A Five-Layered Business Intelligence Architecture

In Lih Ong[1], Pei Hwa Siew[1] and Siew Fan Wong[2]

[1]Universiti Tunku Abdul Rahman, Selangor, Malaysia

[2]Sunway University, Selangor, Malaysia

_____

Abstract

Many organizations today have adopted business intelligence (BI) as a catalyst to meet specific business needs and to improve organizational effectiveness. Although BI has become more robust and pervasive, some organizations are still unable to maximize the return on their BI investments. One contributing reason is the lack of a good guiding BI architecture to support the implementation of such a system. Having a solid architecture can help organizations to better control the implementation process as well as the operation of the entire BI environment. A review of the existing literature shows that although the importance of a good BI architecture is non-arguable, research in this area is still lacking. To fill the gap, this paper proposes a framework of BI architecture which consists of five layers: data source, ETL, data warehouse, end user, and metadata layers. These five layers are essential to ensure high data quality and smooth information flow in a BI system.

Keywords: business intelligence, BI architecture, framework, BI system
_____

Introduction

Business intelligence (BI) is "about how to capture, access, understand, analyze and turn one of the most valuable assets of an enterprise - raw data - into actionable information in order to improve business performance" (Azvine et al., 2005, p. 215). It pairs data gathering, data storage, and knowledge management with analytic tools to provide decision-makers with competitive information that often act as differentiators in today's fierce business environment (Negash, 2004). This explains why BI has remained as the top technology priority for Chief Information Officers for the past five years (Gartner Research, 2006, 2007, 2008, 2009a, 2009b). Specifically, a 2009 IBM Global Study of more than 2,500 CIOs found that 83 percent of respondents viewed BI as the most important visionary element in enhancing their ability to compete in the marketplace (IBM, 2009). With more and more organizations becoming aware of the value of BI, its market is expected to grow rapidly. According to Gartner (2011), the BI market will grow 9.7 percent to reach $10.8 billion in 2011. By 2014, this BI market will reach $11.3 billion (MarketResearch.com, 2010) while the revenue for BI vendors will hit $7.7 billion by 2012 (Sommer, 2008).

As organizations begin to adopt BI, one very important task is to make sure that they follow a good BI architectural plan in their implementation process so as to ascertain the success of their BI investment. BI architecture is a framework detailing different components of BI (i.e., data, people, processes, technology, and the management) and how these components need to come together to ensure smooth functioning of a BI system (Rob & Coronel, 2007). Examples of information contained

in a BI architecture are the types of data that need to be collected, the methods to be used to analyze data, and the way to present certain information. Having a solid BI architecture is critical. If the underlying architecture is not designed properly, inconsistencies that arise among the different components may lead to problems such as inability to share information among the components, inability to meet business requirements, and poor business performance. In the worst case, a bad BI architecture may lead to the scenario where wrong information is delivered to the wrong person at the wrong time. Even in the case where BI systems are functional despite bad architecture, organizations will not be able to maximize the value they should have gotten from their BI investments (Rasmussen et al., 2009).

Even though the importance of a good BI architecture is non-arguable, a review of the literature shows that there is still a lack of academic research in the area of BI architecture (Negash, 2004). Therefore, the objective of this paper is to propose a framework of BI architecture which contains five important layers that should be included when implementing BI systems. The rest of the paper is structured as follows. Section 2 presents existing literature on BI architecture while Section 3 describes the proposed BI architecture. Section 4 concludes the paper.

Business Intelligence Architecture

Through literature review, it is found that there are several existing BI architectures (e.g., Baars & Kemper, 2008; Balaceanu, 2007; Shariat & Hightower, 2007; Turban et al., 2008; Watson, 2009). These architectures are different in their structures such as layers, components, processes, and relationships to guide BI implementation efforts (Shariat & Hightower, 2007). However, there are some common components among these BI architectures (e.g., source systems, data storage, and reporting tools). For example, both the architectures of Shariat and Hightower (2007) and Turban et al. (2008) contain data warehouse, end user

applications, and BI portal. Nonetheless, one important component missing from these existing BI architectures is that of analytical and reporting such as data mining, predictive analytics, and data visualization. These features are new BI capabilities that are important and should be included in a BI architecture.

