BUSINESS INTELLIGENCE: AN INTEGRATED APPROACH

Rafi Ahmad Khan
Scientist
The Business School
University of Kashmir
Email: mca_rafi@yahoo.com

Dr. S.M. K. Quadri
Head
P.G. Department of Computer Sciences
University of Kashmir

Abstract

Business intelligence systems combine operational and historical data with analytical tools to present valuable and competitive information to business planners and decision makers. The objective of Business intelligence (BI) is to improve the timeliness and quality of information, and enable managers to be able to better understand the position of their firm as in comparison to competitors. Business intelligence applications and technologies can help companies to analyze changing trends in market share; changes in customer behavior and spending patterns; customers' preferences; company capabilities; and market conditions. Business intelligence can be used to help analysts and managers determine which adjustments are most likely to respond to changing trends. The emergence of the data warehouse as a repository, advances in data cleansing, increased capabilities of hardware and software, and the emergence of the web architecture all combine to create a richer business intelligence environment than was available previously. In this paper, an attempt has been made to present a framework for building a BI system.

While the business world is rapidly changing and the business processes are becoming more and more complex making it more difficult for managers to have comprehensive understanding of business environment. The factors of globalization, deregulation, mergers and acquisitions, competition and technological innovation, have forced companies to re-think their business strategies and many large companies have resorted to Business Intelligence (BI) techniques to help them understand and control business processes to gain competitive advantage. BI is primarily used to improve the timeliness and quality of information, and enable managers better understand the position of their firm as in comparison to competitors. BI applications and technologies help companies to analyze changing trends in market share; changes in customer behavior and spending patterns; customers' preferences; company capabilities; and market conditions. It is used to help analysts and managers determine which adjustments are most likely to respond to changing trends. It has emerged as a concept for analyzing collected data with the purpose to help decision making units get a better comprehensive knowledge of an organization's operations, and thereby make better business decisions.

BI is an area of Decision Support System (DSS) that which is an information system that can be used to support complex decision making, and solving complex, semi-structured, or ill-structured problems (Azevedo & Santos, 2009; Nematiet al., 2002; Shim, et al., 2002). The first reference to BI was made by Lunh (1958), which has replaced other terms such as Executive Information Systems and Management Information Systems (Negash, 2004; Turban et al., 2008; Thomsen, 2003).

Being rooted in the DSS discipline, BI has suffered a considerable evolution over the last years and is, nowadays, an area of DSS that attracts a great deal of interest from both the industry and researchers (Arnott & Pervan, 2008; Clark et al., 2007; Hannula & Pirttimaki, 2003; Hoffman, 2009; Negash, 2004; Richardson et al., 2008; Richardson et al., 2009). It can be presented as an architecture, tool, technology or system that gathers and stores data, analyzes it using analytical tools, facilites reporting, querying and delivers information and/or knowledge that ultimately allows organizations to improve decision making (Clark et al., 2007; Kudyba & Hoptroff, 2001; Michalewicz et al., 2007; Moss & Shaku, 2003; Negash, 2004; Raisinghani, 2004; Thierauf, 2001; Turban, et al., 2008).

Golfarelli at al., (2004) argue that BI is the process that transforms data into information and then into knowledge.

Although BI is a type of DSS, but it often has a broader meaning. It is the process of gathering high-quality and meaningful information about the subject matter being researched that will help the individual(s) to analyze the information, draw conclusions or make assumptions (Jonathan, 2000).

Stackowiak et al. (2007) opine that BI is 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. Cui et al. (2006) argue that BI is the way and method of improving business performance by providing powerful assistance to executive decision maker which enables them to have actionable information at hand. BI tools are viewed as technology that enhances 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. (2007) have put forth that BI is "The process of collection, treatment and diffusion of information that has an objective, the reduction of uncertainty in the making of all strategic decisions." While (Wu et al. 2007) argues that BI is a "business management term used to describe applications and technologies which are used to gather, provide access to analyzed data and information about an enterprise, in order to help them make better informed business decisions." Van Drunen, (1999) has considered BI as different as its predecessor, "decision support," in that it is a strategic tool intended to help with planning and performance measurement, rather than with Purdy operational decisions. Likewise Cui et al. (2006) argue that BI tools have evolved from being an Executive Information Systems (EIS) and Decision Support System (DSS) to provide much more flair in information delivery and ability to support techniques such as query, reporting, ad hoc analysis and multidimensional analysis which are also known as Online Analytical Processing (OLAP). This variety in capability attracted organizations to start investing in building these types of intelligence systems. However, organizations should have clear BI strategy as a part of IT strategy. While BI is the ability of an organization to understand and use information to its gainful operation (Osterfelt, 2000), the Enterprise BI is a way that brings synergies to business processes and new efficiencies across business areas (Liautaud & Hammond, 2000). BI offers to enterprises "one version of truth", providing consistent and harmonized data to every department in an organization (Bochner & Vaughan 2004). Arents, (2003) argues that there are three important goals that need to be accomplished in order to achieve data consistency across different applications in a complex organization viz:

- Timeliness: the data within system should be synchronized with all other applications;
- Accuracy: the data should encompasses every data from any other application;
- Acceptance: the users convinced of timeliness and accuracy of data, should be able to actively use the system as support for decision making.

The rapidly changing business factors such as globalization, deregulation, mergers and acquisitions and technological innovation, have forced companies to re-think their business strategy. In this competitive environment, BI plays an important role in supporting of the decision making process to augment competitiveness, marking an efficient link between business strategies and IT. BI technology has been continuously expanding and improving to answer more and more complex business. The most widely applied BI enabling technologies, that has emerged include data warehousing (DW), on-line analytical processing (OLAP), and data mining (DM).

BI technology aims to help people make "better" business decisions by making accurate, current, and relevant information available to them when they need it. Competitive organizations accumulate BI in order to assess environment to gain sustainable competitive advantage, and may regard such intelligence as a valuable core competence in some instances.

BI Framework

Although BI is adapted by organizations as per their requirements, history, environment etc to make informed, valuable customer oriented decisions. The main approaches are:

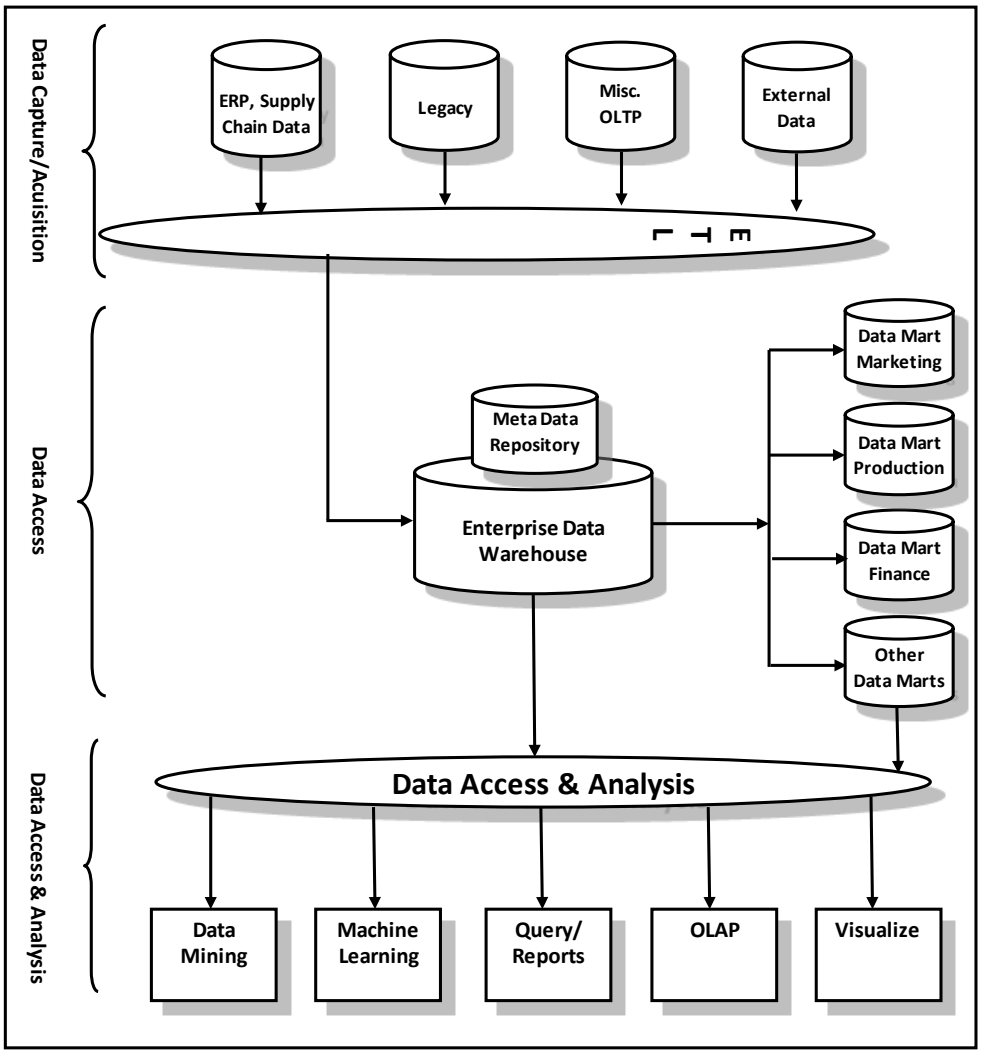
- The traditional approach to BI is concerned with, data aggregation, business analytics and data visualization (Kudyba & Hoptroff, 2001; Raisinghani, 2004; Turban et al., 2008). According to this approach, BI explores several technological tools, producing reports and forecasts, in order to improve the efficiency of the decision making. Such tools include Data Warehouse (DW), Extract-Transform and Load (ETL), On-Line Analytical Processing (OLAP), Data Mining (DM), Text Mining, Web Mining, Data Visualization, Geographic Information Systems (GIS), and Web Portals.

