Data Warehouse as a Backbone for Business Intelligence: Issues and Challenges

Mutaz M. Al-Debei
Data Warehouse as a Backbone for Business Intelligence: Issues and Challenges

Mutaz M. Al-Debei
Department of Management Information Systems
The University of Jordan, Amman, Jordan
Tel: +962-6-5355000; Ext.24247; Fax:+962-6-5355522
E-mail: m.aldebei@ju.edu.jo

Abstract
The aim of this research is to identify and classify the main issues and challenges facing different business organizations when implementing Data Warehouse (DW) technologies. This is highly significant given the theoretical and practical implications and importance of such technologies. It is also important to highlight these challenges given the scarcity of research in this domain despite its value. To determine DW issues and challenges, a qualitative research methodology was followed. A semi-structured interview protocol was used with 17 DW project managers and seniors’ members. The gathered data were analyzed by utilizing a bottom-up content analysis technique where content is coded and classified thematically so as to let concepts emerge naturally. The study results show that DW issues and challenges are both managerial/organizational and technological. According to the results of the study, the main managerial/organizational DW challenges are: (1) management commitment and support; (2) project champion; (3) user involvement and participation; and (4) team skills and composition, whereas the technological challenges are: (1) the selection of DW architecture; (2) the creation of the enterprise schema; (3) data integration and scalability; (4) data quality; (5) the design of human-computer interfaces; (6) mining the DW; (7) security and privacy risks; and (8) networks and telecommunications.

Keywords: Data Warehouse, Data Warehousing, Business Intelligence, Data Mining, Challenges.

1. Introduction
The digital era has meant that the availability of appropriate information and knowledge have become critical to the success of the business. The next information revolution is about information content and its purpose (Drucker, 1999). However, organizations need to adapt in order to survive and succeed as their business domains, processes and technologies change in a world of increasing environmental complexity (Bhatt and Zaveri, 2002). Enhancing their performance and competitive positions by improving their ability to respond quickly to rapid environmental changes with high quality business decisions can be supported by exploiting technologies such as Data Warehousing and Business intelligence (BI) analytical tools (Park, 2006; Arnott and Pervan, 2008). Nevertheless, assuring business intelligence basically requires having relevant, reliable, accessible, accurate, timely, complete, coherent, and consistent quality information to the decision at hand (Vassiliadis et al., 2000; March and
Hence, business intelligence through decision making improvement is a major concern for business managers nowadays (Park, 2006).

No doubt, accessing large and vast amount of the data stored in business organizations operational systems has become progressively more time consuming and cumbersome (Ang and Teo, 2000). Hence, business organizations have embarked on data warehousing to overcome these problems through integrating heterogeneous operational data sources (Shin, 2002). Data Warehouse (DW) provides a technological infrastructure enabling business organizations to extract data from source systems, cleanse/scrub the extracted data, and transform enormous amounts of data to be stored in it (Wixom and Waston, 2001; Nemati et al., 2002); a process known as ETL (Extraction, Transformation, and Loading). DW is considered one of the most powerful decision support and business intelligence technologies that have emerged in the last decade (Ramamurthy et al., 2008). Nevertheless, the realization of DW benefits by business organizations has been below expectations (Stackowiak, 1997; Waston et al., 2001; Zimmer, 2004). Hence, this study is mainly focused on two points: first, it presents, illustrates, and discusses the role and value of DW as an aspect or driver for business intelligence, and secondly it is critically analyzing both organizational and technological issues and challenges of DW development/implementation with current approaches and technologies.

The remainder sections of this study are structured as follows. In the next section, we provide the research goals and objectives. Then, a theoretical background about the study area is provided by illustrating the main area concepts followed by discussing the main motives behind deploying DWs. Thereafter, the research methods are discussed. Then, the main challenges and issues of DWs development are discussed as the main results of the study. Finally, the study conclusions are presented.

2. Research Goals and Objectives
Data warehouse technologies and projects are highly significant to business organizations in terms of investment, time, and effort in addition to their perceived value. However, the failure rate of such projects is high. We therefore aim – in this study- to identify the main issues and challenges facing different business organizations throughout the implementation of data warehouse technologies. Achieving this goal is significant as the results can be used as a roadmap or guidelines for business organizations aiming to implement data warehouse technologies.