Furthermore, existing BI architectures typically feature a uni-directional communication flow between different components. The architectures proposed in Baars and Kemper (2008) and Shariat and Hightower (2007) are good examples where they only feature a one-way data flow from data sources to data warehouse. The limitation of uni-directional data flow (i.e., no backward data flow from data warehouse to data sources) is that no adjustment or correction is allowed on data source even if an error is found. This may lead to the garbage-in-garbage out situation. If organizations want to correct the error, they have to repeat the entire BI process especially that of the cleansing procedures again. To overcome these problems, Dayal et al. (2009) suggested a two-way data integration flow whereby the cleansed data can be sent back to data sources to improve accuracy and reduce cleansing work.

Another issue with existing BI architectures is the lack of support on metadata management. A good BI architecture should include the layer of metadata. A metadata repository is essential for business users to store and standardize metadata across different systems. By having a well-structured metadata, organizations will be able to track and monitor data flows within their BI environment (Pant, 2009). In addition, they will be able to ensure the consistency of definitions and descriptions of data that support BI components and thus avoid misunderstanding and misinterpretation of data.

Aside from that, some of the architectures do not include operational data store (ODS) within the BI environment. For instance, Watson's BI architecture (2009) contains only data warehouse and data marts

whereas Baars and Kemper (2008) and Turban et al. (2008) include only data warehouse. In order to address operational data needs of an organization, it is essential to implement ODS to provide current or near current integrated information that can be accessed or updated directly by users. Through this way, decision makers will be able to react faster to changing business environment and requirements. Furthermore, it is necessary to consider data staging area in the ETL (Extract-Transform-Load) process. As most of the data from data source require cleansing and transformation, it is important to create a temporary storage for data to reside prior to loading into ODS or data warehouse. Without building this staging area, the process of working on the data from data source and the process of loading the data into data warehouse can be very time consuming and resource-intensive (Melchert et al., 2004).

The Proposed Framework of Business Intelligence Architecture

This paper proposes a framework of a five-layered BI architecture (see Figure 1), taking into consideration the value and quality of data as well as information flow in the system. The five layers are data source, ETL (Extract-Transform-Load), data warehouse, end user, and metadata layers. The rest of this section describes each of the layers.
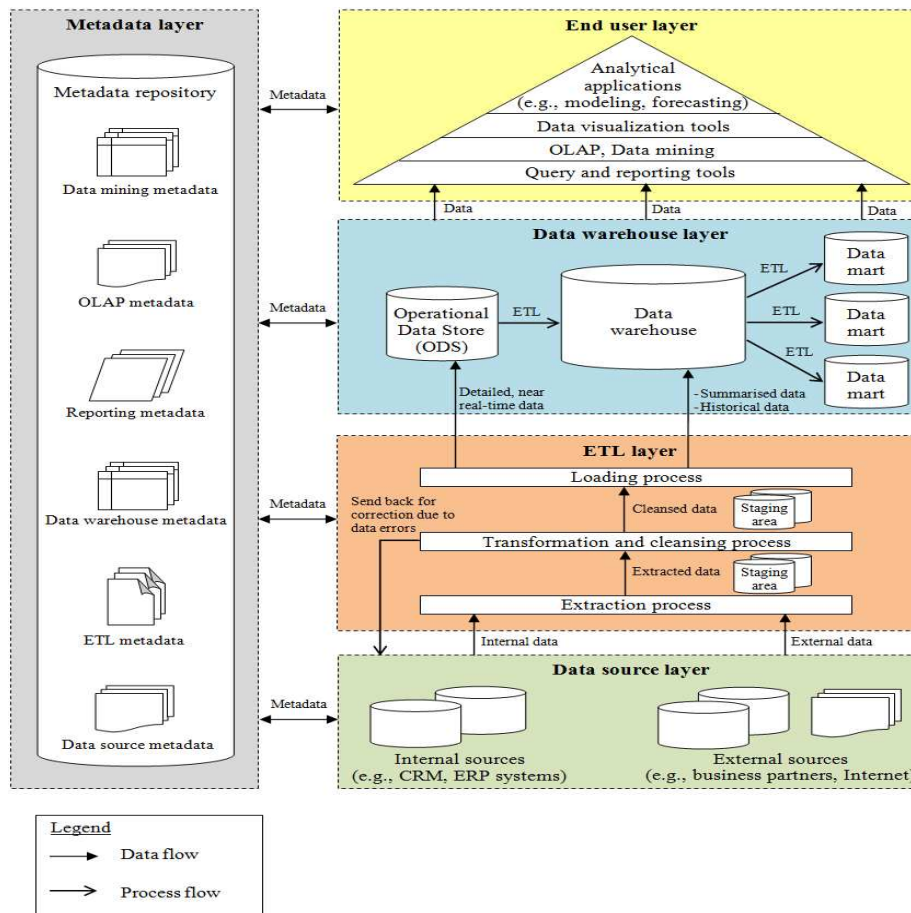

