multi user's demands. Zeng et al. (2006) categorized BI technology based on the method of information delivery; reporting, statistical analysis, ad-hoc analysis and predicative analysis.[21]

The concept of Business Intelligence (BI) is brought up by Gartner Group since 1996. It is defined as the application of a set of methodologies and technologies, such as J2EE, DOTNET, Web Services, XML, data warehouse, OLAP, Data Mining, representation technologies, etc, to improve enterprise operation effectiveness, support management/decision to achieve competitive advantages.

Golfarelli et al. (2004) defined BI that includes effective data warehouse and also a reactive component capable of monitoring the time critical operational processes to allow tactical and operational decision-makers to tune their actions according to the company strategy [14].

Gangadharan and Swamy (2004) define BI as the result of in-depth analysis of detailed business data, including database and application technologies, as well as analysis practices. They widen the definition of BI as technically much broader tools, that includes potentially encompassing knowledge management, enterprise resource planning, decision support systems and data mining.[13]

Berson et.al (2002); Curt Hall (1999) BI includes several software for Extraction, Transformation and Loading (ETL), data warehousing, database query and reporting, OLAP, data analysis, data mining and visualization[2].

## III. TYPES OF DATA PROBLEMS

There are many types of data problems that effect business and its processes; they can generally be divided into two categories: defective data and inconsistent data [11]. Most of the companies don't take the necessary steps to figure out and address these issues until it has a bad impact on their business.

### *A. Defective Data*

#### *(i)Types of Data Defects*

There are many types of data defects. Inaccurate, incomplete, unavailable or obsolete data is defined as defective data. For example, manually entered data is often full of spelling errors, typographic errors, and perhaps falsified. Sometimes people neglect to fill in all the fields or enter correct data in the wrong fields. Data defects can also appear when a system is moved from one platform to another, an old application is replaced with a new one, or data is moved from one application to another on an automated and scheduled basis.

ISSN (Print) : 2320 – 9798
ISSN (Online): 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
*Vol. 1, Issue 4, June 2013*

*(ii) Fixing Data Defects*

The problem with defective data is that it is difficult to find out once it's entered into the system. The best way to prevent defective data is to prevent it from being entered in the first place, so for this the businesses should invest in systems that validate and fix data at the source when it's entered into the system or moved between systems through an application interface. To fix defective data already in the system, the business must invest in data profiling and cleansing tools to cleanse and validate data sets before they are loaded into data warehouses for further use in BI.

*B. Inconsistent Data*

*(i) External Data*

Inconsistency in data occurs over time as the data is either duplicates or out dated. For example, customer data degrades over time as people marry, divorce, die, move, or change their names, so, this results in inconsistent data.

*(ii) Internal Data*

Internally, companies fragment data to suit their needs, largely because of the way they are organized and also because they lack centralized data management. Each analyst, department, or division sees the world through its own lens and thus defines and models data differently.

## IV. CHARACTERIZING DATA QUALITY IN BI

Data quality can be measured and quantified according to various parameters. Previous work provides different classifications of the data quality dimensions. By analyzing these classifications, it is possible to define a basic set of data quality dimensions including accuracy, completeness, consistency, timeliness, interpretability, and accessibility, which represents the dimensions considered by the majority of the authors [19].

### *A. BI-specific data quality dimensions*

***Completeness:*** deals with to ensure "Is all the requisite information available? Are some data values missing, or in an unusable state?

Completeness measures to which extent data should have been recorded in a table is effectively present and also checks whether it is according to the warehouse specifications. We refer to vertical incompleteness when we measure completeness of data in a column. Horizontal incompleteness refers instead to the quantity or percentage of entire tuples. Completeness can be modeled as an extensional or an intensional metadata, and at different levels of granularity. Extensionally, completeness for cells is expressed as a binary value (i.e., true/false). Intensional measure of incompleteness refers to a situation where a rule is stored rather than a measure. We can define rules as functions over the dataset that identify a set of tuples, and for these tuples, the rules define a completeness measure in terms of its percentage.