The objectives of this research are multifold. First, we aim to determine various managerial and organizational factors those deemed significant to the success (or failure) of data warehouse implementations. It is very important to highlight these factors as many decision makers within organizations believe that data warehouse projects are purely technical ones and thus managerial and organizational aspects and domains are irrelevant.

Given the technical complexity of such technologies and projects, there are also some technological challenges to be determined and this represents our second objective. Thirdly, we aim to provide a solid and clear theoretical background about data warehouse, data warehousing, and business intelligence concepts. Highlighting the major drivers for implementing data warehouse technologies and its relationship with business performance represents our fourth objective in this research.

3. Theoretical Background
3.1. Introduction to Data Warehouse, Data Warehousing, and Business Intelligence Concepts
One of the key developments in the Information System (IS) field is data warehousing (Waston et al., 2002). Unlike On-Line Transaction Processing (OLTP) databases, which are application-oriented, detailed, and operational (Ramamurthy et al., 2008), DW (see table 1) is “a subject-oriented, integrated, non-volatile, and time variant, non-updatable collection of data to support management decision-making processes and business intelligence.” (Inmon, 2002: p. 31). DWs are widely perceived
as valuable devices for acquiring information from multiple sources and delivering it to managers and analysts who may be using different software or computer platforms with special features and capabilities (Subramanian et al., 1997). DWs meant to support managers with answers to important business questions that require analytics such as pivoting, drill-downs, roll-ups, aggregations and data slicing and dicing (Ramamurthy et al., 2008). Moreover, all levels of management decision-making processes are supported by DW through the collection, integration, transformation, and interpretation of both internal and external data (Negash, 2004). Moreover, Alshawi et al. (2003) have elaborated how a DW could provide useful and valuable information and knowledge at a strategic, management control, Knowledge and operational levels.

Table 1: Characteristics of Data (adopted from Ang and Teo, 2000)

<table>
<thead>
<tr>
<th>Characteristics of data</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject-oriented</td>
<td>Data are grouped by subjects. For example, data on customers are grouped and stored as an interrelated set.</td>
</tr>
<tr>
<td>Integrated</td>
<td>Data are stored in a globally consistent format. This implies cleansing the data so that data have consistent naming conventions and physical attributes.</td>
</tr>
<tr>
<td>Time-Variant</td>
<td>Data captured are for long-term use often 5–10 years. So they are captured in a series of snapshots.</td>
</tr>
<tr>
<td>Non-volatile</td>
<td>Once data at a particular time, say t1, are captured and stored, their attributes are preserved.</td>
</tr>
</tbody>
</table>

Before digging more into the main study objectives, it’s crucial to differentiate a DW, the repository of summarized data (Ang and Teo, 2000), from data warehousing which revolve around the development, management, methods, and practices that defines how these summarized data are acquired, integrated, interpreted, managed, and used within business organizations (March and Hevner, 2007). On the other hand, business intelligence is rooted in interpreting the data acquired through environmental scanning with respect to a business task “contextualization” and it’s supposed to provide tactical and strategic information to decision-makers so as to be able to manage and coordinate operations and processes in their business organizations (Tseng and Chou, 2006). For the purpose of business intelligence, many analytical tools have been developed such as: excel, reporting tools, dashboards, OLAP, and data mining. Business intelligence revolves around knowledge discovery and inferences by analysing the data stored in DW to acquire valuable information; in other words, “DW is a repository of intelligence from which business intelligence can be derived” (March and Hevner, 2007).

Another important issue is the differentiation between Operational Data Store (ODS) and DW. Although ODS uses DW technology (i.e. star schema) to provide integrated view of data, but its intended to assist day-to-day operations and not decision making (Shin, 2002), since ODS is a subject-oriented, integrated, and volatile (updatable) data store that contain only business organization detailed data for operational usage (Boar, 1997).

Furthermore, the differentiation between the data stored in a DW and the metadata is important to both DW users and maintainers. Metadata (data about data) provides the following information about the DW (Gray and Waston, 1998):

1. A directory about the data stored in a DW (location, index, etc.).
2. A guide to mapping data from operational sources to DW form.

3.2. Reasons Underlying the Implementation of Data Warehouse Technologies

A wide variety of tangible and intangible benefits can be gained from DWs applications “see figure 1” (Waston et al., 2002). DW is a data repository which is relevant to the management of business organization and from which the needed information and knowledge to effectively manage the organization are emerged (Waston, 2001). Initially, Data warehousing was viewed as a way by which business organization could solve the problems associated to their independently legacy systems which
often contains inaccurate, duplicate, and dissimilar data about the same entity (Grant, 2003). DW technology (i.e. star schema, see figure 2) can help managers make more effective decisions (Griffin, 1998) by providing them with suitable information which is fundamentally different from the type of information that businesses use in their day-to-day operations (Summer and Ali, 1996).