Fig 1. Proposed BI Architecture

Data Source Layer

Nowadays, many application domains require the use of structured data as well as unstructured and semi-structured data to make effective and timely decision (Baars & Kemper, 2008). All these data can be acquired from two types of sources: internal and external. Internal data source refers to data that is captured and maintained by operational systems inside an organization such as Customer Relationship Management and Enterprise Resource Planning systems. Internal data sources include the data related to business operations (i.e., customers, products, and sales data). These operational systems are also known as online transaction processing systems because they process large amount of transactions in real time and update data whenever it is needed. Operational systems contain only current data that is used to support daily business operations of an organization. Generally, operational systems are process-oriented as they focus mainly on specific business operations such as sales, accounting, and purchasing (Hoffer et al., 2007; Imhoff et al., 2003).

External data source refers to those that originate outside an organization. This type of data can be collected from external sources such as business partners, syndicate data suppliers, the Internet, governments, and market research organizations (Ranjan, 2009; Reinschmidt & Francoise, 2000; Strand et al., 2003). These data are often related to competitors, market, environment (e.g., customer demographic and economic), and technology (Haag et al., 2007).

It is important for organizations to clearly identify their data sources. Knowing where the required data can be obtained is useful in addressing specific business questions and requirements, thereby resulting in significant time savings and greater speed of information delivery. Furthermore, the knowledge can also be used to facilitate data replication, data cleansing, and data extraction (Reinschmidt & Francoise, 2000). This is because even though there are many existing data sources, some of them might be inaccessible, unreliable or irrelevant to current business needs. With correct identification of data sources, problems such as inconsistent information, difficulty in finding root causes, and issues of data isolation can be avoided.

ETL (Extract-Transform-Load) Layer

This layer focuses on three main processes: extraction, transformation and loading (Baars & Kemper, 2008; Sen & Sinha, 2005). Extraction is the process of identifying and collecting relevant data from different sources (Reinschmidt & Francoise, 2000). Usually, the data collected from internal and external sources are not integrated, incomplete, and may be duplicated. Therefore, the extraction process is needed to select data that are significant in supporting organizational decision making.

The extracted data are then sent to a temporary storage area called the data staging area prior to the transformation and cleansing process (Ranjan, 2009). This is done to avoid the need of extracting data again should any problem occurs. After that, the data will go through the transformation and the cleansing process. Transformation is the process of converting data using a set of business rules (such as aggregation functions) into consistent formats for reporting and analysis. Data transformation process also includes defining business logic for data mapping and standardizing data definitions in order to ensure consistency across an organization (Davenport & Harris, 2007). As for data cleansing, it refers to the process of identifying and correcting data errors based on pre-specified rules (Reinschmidt & Francoise, 2000). If there is an error found on the extracted data, then it is sent back to the data source for correction (Dayal et al., 2009). Once data have been transformed and cleansed, they are stored in the staging area. This can prevent the need of transforming data again if the loading processes fail or terminate (Kimball & Caserta, 2004). Loading is the last phase of the ETL process. The data in staging area are loaded into target repository.

Data Warehouse Layer

There are three components in the data warehouse layer, namely operational data store, data warehouse, and data marts. Data flows from operational data store to data warehouse and subsequently to data mart.