***Consistency:*** Consistency denotes the uniformity of the information in a given table. Syntactic consistency refers to uniformity in the data format for a specific field. This is typically something that is detected and corrected at data cleaning time (e.g. via normalization). Semantic consistency refers to the satisfaction of semantic rules defined over a set of data items.

***Validity:*** Validity refers to the correctness of data.

***Conformity:*** Conformity refers to expectations that data values conform to specified formats? If yes, do all the values conform to those formats?

***Accuracy:*** Do the data objects accurately represent the "real world" values they are expected to model? Incorrect spellings fields and outdated data can impact operational and analytical BI applications. Confidence describes the perceived accuracy and precision of the data, or the degree of trust that the data present in a table or set of tables is accurate.

***Integrity:*** What data is missing important relationship linkages? The inability to link related records together may actually introduce duplication across the systems.

Some other data quality characteristics are ***Assurance, Delivery, Openness/ Transparency, Responsibility.***

ISSN (Print) : 2320 – 9798
ISSN (Online): 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
*Vol. 1, Issue 4, June 2013*

## V. CATEGORIES OF BUSINESS INTELLIGENCE

*A. Types of business intelligence tools*

The key general categories of business intelligence tools are [4]:

- Spreadsheets
- Reporting and querying software: tools that extract, sort, summarize, and present selected data
- OLAP: Online analytical processing
- Digital dashboards
- Data mining
- Data warehousing
- Decision engineering
- Process mining
- Business performance management
- Local information systems

Incremental and Iterative process followed in BI architecture is shown in Fig 2.

*B. BI Delivery Mechanisms*

**Scheduled Reports**: These reports are generated at regular intervals and provide the "as – is" information to business. These typically answer the question – 'What happened?'

**Ad hoc/User Query Tools**: With the help of the available BI tools, business can query the underlying data and make decisions. These typically answer the question – 'Why did it happen?'

**Dashboards**: These are BI solutions where analysis and reporting tools are used to provide feedback on the achievement of KPIs. They are also used to do ad-hoc reporting to improve organization's ability to make correct decisions. Dashboards can be prepared on claim reporting and settlement time, sales service support team's response times, etc. These typically answer the question 'As a company, how are we performing?'

**Trend Analysis Reports**: These business analytics solutions involve data mining to determine the historic behavior of a group, or the performance of a grouping of people, risks or transaction types. Trend analysis reports can be prepared on segment of customers by risk, types of claims by occurrence, etc.

**Forecasting**: This BI solution allows an insurer to attempt to figure out what will happen in the future based on evaluation of current and historic data. This typically answers the question –'What is going to happen next?'

*C. Key Features of BI Architecture*

Listed below are some of the key features of a flexible and highly adaptive BI architecture:

- High quality data: Data presented to the business should be of high quality and consistent. Users from various groups should not end up arguing regarding whose data is correct. Everyone should get the same numbers when querying a metric.
- Data which is presented to business should be correct, complete and current.
- Data sources should be accurately identified. Business and IT should work together to find most suitable data sources.
- Data storage and maintenance – DBA's should establish processes, checks and balances for data storage and maintenance.
- Selection of suitable ETL (Extract – Transform – Load) Tools: Depending on the organization's requirement appropriate ETL tools should be chosen.