**Figure 1:** The Benefits from data warehousing (adopted from Waston et al., 2002).

![Figure 1: The Benefits from data warehousing](image)

DWs contain cleaned, aggregated, consolidated large volumes of data that is accumulated in multidimensional data structure to support multidimensional analysis (Boehnlein and Ende, 1999). Not only a DW recognizes the need for current and future data, but also recognizes the need for historical data (Harmon, 1998); for instance, trend analysis requires a great deal of historical data regardless the company size (Griffin, 1998). DW allows a business organization to manipulate a great deal of data in ways that are useful to it, such as: cleansing, organizing, describing, summarizing and storing large volumes of data to be transformed, analyzed and reported (Griffin, 1998).

**Figure 2:** Material Inventory Star Schema (adopted from Chau et al., 2002).

![Figure 2: Material Inventory Star Schema](image)

DW offers effective service data management and data delivery processes by expanding stovepipe knowledge into cross-functional integrative business intelligence (Shin, 2002). Accordingly,
Business organizations could compete better by having the ability to learn from the past, to analyze current situations, and to predict the future scenarios (Boar, 1997).

**Figure 3:** A generic DW architecture (adopted from Ahmad et al., 2004)

**Figure 4:** A sample data cube showing the multidimensional database concepts (adopted from Rivest et al., 2005).

Arnott and Pervan (2005) argue that data warehousing provides the large scale IT infrastructure for contemporary decision support and business intelligence (see figure 3). They argue that the main reasons behind that is the use of Multi-Dimensional Data Model “MDDM” or cubes “see figure 2” (Kimball, 1996), which organizes large data sets in ways that are meaningful to managers besides
being relatively easy to query and analyze. It has been proved that MDDM is the most suitable for On-Line Analytical Processing “OLAP” applications, data mining and advanced reporting functions (Bellatreche, 2001). The conceptualized MDDM can be physically realized in two ways: first, by using trusted relational databases “star schema/snowflake schema” or, by using a specialized multidimensional databases (Chaudhuri, 1997). Many computer assisted analytical processes such as: data mining and OLAP (Ahmad et al., 2004) that used to analyze data from different angles and distilling it into actionable information run over DWs (Gunnarsson et al., 2007). Moreover, the category of DW project is IT infrastructure type that provides a foundation for IS/IT application development (i.e. ERP, CRM) which could provide business organizations with strategic competitive advantages (Swanson, 1994; Duncan, 1995). Another potential benefit of DW is that using a single data source (DW) may facilitate business process re-engineering at business organizations (Waston and Haley, 1998).

Summer and Ali (1996) argued that DW is the way in which business organization converts its data into information which can be represented into different ways (textual, graphically, etc.) based on its reporting capabilities. Also, they argued that ad hoc system which is provided to managers based on DW allow them to generate speculative information such as projections, and allow them to explore “what-if” analysis. Thus, the desire to improve decision-making and business performance has been the fundamental business driver behind data warehousing (Gray and Waston, 1998). Furthermore, DW is analytical processing (Ahmad et al., 2004) and its key role is to offer forceful business intelligence to business organizations’ decision-makers, through enriching their abilities in understanding business problems, exploiting opportunities and improving business performance (March and Hevner, 2007). Hence, by leveraging DW technology for business intelligence initiatives, business organization can gain strategic competitive advantage (Ramamurthy et al., 2008).

3.3. The Role of DWs in Enhancing Business Performance

In this section, the researcher presents a number of DW studies’ applied in different industries a long with their results as an empirical evidences of the role of DWs in enhancing business performance. One example is that of Berndt et al. (2007), who explored the role of data warehousing in surveillance systems; using Florida wildfire data sets from years 1998 to 2001, they presented the adroit application of DW OLAP query tools for identifying and exploring patterns of illness that might indicate the presence of biological or chemical agent in the environment. Ahmad et al. (2004) showed how the application of data warehousing technique could be used in developing DSS system to help in decision-making task, more particularly for use in selecting sites for residential housing. Park (2006) showed through a laboratory experiment that implementing DW to support DSS systems results in significantly better performance by providing more reliable and consistent data for business decision-making. And finally, Griffin (1998) argued that data warehousing can provide hospitality companies with strategic competitive advantage since it help managers make more effective decisions.