Operational Data Store

An operational data store (ODS) is used to integrate all data from the ETL layer and load them into data warehouses. ODS is a database that stores subject-oriented, detailed, and current data from multiple sources to support tactical decision making (Imhoff et al., 2003). It provides an integrated view of near real-time data such as transactions and prices. In addition, the data stored in ODS is volatile, which means it can be over-written or updated with new data that flow into ODS (Imhoff et al., 2003; Walker, 2006). As such, ODS does not store any historical data. Generally, ODS is designed to support operational processing and reporting needs of a specific application by providing an integrated view of data across many different business applications (Chan, 2005). It is normally used by middle management level for daily management and short-term decision making (Li et al., 2007). Since the data stored in ODS are updated frequently (i.e., in minutes or hours), it is useful for reporting types that require real time (within 15 minutes) or near time (updated in 15 minutes to 1 hour) information (Walker, 2006).

Data Warehouse

Data warehouse is one of the most important components in BI architecture. Inmon (2005) defines data warehouse as "a subject-oriented, integrated, time-variant, and non-volatile collection of data in support of management's decision making process" (p. 29). The characteristics of a data warehouse are described as follows (Hoffer et al., 2007; Inmon, 2005):

- **Subject-oriented:** Data from various sources are organized into groups based on common subject areas that an organization would like to focus on, such as customers, sales, and products.

- **Integrated:** Data warehouse gathers data from various sources. All of these data must be consistent in terms of naming conventions, formats, and other related characteristics.

- **Time-variant:** Each data stored in the data warehouse has time dimension to keep track of the changes or trends on the data. In other words, data warehouse will store historical changes on each piece of data.

- **Non-volatile:** New data can be added into data warehouse regularly. But, all the data stored in data warehouse are read-only. This means users are not allowed to update, over-write or delete the stored data.

In summary, data warehouse is a central storage that collects and stores data from internal and external sources for strategic decision making, queries, and analysis (Bara et al., 2009; Imhoff et al., 2003). Data warehouse stores aggregated or summarised data. In addition, it also stores large amount of historical data for the purpose of long term analysis (Li et al., 2007). Data are stored in data warehouse longer (5 to 10 years) than in ODS (60 to 90 days) (Chan, 2005). Data in a data warehouse is updated regularly, for instance weekly or sometimes daily (Al-Noukari & Al-Hussan, 2008). As a result, it does not contain the latest data as in operational systems and ODS. Aside from that, data warehouses are designed to support OLAP (Online Analytical Processing) applications by storing and maintaining data in multi-dimensional structures for query, reporting, and analysis (Sen & Sinha, 2005).

Data Mart

While the data in a data warehouse is mainly used to support various needs across the whole organization, it is not equipped to support the needs and requirements of specific departments. Consequently, it is necessary to have data

marts to support them. A data mart is a subset of the data warehouse that is used to support analytical needs of a particular business function or department (Bukhbinder et al., 2005). Like data warehouses, it contains historical data that can help users to access and analyze different data trends (Ranjan, 2009). However, it can only keep data for 60 to 90 days. Therefore, the amount of data stored in a data mart is much lesser than the data stored in a data warehouse. There can be many data marts inside an organization. Data warehouses and data marts are built based on multi-dimensional data model which consists of fact and dimension tables. Fact table contains quantitative data about business entities such as sales amount, quantity, and price. Dimension table contains data (such as product, customer, data, and location) that describes facts (Kimball et al., 2008).

Metadata Layer

Metadata refers to data about data. It describes where data are being used and stored, the source of data, what changes have been made to the data, and how one piece of data relates to other information (Giovinazzo, 2003). Metadata repository is used to store technical and business information about data as well as business rules and data definitions (Davenport & Harris, 2007). Good management and use of metadata can reduce development time, simplify on-going maintenance, and provide users with information about data source (Bryan, 2009). For instance, users do not have to re-design data structure (such as table name and data types) for data modelling since the data structures needed have been stored as metadata. Users can just query and retrieve these metadata from repositories. Therefore, it is essential to ensure that metadata in repositories are maintained and updated regularly.