- On the next level there is a concern with the integration of business processes on BI (Eckerson, 2009; Golfarelli et al., 2004; Turban et al., 2008; Wormus, 2008; Zeller, 2007). According to this approach, "BI is a mechanism to bridge the gap between the business process management to the business strategy" (Zeller, 2008). In addition to all the tools in traditional BI, tools such as Business Performance Management (BPM), Business Activity Monitoring (BAM), Service-Oriented Architecture (SOA), Automatic Decision Systems (ADS), and dashboards, are included.
- Adaptive BI is concerned with self-learning adaptive systems, that can recommend the best actions, and that can learn with previous decisions, in order to improve continuously (Michalewicz et al., 2007). Artificial Intelligence is incorporated on BI systems in this manner.

However, the general framework for understanding and guidance of practitioners, academicians and researchers is presented here.

The concept of BI can be decomposed into three parts: (i) Data Capture/Acquisition, (ii) Data Storage and (iii) Data Access & Analysis. Data is collected from internal as well external sources. Internal sources of data are organizations operational database and data warehouse. External data sources include the data from customers, suppliers, government agencies, competitors, internet etc. The collected heterogeneous data is stored in a data warehouse after Extract, Transform and Load (ETL) processes. Finally the data stored in the data warehouse is analyzed for decision making. An attempt has been made in Fig.1 to present a framework of BI, which depicts different components of BI. These components include:

Fig. 1: Framework of Business Intelligence



(i) Data Capture/ Acquisition

The acquisition component is the back end of the data warehousing system and consists of systems that have interface with the operational systems to load data into the data warehouse. Data is first entered or processed by a daily business process that is based on Online Transaction Processing(OLTP) environment and stored in operational database, which may consist of databases such as Oracle, DB2, Informix, SQL Server, SAP R/3, etc. Before data is loaded from operational database and external sources into the data warehouse, it needs to be processed through following stages:

- ***Extraction & Cleanse:*** During data extraction data is acquired from multiple sources including the operational systems. The selected data is consolidated and filtered out from various forms of pollution. Data cleansing validates and cleans up the extracted data to correct inconsistent, missing, or invalid values. This step applies triggers, error reports and corrective processes.
- ***Transform.*** Data transformation integrates data into standard formats and applies business rules that map data to the warehouse schema. Aggregates (e.g., summary table data) and imputed characteristics are generated.
- ***Load.*** Data loading loads the cleansed data into the data warehouse.

(ii) Data Storage

After ETL data is stored in data warehouse or data marts for future analysis.

Data Warehousing: Data warehouse is a copy of transaction data specifically structured for query and analysis and is informational, analysis and decision support oriented, not operational or transaction processing oriented (Kimball, 1996). Corey & Abbey, (1997) views data warehouse as a collection of corporate information derived directly from operational systems and some external data sources. Its specific purpose is to support business decisions, not business operations. Inmon who coined the term "data warehouse" in 1990, argues that a data warehouse is a subject oriented, integrated, time-variant, non-volatile collection of data that is used primarily in organizational decision making (Inmon,1996).