ISSN (Print) : 2320 – 9798
ISSN (Online): 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
*Vol. 1, Issue 4, June 2013*

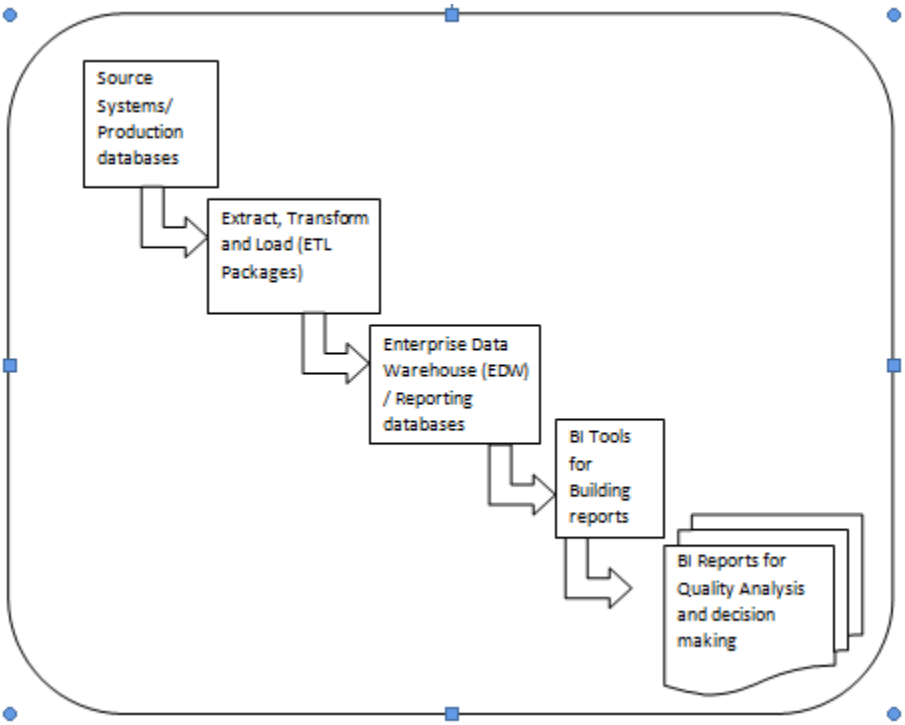


**Fig 1: Business Intelligence stages as data quality sources.**

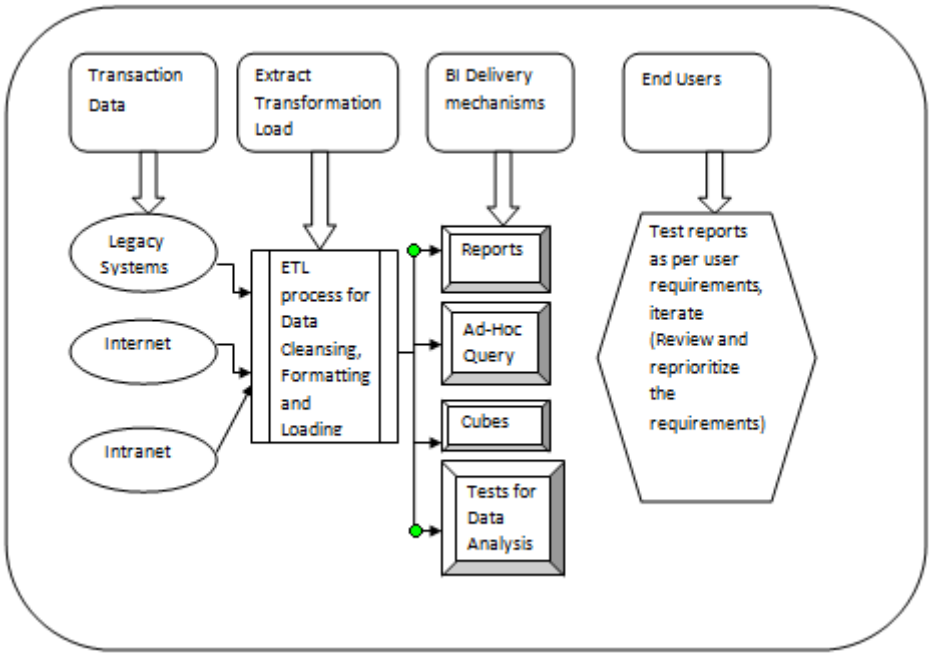


**Fig 2: Incremental and Iterative process followed in BI architecture**

ISSN (Print) : 2320 – 9798
ISSN (Online): 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
*Vol. 1, Issue 4, June 2013*

## VI. AGILE BUSINESS INTELLEGENCE

Agile business intelligence is an approach that combines processes, methodologies, organizational structure, tools, and technologies that enable strategic, tactical, and operational decision-makers to be more flexible and more responsive to the fast pace of changes to business and regulatory requirements[6].