There are many different types of metadata to support a BI architecture such as data source, ETL, reporting, OLAP, and data mining metadata. Data source metadata consists of information about access mode, structure of data sets (e.g., relational tables,

views, stored procedures), and referential integrity constraints (Ma et al., 2011; Wang & Ye, 2010). As data are integrated into the data warehouse layer using ETL tools, an extraction log is maintained to record the changes made to data element during the extraction process to ensure the quality of data. This log is ETL metadata and it is stored in metadata repository. ETL metadata generally contains information about sources, targets, transformation rules, and mapping. Metadata repository is also used to document the information about data contained in the data warehouse layer. It includes description of data structure (schema, dimensions, and hierarchies) and definitions of conformed dimensions and conformed facts (Chaudhuri & Dayal, 1997; Sen & Sinha, 2005). These metadata guide the process of extracting, transforming, and loading data into target repository (Shariat & Hightower, 2007). OLAP metadata provides descriptions about structure of cubes, dimensions, hierarchies, levels, and the type of drill paths being taken. Data mining metadata include descriptions about algorithms and queries (Nelson, 2008). Reporting metadata are XML-based and are used to store report templates and reporting descriptions such as report name, start date, and end date (Al-Noukari & Al-Hussan, 2008). These metadata also contain information about structures of charts and queries.

End User Layer

The end user layer consists of tools that display information in different formats to different users. These tools can be grouped hierarchically in a pyramid shape (as shown in Figure 1). As one moves from the bottom to the top of the pyramid, the degree of comprehensiveness at which data are being processed and presented increases. This is to tailor to increasing complexity in decision-making as one moves up organizational hierarchy. For instance, the highest level of pyramid consists of analytical applications which are usually used by top management while the lowest level consists of query and reporting tools which are used mostly by operational management level.

Query and Reporting Tools

Query and reporting tools are very useful tools which allow end users to access and query data quickly, and to produce reports for decision making and management purposes. There are many different types of reports including standard reports, ad-hoc reports, budgeting and planning reports, and metadata reports. Both internal and external users can manage reports and other information easier and faster through BI portals. BI portal is a popular end user tool to deliver information. It is a single, secure interface that integrates data and information from various sources so that users can have a one-stop access to different types of information.

OLAP (Online Analytical Processing)

One or more OLAP servers can manage data in the data warehouse layer for reporting, analysis, modelling, and planning to optimize business (Ranjan, 2009). OLAP server is a "data manipulation engine that is designed to support multi-dimensional data structures" (Reinschmidt & Francoise, 2000, p. 13). OLAP server can provide multi-dimensional and summarized views of aggregated data. OLAP is a user-friendly graphical tool that allows users to quickly view and analyze business data from different perspectives. Besides that, OLAP also allows users to easily compare different types of data and complex computations.

In order to reduce query time, data in OLAP server are organized in the form of data cubes instead of tables (rows and columns) as in relational data model (Wang et al., 2005). Data cubes are dimensional models stored in multi-dimensional OLAP structures. They contain fact and dimensional tables to store and manage multi-dimensional data so that users can analyze data easily and in a faster manner (Prevedello et al., 2010). Four basic OLAP operations used in analyzing multi-dimensional data are (Chaudhuri & Dayal, 1997; Han & Kamber, 2006):

- **Roll-up or drill-up:** It increases the level of aggregation, either by moving up to a higher level (more detailed data) along a dimensional hierarchy or by reducing one or more dimensions from a given data cube.

- **Drill-down:** It is the opposite of roll-up. It decreases the level of aggregation by moving down to a lower level (less detailed data) along a dimensional hierarchy or by adding one or more dimensions to a data cube.

- **Slice and dice:** The slice operation can be performed by selecting a specific value on a single dimension, resulting in a sub-cube. The dice operation performs a projection on a data cube by selecting a range of values on two or more dimensions.

- **Pivot:** It enables users to rotate the axes of the data cube, meaning swapping the dimensions to get different views of data.

Data Mining

Data mining process can be achieved with the integration of data warehouses and OLAP servers by performing further data analysis in OLAP cubes. Since the amount of data in an organization is growing rapidly, it is necessary to have data mining to make decisions faster. Basically, data mining is a process that automatically identifies useful information such as unusual patterns, trends, and relationships that are hidden within large amount of data. This can be achieved by applying statistical techniques such as classification, time-series analysis or clustering (Al-Noukari & Al-Hussan, 2008; Kerdprasop & Kerdpraso, 2007; Kimball et al., 2008). Data mining techniques have been used in many application areas such as marketing, financial, medical, and manufacturing to predict future results and summarize details of data (Al-Noukari & Al-Hussan, 2008).