4. Research Methods

Despite the importance of data warehousing technologies in theory and practice, only few research (e.g. Watson and Haley, 1997; Wixom and Watson, 2001; and Hwang and Xu, 2007) have been conducted to study the data warehousing challenges and success factors. Therefore, in this research, we follow a qualitative paradigm so as to comprehensively and clearly identify the main issues and challenges facing business organizations when implementing data warehouse technologies. A semi-structured interview protocol was used with 17 data warehouse project managers and senior members with both technical and managerial backgrounds and positions in order to collect relevant data. The duration of each interview was about 45 minutes. The main themes discussed within the interviews were “the role of management in data warehouse implementations”, “the role of team members in data
The collected data were analyzed following a content analysis technique. Content analysis can be defined as any research technique for making inferences by systematically and objectively identifying shared common properties for the phenomenon under investigation (Stone et al., 1966). To end up with systematic and objective inferences, a data classification technique where data are read and categorized into concepts that are suggested by the data rather than imposed from outside is mandatory (Agar, 1980; Al-Debei and Avison, 2010).

To be aligned with the aforementioned principles, we analyzed the collected data thematically. In fact, a bottom-up open coding procedure was followed here where interviews have been broken into segments or incidents which are found in a phrase; a sentence or two (Charmaz, 2006). Each incident was coded by an indicator characterized as short and effective. The indicators of different incidents were compared to each other to develop indicators of a higher level (i.e. indexes). Thereafter, the codes or indexes were analyzed where those pointing towards the same theme were aggregated together. Aggregating the codes based on themes allowed 12 concepts to emerge. Then, these 12 concepts were categorized into two different classes: technological challenges including 8 concepts and managerial/organizational challenges including 4 concepts (see Table 2).

Table 2:  Issues and Challenges of Data Warehouse Implementations

<table>
<thead>
<tr>
<th>Managerial/Organizational Challenges</th>
<th>Technological Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Management Commitment and Support</td>
<td>• The Selection of Data Warehouse Architecture</td>
</tr>
<tr>
<td>• Project Champion</td>
<td>• The Creation of the Enterprise Schema</td>
</tr>
<tr>
<td>• User Involvement and Participation</td>
<td>• Data Integrity and Scalability</td>
</tr>
<tr>
<td>• Team Skills and Composition</td>
<td>• Data Quality</td>
</tr>
<tr>
<td></td>
<td>• Human-Computer Interfaces</td>
</tr>
<tr>
<td></td>
<td>• Mining the Data Warehouse</td>
</tr>
<tr>
<td></td>
<td>• Security and Privacy Risks</td>
</tr>
<tr>
<td></td>
<td>• Networks and Telecommunications</td>
</tr>
</tbody>
</table>

5. Results and Discussion: Issues and Challenges of DWs’ Developments/Implementations

Although many business organizations appear to have implemented DW, the road to its success has been plagued with failures (Hall, 2003; Kotler, 2003). In addition to the high cost and resource requirements needed for such project, many other societal/cultural, organizational and technical reasons may be responsible for their high failure rates (Ramamurthy et al., 2008). Implementing large-scale IS/IT projects such as DW is complex undertaking; a variety of many organizational and technological challenges and issues facing business organizations when they starting DW projects. This could be referred to the fact that DW project is an enterprise-wide scale initiative that involves acquisition and/or development of tools and applications for user access, database maintenance, and data transfer and scrubbing (Subramanian et al., 1997). On the other hand, lack of organizational readiness, project size and structure, incomplete risk management, familiarity with technology, and skills of system analysts can negatively impact the project’s outcome (Shin, 2002). In the following sub-sections, both managerial/organizational and technological issues and challenges have been discussed in detail.

5.1. Managerial/Organizational Issues and Challenges

5.1.1. Management Commitment & Support
Ramamurthy et al. (2008) argue that organizational commitment through senior management’s participation, involvement, and support is a key element for an innovation such as DW to be adopted
and subsequently used efficiently and effectively. Organizational commitment ensures the consistency between DW project and strategic vision and direction as well as shows the importance of this innovation to business value in order to sustain favourable attitude (Griffith et al., 2003). Moreover, if DW is backed by the organization’s management, the users are more likely to accept it (Wixom and Waston, 2001). On the other hand and since DW is a high multiyear investment, significant politically charged issues could be triggered; thus, senior management and other stakeholders’ attention, involvement, support, and commitment is highly needed (Wixom and Waston, 2001; Ramamurthy et al., 2008).