Data warehouses, targeted for decision support, are maintained separately from the operational databases. The architecture of data warehouse can take a variety of forms in practice. But before designing a data warehouse, the requirements and resources of the organization should be taken into consideration. However, some of the options of architecture from which organizations may choose under different circumstances may include: Data Mart; Central Data Warehouse; Distributed Data Warehouse; Virtual Data Warehouse.

Data Marts: Data marts or localized data warehouses are small sized data warehouses, typically created by individual departments or divisions to facilitate their own decision support activities. For example, a data mart can be created for specific products or functions, like customer management, marketing, finance etc.

One of the purposes to build a data mart is to get prototype as soon as possible without waiting for a larger corporate data warehouse, because it's small and easy to develop. But after having several data marts, organizations face operational difficulties in using them in an overall corporate data warehouse strategy, because individual data marts are not consistent with each other.

Metadata: To understand and locate data in the data warehouse users need information about the data warehousing system and its content. This information known as metadata, data about data, includes format, encoding/decoding algorithms, domain constraints, and definitions of the data. It also includes business definitions, data quality alerts, organizational changes, business rules and assumptions, as well as other items of business interest. Metadata help the business user to understand what is available, how to access it, what it means, which data to use, when to use them, etc. Metadata browsers provide an easy to understand view of the data warehouse.

(iii) Data Access and Analysis

The access component of the BI is referred to as the front end. It consists of access tools and techniques that provide a business user with direct, interactive, or batch access to data, while hiding the technical complexity of data retrieval. The interface provides an intuitive, business-like presentation of information, friendly enough for use by a no technical person. This is accomplished by use of BI tools, a suite of software tools that presents a graphical user interface (GUI) with rich reporting and business analysis features. A variety of tools are typically used in an integrated fashion to serve the needs of different groups of users viz:

- Query and reporting tools in packaged software;
- Sophisticated data analysis tools (OLAP/ROLAP);
- Data mining or Knowledge discovery tools;
- Machine Learning tools
- Visualization tools

OLAP: The best known knowledge discovery techniques are Online Analytical Processing (OLAP) and data mining (DM) techniques (Turban et al., 1999). OLAP provide users with the means to explore and analyse large amounts of data, involving complex computations, their relationships, and visually present results in different perspectives. OLAP tools are a combination of analytical processing procedures and graphical user interface. The key features of an OLAP application are: multidimensional views of data, calculation intensive capabilities and time intelligence (Forsman, 1997).

A multidimensional view of data that is usually used in OLAP applications provides quick and flexible access to data and information. Typical applications performed on multidimensional data views are: roll-up (data is summarized with increasing generalization), drill-down (increasing levels of detail are revealed), slice and dice (performing projection operations on the dimensions), and pivoting (cross tabulation is performed) (Jarke et al, 2000). Complex analyses are possible, such as time series (sequence of events) and model charting, forecasting, modelling, statistical, and "what-if" analysis. Analytical processing procedures represent methods of detecting different forms of information needed in the decision process.

Data Mining: It refers to using a variety of techniques to identify nuggets of information or decision-making knowledge in bodies of data, and extracting these in such a way that they can be put to use in the areas such as decision support, prediction, forecasting and estimation. The data is often voluminous, but as it stands of low value as no direct use can be made of it; it is the hidden information in the data that is useful. Data mining, as it is also known, is the nontrivial extraction of implicit, previously unknown, and potentially useful information from data. This encompasses a number of different technical approaches, such as clustering, data summarization, learning classification rules, finding dependency networks, analysing changes, and detecting anomalies. (William et. at ,1994). Data mining is the search for relationships and global patterns that exist in large databases but are hidden among the vast amount of data, such as a relationship between patient data and their medical diagnosis (Holsheimer & Siebes,1994)).

Data mining is concerned basically with the analysis of data and the use of software techniques for finding patterns and regularities in sets of data. It is responsible for finding the patterns by identifying the underlying rules and features in the data. The idea is that it is possible to strike gold in unexpected places as the data mining software extracts patterns not previously discernable or so obvious that no one has noticed them before. Historically the finding of useful patterns in data has been referred to as knowledge extraction, information discovery, information harvesting, data archaeology, and data pattern processing in addition to data mining. As the evolution of data mining has matured, it is widely accepted to be a single phase in a larger life cycle known as Knowledge Discovery in Databases (KDD). The term KDD refers to the broad process of finding knowledge in data stores. The field of KDD is particularly focused on the activities leading up to the actual data analysis and including the evaluation and deployment of results (Collier et. al 1998).