Data Visualisation Tools

Data visualisation tools such as dashboard and scorecards can be provided to managers and executives who need an overall view of their business performance. Dashboard is a useful tool that allows users to visualize data using charts, coloured metrics or tables. Users can also view more detailed information about key performance indicators across their organizations (Ranjan, 2009). By doing so, managers can closely and more effectively monitor their business performance and progress toward defined goals.

Analytical Applications

Analytical applications provide functionalities such as modeling, forecasting, sales analysis, and what-if scenarios (Hobbs, 2007; Parida, 2006; Popovic et al., 2009). These applications can be used to support both internal and external business processes (Davenport & Jarvenpaa, 2008). Applications that are equipped with analytical capabilities allow users to gain insights into improving the performance of business operations. By employing analytical applications, decision makers can also identify and understand what factors drive their business value, and thus able to leverage opportunities faster than their competitors (Avantikumar, 2008).

For a BI system to work smoothly, all five layers described above have to be linked together in a systematic manner. Data originating from internal and external sources have to be extracted, transformed, and loaded into the data warehouse layer. When the data passes through the ETL layer, it can flow to both directions, either to ODS (and then to data warehouse) or to data warehouse directly. Since data warehouse is developed for the usage of the entire organization, data from the warehouse is sent to data marts to fulfill specific operational needs. At end user layer, data in the ODS, data warehouse, and data marts can be accessed by using a variety of tools such as query and reporting tools, data visualisation tools, and analytical applications. Finally, there is a centralized metadata repository that is connected with various components such as the ETL layer, the data warehouse layer, and the end user layer.

Note that the data flows among the components as proposed in the framework for this paper are multi-directional. Such consideration overcomes limitations of uni-directional data flow in many existing BI architectures. A multi-directional flow can enhance query performance and improve accuracy because data error at one layer can be returned to the previous layer for clarification if error occurs. For instance, if an error data is found in the ETL layer, that particular piece of data can be sent back to the data source layer (i.e., internal sources) for modification. Nevertheless, data flow from external sources to ETL layer is only uni-directional. No correction or adjustment is feasible to external source since it originates from the outside of an organization.

Conclusion

This paper has proposed a framework of five-layered BI architecture with various components. BI architecture plays an important role in affecting the success of a BI implementation. To have a smooth BI operation, organizations can benchmark their architectural plan against the framework proposed here. By having a good BI architecture, organizations will be able to maximize the value from their BI investments, and thereby meet their business requirements and improve business performance. However, at this point, the framework proposed in this paper remains conceptual in nature. Though it is built based on existing literature, the framework still needs to be validated using real-life BI cases to affirm its usability. Future research therefore can go along this line to validate the framework.

Acknowledgment

References

Al-Noukari, M. & Al-Hussan, W. (2008). "Using Data Mining Techniques for Predicting Future Car Market Demand," *3rd International Conference on Information and Communication Technologies*: From Theory to Applications, 1-5.

Avantikumar. (2008). 'BI can help in downturn: SAS Malaysia,' [Online], [Retrieved December 31, 2010], http://mis-asia.com/news/articles/bi-can-help-in-downturn-sas-malaysia?SQ_DESIGN_NAM E=print.

Azvine, B., Cui, Z. & Nauck, D. D. (2005). "Towards Real-time Business Intelligence," *BT Technology Journal*, 23(3). 214-225.

Baars, H. & Kemper, H.-G. (2008). "Management Support with Structured and Unstructured Data: An Integrated Business Intelligence Framework," *Information Systems Management*, 25(2). 132-148.

Balaceanu, D. (2007). "Components of a Business Intelligence Software Solution," *Informatica Economica*, 2(42). 67-73.

Bara, A., Botha, I., Diaconita, V., Lungu, I., Velicanu, A. & Velicanu, M. (2009). "A Model for Business Intelligence Systems" Development," *Informatica Economica*, 13(4). 99-108.

Bryan, C. (2009). "Quick Guide to Implementing Business Intelligence, Data Warehousing & BPM," [Online], [Retrieved December 5, 2010],

http://ezinearticles.com/?Quick-Guide-to-Implementing-Business-Intelligence,-Data-Warehousing-and-BPM&id=3465076.

Bukhbinder, G., Krumenaker, M. & Phillips, A. (2005). "Insurance Industry Decision Support: Data Marts, OLAP and Predictive Analytics," *Casualty Actuarial Society Forum*, 171-197.