5.1.2. Project Champion
The planning process for data warehousing initiative falls very much into a broader strategic planning exercise for municipal Information Systems (Subramanian et al., 1997). Therefore, DW project needs proper executive sponsorship such as a champion at the top management level to fund the project toward its completion and to be properly maintained (Griffin, 1998). When a champion is strongly supportive, users are willing to acquire knowledge about DW technology continuously even after the introductory training (Chenoweth et al., 2006). Thus, champions exhibit “transformational leadership behaviour”; in other words, when they support the DW project, they posses the needed skills to overcome resistance that may arise within the organization while implementing the DW system (Wixom and Waston, 2001).

5.1.3. User Involvement & Participation
Generally speaking and in order to get the most from DW, users must endure continual, formal, and systematic training for such large-scale initiative; this will enable users to understand functions supported by DW much better as well as being more accountable for making DW produce higher quality information (Ang and Teo, 2000). User involvement is crucial especially when the requirements for a system are unclear (Wixom and Waston, 2001). Moreover, such training will be useful in eliminating or at least reducing users’ resistance. If users are not willing to use DW, then its necessary to change their attitudes by offering them a suitable motivation and/or giving them additional training (Chenoweth et al., 2006). However, Subramanian et al. (1997) argued that assessing the talents and mind-sets of DW users is a necessary preliminary step to determine the kind of training they need and their readiness for such computing practice. Hence, if the DW project to go forward, both management and users must be convinced of its value and its deliverables (Chenoweth et al., 2006).

5.1.4. Team Skills & Composition
The DW development team skills have a major impact on its outcomes (Wixom and Waston, 2001). Further, Ang and Teo (2000) argued that even though DW project already had a great deal of top management support, users and IS staff should be part of the DW project team and should be viewed as partners with common objective of creating a useful and reliable DW. They also argued that when users are part of the team, they will be more proactive and patient in helping to exploit such a technology as well as understanding its capabilities and limitation and being more able to provide valuable inputs into future enhancements. Thus, the adoption of DW could be facilitated by ensuring the existence of key users and stakeholders within the project team since they can creatively identify ways through which needed knowledge can be extracted from multiple functional areas within the business organization (Nambisan et al., 1999; Ramamurthy et al., 2008). In addition of getting a good combination of members in the DW project team, getting the right person to lead the DW project, who is technically competent, has adequate business knowledge and interpersonal skills is crucial to DW success (Ang and Teo, 2000).
5.2. Technological Issues and Challenges

Obviously, a DW development with its supporting middleware and analytical tools (i.e. OLAP, data mining, and dashboards) represents a large investment for business organization. On average, a DW installation cost is $1.5 million and in some cases exceeds $50 million (Goeke and Faley, 2007). Thus, a great deal of care must be taken to determine the nature of its ultimate development and use (Subramanian et al., 1997). In this section, a number of technological issues and challenges in building a DW are discussed.

5.2.1. The Selection of DW Architecture

There are two fundamental approaches to DWs: enterprise-level DWs (top-down approach) and departmental-level data marts (bottom-up approach). Unlike data marts which fulfil the information needs of a discrete group or a specific department, DWs are designed to satisfy the information needs of an entire business organization thus being more complicated and expensive to build and to maintain. Data marts main advantages are: it requires less time to be built, thus delivering value to customer more quickly. Further, it reduces the risk associated with such initiative since its design and development is simpler, and the amount of investment is lower than in DW. On the other hand, data marts within the same enterprise are often different, thus integrating them in an enterprise-wide DW can be a cumbersome and challenging process. To overcome this inconsistency problem, two approaches have been proposed by Gray and Waston (1998) as follows.

1. Following a step-by-step approach; an organization starts by a stand-alone data mart, but have a plan in place for integrating them later.
2. Create an enterprise-wide DW; then populate data marts from it and tune to local departmental needs.

Griffin (1998) elaborates more on the second previous solution. He argues that DW reduces redundancy of data storage among different departments as well as allowing users to perform a severe variety of analyses.