BI solutions are expected to play a significant part in swift response to market demands and in the formulation of the strategy of many of business organizations, with those organizations that will not adapt to take market challenges seriously shall face threat of survival. With easy access to large amounts of complex data from disperse sources, business organizations are able to manage costs and performance, and acquire and increase the profitability.

Machine Learning (ML): ML is part of an emerging Artificial Intelligence (AI) technology that over the past few years has been employed by an increasing number of disciplines to automate complex decision making and problem solving tasks. ML is a family of methods that attempt to allow machines to acquire knowledge for problem solving by showing them historical cases. Among the various methods available, Artificial Neural Network (ANN) is the most popular which has been inspired by the biological neural networks of the human brain and started as an attempt to model the learning capabilities of humans. Other techniques include inductive learning, case-based reasoning, genetic algorithms, NLP etc.

Conclusion: In today's highly competitive world, the quality and timeliness of business information for an organization is not just a choice between profit and loss; it may be a question of survival or bankruptcy. No business organization can deny the benefit of BI. Recent industry analyst reports show that in the coming years millions of people will use BI visual tools and analytics everyday

(Buam, 2006). Today's organizations are deriving more value from BI by extending actionable information to many types of employees, maximizing the use of existing data assets. BI is spreading its wings to cover small, medium and large companies, more and more analytical tools are penetrating the market to do any kind of analysis and help to make informed decision making. The rapidly changing business environment will increase the need for BI. In this paper an attempt has been made to educate the practitioners and academics about the formidable development and application of BI.

References

Arents-Gregory M. (2003). One Version of the Truth, DM Direct, DMReview and SourceMedia, articleID=6359.

Arnott, D., & Pervan, G. (2008). Eigth Key Is- sues for the Decision Support Systems Discipline. Decision Support Systems, 44(3), 657–672. doi:10.1016/j.dss.2007.09.003

Azevedo, A., & Santos, M. F. (2009). Business Intelligence: State of the Art, Trends, and Open Issues. In Proceedings of the First International Conference on Knowledge Management and Information Sharing - KMIS 2009 (pp.296-300).

Bochner, P., Vaughan, J. (2004). BI today: One version of the truth, Application Development Trends, 2004, VOL 11; NUMB 9, pages 18-24, ISSN 1073-9564.

Buam, D. (2006). "The Face of Intelligence." ORACLE Magazine.

Clark, T. D., Jones, M. C., & Armstrong, C. P. (2007). The Dynamic Structure of Management Support Systems: Theory Development, Research, Focus, and Direction. Management Information Systems Quarterly, 31(3), 579–615.

Collier, K., Carey, B., Grusy, E., Marjaniemi, C. & Sautter, D. (1998). "A Perspective on Data Mining", The Center for Data Insight, Northern Arizona University.

Corey, M. J. and Abbey, M. (1997). Oracle Data Warehousing. Osborne McGraw-Hill.

Cui, Z., Damiani, E. and Leida, M. (2007), "Benefits of Ontologies in Real Time Data Access", Digital EcoSystems and Technologies Conference, 2007. DEST '07. Inaugural IEEE-IES, , pp. 392-397, 21-23

Eckerson, W. W. (2009). Research Q&A: Performance Management Strategies. Business Intelligence Journal, 14(1), 24–27.

Forsman, S., (1997). OLAP Council White Paper, OLAP Council, San Rafael CA 1996 www.olapcouncil.org/research/whtpaply.htm.

Golfarelli, M., Rizzi, S., & Cella, I. (2004). What`s Next in Business Intelligence. In DOLAP´04 (pp. 1–6). Beyond Data Warehousing.

Hannula, M., & Pirttimäki, V. (2003). Business Intelligence Empirical Study on the Top 50 Finnish Companies. Journal of American Academy of Business, 2(2), 593–599.

Hoffman, T. (2009). 9 Hottest Skills for '09. Computer World, January 1 (1), 26-27.

Holsheimer, M. , Siebes, A.P.J.M. (1994), " Data Mining: The Search for Knowledge in Databases", CWI (Centre for Mathematics and Computer Science), Amsterdam, The Netherlands .

Inmon, W. H. (1996), Building the Data Warehouse, 2nd ed. ,New York, NY: John Wiley & Sons.