Chan, J. O. (2005). "Optimizing Data Warehousing Strategies," *Communications of the IIMA*, 5(1). 1-13.

Chaudhuri, S. & Dayal, U. (1997). "An Overview of Data Warehousing and OLAP Technology," *ACM Sigmod Record,* 26(1). 65-74.

Davenport, T. H. & Harris, J. G. (2007). "The Architecture of Business Intelligence," [Online], [Retrieved December 5, 2010], http://www.accenture.com/NR/rdonlyres/15DCFF6A-4DE0-44D8-B778-630BE3A677A2/0/ArchBIAIMS.pdf.

Davenport, T. H. & Jarvenpaa, S. (2008). "Strategic Use of Analytics in Government," *IBM Center for the Business of Government*, Washington, District of Columbia.

Dayal, U., Castellanos, M., Simitsis, A. & Wilkinson, K. (2009). "Data Integration Flows for Business Intelligence," *Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology*, 1-11.

Gartner (2006). "Gartner Survey of 1,400 CIOs Shows Transformation of IT Organisation is Accelerating," [Online], [Retrieved November 27, 2010], http://www.gartner.com/presseleases/asset14367811.html.

Gartner (2007). "Gartner EXP Survey of More than 1,400 CIOs Shows CIOs must Create Leverage to Remain Relevant to the Business," [Online], [Retrieved November 27, 2010], http://www.gartner.scom/it/page.jsp?id=501189.

Gartner (2008). "Gartner EXP Worldwide Survey of 1,500 CIOs Shows 85 Percent of CIOs Expect "Significant Change" over Next Three Years," [Online], [Retrieved November 27, 2010], http://www.gartner.com/it/page.jsp?id=587309.

Gartner (2009a). "Gartner EXP Worldwide Survey of More than 1,500 CIOs Shows IT Spending to be Flat in 2009," [Online], [Retrieved November 27, 2010], http://www.gartner.com/it/page.jsp?id=855612.

Gartner (2009b). "Business Intelligence Ranked Top Technology Priority by CIOs for Fourth Year in a Row," [Online], [Retrieved November 27, 2010], http://www.gartner.com/it/page.jsp?id=888412.

Gartner (2011). "Gartner Forecasts Global Business Intelligence Market to Grow 9.7 Percent in 2011," [Online], [Retrieved February 27, 2011], http://www.gartner.com/it/page.jsp?id=1553215.

Giovinazzo, W. A. (2003). Internet-enabled Business Intelligence, *Prentice-Hall*, Upper Saddle River, New Jersey.

Haag, S., Cummings, M. & Philips, A. (2007). Management Information Systems for the Information Age, *McGraw-Hill*, Boston.

Han„ J. & Kamber, M. (2006)."Data Mining: Concepts and Techniques," *Morgan Kaufmann*, San Francisco, California.

Hobbs, L. (2007). Oracle Database 10g Data Warehousing, *Elsevier Digital Press*, Burlington, Massachusetts.

Hoffer, J. A., Prescott, M. B.& McFadden, F. R. (2007). Modern Database Management, *Pearson/Prentice Hall,* Upper Saddle River, New Jersey.

IBM. (2009). "The New Voice of the CIO: Insights from the Global Chief Information Officer Study," [Online], [Retrieved December 4, 2010], http://www-935.ibm.com/services/us/cio/ciostudy/pdf/ midsize.pdf.

Imhoff, C., Galemmo, N. & Geiger, J. G. (2003). Mastering Data Warehouse Design: Relational and Dimensional Techniques, *John Wiley & Sons*, Indianapolis, Indiana.

Inmon, W. H. (2005). Building the Data Warehouse, *Wiley*, Indianapolis, Indiana.

Kerdprasop, N. & Kerdprasop, K. (2007). "Moving Data Mining Tools Toward a Business Intelligence System," *International Journal of Intelligent Technology*, 2(2). 99-104.

Kimball, R. & Caserta, J. (2004). The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data, *Wiley, Indianapolis, Indiana*.

Kimball, R., Ross, M., Thornthwaite, W., Mundy, J. & Becker, B. (2008). The Data Warehouse Lifecycle Toolkit, Wiley, New York.