The future of agile business intelligence:

In the past, BI vendors and BI application developers focused on business and operational functionality and architectural robustness. BI vendors and developers now need to concentrate on technologies that are categorized as "agile" and refer to four major subcategories of agility: automated, pervasive, unified and limitless. Each of these new technologies stands on its own and is independent of the others. Firstly, firms need to automate BI processes and steps as fully as possible to eliminate manual work and free up the human resources for analysis and other value-added tasks. Next, as BI initiatives attempt to bridge data and information silos, BI technology itself is not unified. Today, different BI tools address various BI use cases, such as historical and predictive or batch based and real time BI applications. Next-generation BI brings all of them together in a unified platform. After automation and unification, companies should address pervasiveness. How? Make enterprise BI applications available wherever and whenever strategic, tactical, and operational decision-makers need to analyze information, make decisions and act. This category includes technologies such as mobile, cloud and embedded BI. Finally, the next-generation BI to be able to face the challenges of the modern business world - a world that does not fit into nice, neat models - it must operate on information without any borders or restrictions using technologies line in-memory and other BI specific databases.

## VII. CONCLUSION

In summary, comprehensive data quality management is a necessity for BI success in the organization as data quality represents a massive risk for the BI effort. BI processes have to incorporate steps for dealing with data quality issues originating from source systems and assure data quality within BI processes. But data quality management is not only about monitoring and fixing data quality issues. It is also about educating all types of BI users about data quality, data collection from legacy systems and the BI environment including data changes while ETL processes. In this paper we have discussed data quality characteristics, types of data errors, various stages of business intelligence that act as sources for data quality issues. We have also covered the categories and key features of Business intelligence. Further research can be done on various BI tools available and exploring and analyzing which domain they are used, their technical specifications and requirements, analyzing the most widely used BI tool. Another area of further research is mobile, cloud and embedded BI.

### ACKNOWLEDGMENT

I would like to thank Dr. Sawtantar Singh Khurmi, Professor, Bhai Maha Singh College of Engineering, Muktsar for his support, encouragement and valuable comments. He guided me and offered invaluable detailed advices on grammar, organization, and the theme of the paper.

## REFERENCES

[1] Andrea L, Nylund. "Tracing the BI family tree." Knowledge Management, July 1999, URL: http://www.scholarosity.net/documents/dw_family_tree.pdf on September 25, 2012.
[2] Berson, Alex and Stephen J. Smith. Building data mining applications for CRM. McGraw-Hill, Inc., 2002.
[3] Boris Evelson's blog - Forrester Blogs, URL: http://blogs.forrester.com/boris_evelson on September 17, 2012.
[4] Business intelligence tools, URL: http://en.wikipedia.org/wiki/Business_intelligence_tools on September 15, 2012.
[5] Business intelligence, URL: http://en.wikipedia.org/wiki/Business_intelligence on September 15, 2012.
[6] Buyer's Guide: How agility will shape the future of business intelligence, URL: http://www.computerweekly.com/feature/Buyers-Guide-How-agility-will-shape-the-future-of-business-intelligence on August 26, 2012.
[7] Chavaly, Sunita and Vemuri, Satish," Business Intelligence-Driver for Business Transformations in Insurance Industry", Wipro Technologies, Wipro Council for Industry Research, URL: www.wipro.com/industryresearch , pp. 5-7.
[8] Cui, Zhan, E. R. N. E. S. T. O. Damiani and Marcello Leida. "Benefits of ontologies in real time data access." Digital EcoSystems and Technologies Conference, 2007. DEST'07. Inaugural IEEE-IES. IEEE, pp. 392-397, 2007.
[9] Power, Daniel J. "A brief history of decision support systems." DSSResources. COM, World Wide Web, http://dssresources.com/history/dsshistory.html, version 4, March 10, 2007.
[10] Data Quality, URL: http://www.business-intelligence-secrets.com/articles-directory/data-quality/ on August 28, 2012.