Jarke, M., Lenzerini, M., Vassiliou, Y., Vassiliadis, P., (2000). Fundamentals of Data Warehouses, Springer Verlag.

Jonathan, Wu. (2000), "Business Intelligence: What is Business Intelligence?", DM Review.

Kimball, R. (1996), The Data Warehouse Toolkit : Practical Techniques for Building Dimensional Data Warehouses. John Willy & Sons.

Kudyba, S., & Hoptroff, R. (2001). Data Mining and Business Intelligence: A Guide to Productivity. Hershey, PA: Idea Group Publishing.

Liautaud, B., Hammond, M., (2000). e-Business Intelligence turning information into knowledge into profit, McGraw Hill, New York.

Lunh, H. P. (1958). A Business Intelligence Sys- tem. IBM Journal of Research and Development, 2(4), 314–319. doi:10.1147/rd.24.0314

Michalewicz, Z., Schmidt, M., Michalewicz, M., & Chiriac, C. (2007). Adaptive Business Intelligence. Berlin, Heidelberg: Springer-Verlag.

Moss, L. T., & Shaku, A. (2003). Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications. Upper Saddle River, NJ: Pearson Education.

Negash, S. (2004). Business Intelligence. Communications of the Association for Information Systems, 13(1), 177–195.

Nemati, H. R., Steiger, D. M., Iyer, L. S., & Herschel, R. T. (2002). Knowledge Warehouse: An Architectural Integration of Knowledge Management, Decision Support, Artificial Intel- ligence and Data Warehousing. Decision Support Systems, 33(1), 143–161. doi:10.1016/S0167- 9236(01)00141-5

Osterfelt, S., (2000), Business Intelligence: The Intelligent Customer, DM Review www.dmreview.com/editorial/dmreview/print_action.cfm?EdID=3107

Raisinghani, M. (2004). Business Intelligence in the Digital Economy: Opportunities, Limitations and Risks. Hershey, PA: Idea Group Publishing.

Richardson, J., Schlegel, K., & Hostmann, B. (2009). Magic Quadrant for Business Intelligence Platforms. Core Research Note: G00163529, Gartner.

Richardson, J., Schlegel, K., Hostmann, B., & McMurchy, N. (2008). Magic Quadrant for Business Intelligence Platforms, 2008. Core Research Note: G00154227, Gartner.

Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R., & Carlsson, C. (2002). Past, Present, and Future of Decision Support Technology. Decision Support Systems, 32(1), 111–126. doi:10.1016/S0167-9236(01)00139-7

Stackowiak, R., Rayman, J. and Greenwald, R. (2007), Oracle Data Warehousing and Business Intelligence Solutions, Wiley Publishing, Inc, Indianapolis.

Thierauf, R. J. (2001). Effective Business Intelligence Systems. West Port, CP: Quorum Books.

Thomsen, E. (2003), "BI's Promised Land", Intelligent Enterprise, (6)4, pp. 21-25.

Turban, E., McLean, E., Wetherbe, J. (1999). Information Technology for Management: Making Connections for Strategic Advantage, John Wiley & Sons.

Turban, E., Sharda, R., Aroson, J. E., & King, D. (2008). Business Intelligence: A Managerial Approach. Upper Sadle River, New Jersey: Pearson Prentice Hall.

Van Drunen, H. (Feb,1999) "Three stages can work better as one process. Computing Canada", 25 i5 p 19(l).

William J. F., Gregory P.& Matheus C. J. (1994) "Knowledge Discovery in Databases: An overview", AI Magazine, Volume 13 , Issue 3, Fall 1992.

Wormus, T. (2008). Complex Event Processing: Analytics and Complex Event Processing: Adding Intelligence to the Event Chain. Business Intelligence Journal, 13(4), 53–58.

Wu, L., Barash, G. and Bartolini, C. (2007), "A Service-oriented Architecture for Business Intelligence", Service-Oriented Computing and Applications, SOCA '07. IEEE International Conference, pp. 279-285, 19-20

Zeller, J. (2008). Business Intelligence: The Road Trip. Information Management Special Reports, December 2, 2008, Retrieved from http://www.information-management.com/specialreports/2008112/100002266-1.html.

Zeng, L., Xu, L., Shi, Z., Wang, M. and Wu, W. (2007), "Techniques, process, and enterprise solutions of business intelligence", Vol. 6, pp. 4722.