Li, Z., Huang, Y. & Wan, S. (2007). "Model Analysis of Data Integration of Enterprises and E-commerce Based on ODS," *International Conference on Research and Practical Issues of Enterprise Information Systems (CONFENIS 2007)*. 1, 275-282.

Ma, Z., Wang, C. & Wang, Z. (2011). "Semantic Model Based on Three-layered Metadata for Oil-gas Data Integration," *Advances in Information Sciences and Service Sciences*, 3(7). 216-224.

MarketResearch.com.(2010). "Business Intelligence Tools Market Forecasted to Grow $13 Billion by 2015," [Online], [Retrieved December 3, 2010], http://www.marketwire.com/press-release/Business-Intelligence-Tools-Market-Forecasted-to-Grow-13-Billion-by-2015-1162837.htm.

Melchert, F., Winter, R. & Klesse, M. (2004). "Aligning Process Automation and Business Intelligence to Support Corporate Performance Management," *Proceedings of the Tenth Americas Conference on Information Systems, 4053-4063.*

Negash, S. (2004). "Business Intelligence," *Communications of the Association for Information Systems*, 13, 177-195.

Nelson, G. (2008). Metadata 100 Success Secrets 100 Most Asked Questions on Meta Data How-to Management, Repositories, Software, Standards, Tools and Databases, *Emereo, London, UK.*

Pant, P. (2009). "Essential Components of a Successful BI Strategy," [Online], [Retrieved January 21, 2011], http://www.information-management.com/specialreports/2009_15

5/business_intelligence_bi-10015846-1.html?pg=1.

Parida, R. (2006). Principles and Implementation of Data Warehousing, *Firewall Media*, New Delhi, India.

Popovic, A., Coelho, P. S. & Jaklic, J. (2009). "The Impact of Business Intelligence System Maturity on Information Quality," *Information Research,* 14(4). 417.

Prevedello, L. M., Andriole, K. P., Hanson, R., Kelly, P. & Khorasani, R. (2010). "Business Intelligence Tools for Radiology: Creating a Prototype Model Using Open-source Tools," *Journal of Digital Imaging*, 23(2). 133-141.

Ranjan, J. (2009). "Business Intelligence: Concepts, Components, Techniques and Benefits", *Journal of Theoretical and Applied Information Technology*, 9(1). 60-70.

Rasmussen, N. H., Rasmussen, N., Chen, C. Y. & Bansal, M. (2009). Business Dashboards: A Visual Catalog for Design and Deployment, *John Wiley & Sons,* Hoboken, New Jersey.

Reinschmidt, J. & Francoise, L. (2000). Business Intelligence Certification Guide, *IBM*, San Jose, California.

Rob, P. & Coronel, C. (2009). Database Systems: Design, Implementation, and Management, *Course Technology*, Boston, Massachusetts.

Sen, A. & Sinha, A. P. (2005). "A Comparison of Data Warehousing Methodologies," *Communications of the ACM*, 48(3). 79-84.

Shariat, M. & Hightower Jr, R. (2007). 'Conceptualizing Business Intelligence Architecture,' *Marketing Management Journal*, 17(2). 40-46.

Sommer, D. (2008). "Gartner Says Worldwide Business Intelligence Spending to Grow 11 Percent in 2008," [Online], [Retrieved December 3, 2010], http://www.gartner.com/it/page.jsp?id=636310.

Strand, M., Wangler, B. & Olsson, M. (2003). 'Incorporating External Data into Data Warehouses: Characterizing and Categorizing Suppliers and Types of External Data,' Proceedings of the 9th Americas Conference on Information Systems, 2460-2468.

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

Walker, D. M. (2006). "Overview Architecture for Enterprise Data Warehouses," [Online], [Retrieved December 1, 2010], http://www.datamgmt.com/index.php?module=documents&JAS_Document Manager_op=downloadFile&JAS_File_id=29

Wang, H. & Ye, Z. (2010). "An ETL Services Framework Based on Metadata," *2nd International Workshop on Intelligent Systems and Applications*, 1-4.

Wang, J., Chen, T.-J. & Chiu, S.-H. (2005). "Literature Review on Data Warehouse Development," IACIS Pacific 2005 Conference Program, 987-994.

Watson, H. J. (2009). "Tutorial: Business Intelligence – Past, Present, and Future," *Communications of the Association for Information Systems*, 25, 487-510.