Table 3: DW Approaches

<table>
<thead>
<tr>
<th>DW Approach</th>
<th>Brief Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database Conversion</td>
<td>Taking all the data in the source systems and performing a one-time conversion to the integrated target system.</td>
</tr>
<tr>
<td>Database Synchronization</td>
<td>Data is extracted periodically from the operational source systems to update the target system (DW).</td>
</tr>
<tr>
<td>Federated Database</td>
<td>Only the data required by a query is integrated.</td>
</tr>
</tbody>
</table>

Srivastava and Chen (1999) have proposed three architectural approaches to data warehousing summarized in table 3.

5.2.2. The Creation of the Enterprise Schema

The step that follows the selection of the DW architecture is its logical design. Since DW represents an integration of multiple databases exist within a business organization, their existing schemas could be used as a foundation from which the new enterprise data model will be generated and designed. These existing schemas will be valuable in determining the main entities and their relationships although designers need to refine them and to add any missing elements (Srivastava and Chen, 1999). Taking into account Sheth and Larson (1989), Srivastava and Chen (1999), and March and Hevner (2007), the following challenges and issues regarding the enterprise schema creation have been identified as follows.

1. **Structural heterogeneity:** given that the enterprise data model relies heavily on the overlapping existing different database models, the structural heterogeneity arise. To give only one example, the same entity which represents the same real-world class may exist in multiple database schemas with the same fields, but the fields’ size or data type may differs.
2. **Semantic heterogeneity**: Semantic heterogeneity exists when data is defined differently in different source schemas, thus pose enormous challenges. This includes synonym and homonym problems. The first represents having two different names in different database schemas for the same real-world class, and the later occurs when having the same name for different real-world classes in different database schemas.

3. **Constraints mismatches**: generally speaking, there is no right approach for resolving mismatch incompatibility; such as having manager.salary $\geq$ 1000 in one database, and having manager.salary $\geq$1500 in another database.

5.2.3. **Data Integration & Scalability**

One of the significant technological challenges to DWs development is data integration from heterogeneous sources. The wide variety of critical information requirements for business decision-making and their different sources; both internal {spreadsheets, documents, plain files, etc.} and external {governmental documents, Journals, etc.} (Waston et al., 2001) represents a great challenge to DW designers. However, the internal information is useful for tactical level decision making and evaluation, whilst the external one is useful for strategic decision making and evaluation (Drucker, 1995). The incorporation of unstructured and semi-structured information about partners, policies, rules, competitors, etc. through an environmental scanning process influencing the DW usefulness for decision making purposes (March and Hevner, 2007). On the other hand, the integration of past, present and future information is also challenging. Most often, data has to be extracted from a wide variety of distributed operational information systems that operate on different hardware platform and use different Database Management Systems (DBMS) with different structure (March and Hevner, 2007). Thus, data integration through the need to handle such heterogeneous sources of data leads to considerable complexity (Greenfield, 2003). To give just two examples, determining whether a pair of records coming from two different databases represent the same real-world entity or not is a great challenge, another example is when two records coming from two databases having the same entity identification (primary key), but the other fields’ values are different (Srivastava and Chen, 1999). Hence, ETL function represented the most expensive and time-consuming portion in the DW development (March and Hevner, 2007).

Generally speaking, the common DW development method which relies on capturing and acquiring a large volume of transactional data and loading it into a DW is not sufficient (Wetherbe, 1991) since not all information would be stored in the DW, not all departments would desire the same information with the same physical format (Subramanian et al., 1997), and not all information would be suitable for decision-making purposes. Hence, data may remain relatively useless for business intelligence and decision-making purposes even when it is efficiently acquired and stored (Grant, 2003). For these reasons, performing information system requirements analysis could be useful in determining the types of information required for effective management of business operations, control, strategic planning (Subramanian et al., 1997), and decision-making. Furthermore, March and Hevner (2007) argued that one of the significant challenges for the data warehousing research community nowadays is that DW must be understandable and adaptable, and must include experience-based organizational knowledge in order to provide information that improves business managers’ ability in understanding and identifying situations requiring actions as well as enabling them to understand decisions’ implications and their impacts over time.

Moreover, DWs stores non-volatile information over time, thus scalability issue should be taken into consideration. DW must be designed for change from the beginning (March and Hevner, 2007) since it has a periodic updates rate of few gigabytes (Hernandez and Stolfo, 1995).