ISSN (Print) : 2320 – 9798
ISSN (Online): 2320 – 9801

International Journal of Innovative Research in Computer and Communication Engineering
*Vol. 1, Issue 4, June 2013*

[11] Eckerson Wayne, "Who Ensures Clean, Consistent Data, TDWI Monograph Series", URL: http://www.sas.com/resources/asset/104222_1009.pdf, pp. 5, Sept 2009.
[12] Daniel, Florian, Fabio Casati, Themis Palpanas, Oleksiy Chayka and Cinzia Cappiello. "Enabling Better Decisions through Quality-Aware Reports in Business Intelligence Applications." 13th International Conference on Information Quality, pp. 6-10, 2008.
[13] Gangadharan, G. R. and Sundaravalli N. Swami. "Business intelligence systems: design and implementation strategies." Information Technology Interfaces, 2004. 26th International Conference on Information Technology Interfaces ITI, June 2004.
[14] Golfarelli, Matteo, Stefano Rizzi and Iuris Cella. "Beyond data warehousing: what's next in business intelligence?" Proceedings of DOLAP-04, Washington, DC, USA., www.acm.org, URL: http://student.bus.olemiss.edu/files/Conlon/Others/Others/BusinessIntelligence/Beyond%20data%20warehousing_whats%20next%20in%20business%20intelligence.pdf, May 17, 2006.
[15] Hancock, John Christian and Roger Toren. "Practical business intelligence with SQL Server 2005". Addison-Wesley Professional, 2007.
[16] Luhn, Hans Peter. "A business intelligence system." IBM Journal of Research and Development Vol. 2.4, pp. 314-319, 1958.
[17] Rud, Olivia Parr. Business intelligence success factors: tools for aligning your business in the global economy. Vol. 18. Wiley, Hoboken, N.J: Wiley & Sons. ISBN 978-0-470-39240-9, 2009.
[18] Ranjan, Jayanthi. "Business intelligence: Concepts, components, techniques and benefits." Journal of Theoretical and Applied Information Technology Vol 9.1, 2009.
[19] Singh, Ranjit, and Kawaljeet Singh. "A descriptive classification of causes of data quality problems in data warehousing." International Journal of Computer Science Issues 7.3: 2, pp. 1-2, 2010.
[20] Stackowiak, Robert, Joseph Rayman, and Rick Greenwald. ORACLE DATA WAREHOUSING & BUSINESS INTELLIGENCE SO. John Wiley & Sons, Wiley Publishing, Inc, Indianapolis, 2007.
[21] Tvrdikova, Milena. "Support of Decision Making by Business Intelligence Tools." Computer Information Systems and Industrial Management Applications, CISIM'07. 6th International Conference on. IEEE, pp. 368, 2007.
[22] Umar, Mukhtar Banday and Fizza, Shabir, "Competitive Advantage through Business Intelligence", pp. 7.
[23] Verbitskiy, Yuriy and William Yeoh. "Data quality management in a business intelligence environment: from the lens of metadata." ICIQ 2011: Proceedings of 16th International Conference on Information Quality. University of South Australia, pp. 5-9, 2012.
[24] Zeng, Li, Lida Xu, Zhongzhi Shi, Maoguang Wang and Wenjuan Wu. "Techniques, process, and enterprise solutions of business intelligence." Systems, Man and Cybernetics, 2006. SMC'06. IEEE International Conference on Systems, Man, and Cybernetics October 8-11, Taipei, Taiwan, Vol. 6, pp. 4722, 2006.