5.2.4. **Data Quality**

Another major concern for DWs is data quality. Further, it has a rigorous impact on the overall business organization’s performance. Thus, relying on inaccurate, incomplete, inconsistent, imprecise, irrelevant, and non-coherent information for decision making purposes in general is a disaster, and it is
more harmful for strategic level of decisions. Loss of information, insufficient information (ambiguity), meaningless data, and incorrect data have been identified by Wand and Wang (1996) as the most observed data problems. Interestingly, they also indicate that there is a proportional relationship between the design & production techniques involved in generating the data and the quality of resulted data. To give just one example, approaches for cleansing data during ETL have been shown to be successful (Berndt et al., 2003; March and Hevner, 2007).

Table 4: Notable data quality dimension (adopted from Wand and Wang, 1996)

<table>
<thead>
<tr>
<th>Dimension</th>
<th>#cited</th>
<th>Dimension</th>
<th>#cited</th>
<th>Dimension</th>
<th>#cited</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>25</td>
<td>Format</td>
<td>4</td>
<td>Comparability</td>
<td>2</td>
</tr>
<tr>
<td>Reliability</td>
<td>22</td>
<td>Interoperability</td>
<td>4</td>
<td>Conciseness</td>
<td>2</td>
</tr>
<tr>
<td>Timeliness</td>
<td>19</td>
<td>Content</td>
<td>3</td>
<td>Freedom from bias</td>
<td>2</td>
</tr>
<tr>
<td>Relevance</td>
<td>16</td>
<td>Efficiency</td>
<td>3</td>
<td>Informativeness</td>
<td>2</td>
</tr>
<tr>
<td>Completeness</td>
<td>15</td>
<td>Importance</td>
<td>3</td>
<td>Level of detail</td>
<td>2</td>
</tr>
<tr>
<td>Currency</td>
<td>9</td>
<td>Sufficiency</td>
<td>3</td>
<td>Quantitativeness</td>
<td>2</td>
</tr>
<tr>
<td>Consistency</td>
<td>8</td>
<td>Usableness</td>
<td>3</td>
<td>Scope</td>
<td>2</td>
</tr>
<tr>
<td>Flexibility</td>
<td>5</td>
<td>Usefulness</td>
<td>3</td>
<td>Understandability</td>
<td>2</td>
</tr>
<tr>
<td>Precision</td>
<td>5</td>
<td>Clarity</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Although disagreement in the literature on data quality dimensions (see table 1), assuring better quality requires thorough understanding of its meaning and dimensions. However, more recent researchers have stated accuracy, timeliness, completeness, and consistency as quality dimensions (i.e. March and Hevner, 2007). Nevertheless, the concept of data quality is relative and the quality of data is of high subjective nature (Vassiliadis et al., 200) which makes data quality assurance more challenging. Thus, more attention to data quality issues in DWs is needed (March and Hevner, 2007).

5.2.5. Human-Computer Interfaces

One of the paramount important issues and the primary determinant of DW success from the end-user perspective is the human-computer interface (March and Hevner, 2007). Providing more user-friendly interface tools such as pull-down menus, menu-based interfaces and drags and drop capabilities for analysis and reporting purposes consider one of the main competitive advantages for DW vendors. Payton and Zahay (2003) claims that one of the major reasons for not using DW is that users perceived it as more technically-oriented reporting tools, thus depending on ‘power-users’ to get an answer to a query. However DW supports ad-hoc queries, thus DW designers should trade-off between the ability to express such queries and its ease-of-use (March and Hevner, 2007). Finally, For DWs to be successful and to enhance business intelligence mainly by supporting organizational decision-making, users must be shielded from its underlying complexity by a user-friendly human-computer interface.

5.2.6. Mining the DW

In order to perform strategic analysis of warehouse data, a number of tools have been developed such as: data mining, OLAP, trend analysis, forecasting and simulation (Srivastava and Chen, 1999). These tools are useful to extract a valuable hidden knowledge from large amount of data through slicing, dicing, rolling-up, and drilling down capabilities. To give just one example, data mining, or Knowledge Discovery from Databases (KDD), techniques could be applied in Customer Relationship Management (CRM) to predict customer profitability (Shen et al., 2007), to conduct click stream analysis (Lee et al., 2001), and for customer segmentation purposes (Ha, 2007). However once a problem is identified, the available commercial tools are not very effective for generating solution alternatives nor in discovering useful knowledge for strategy formulation and implementation (March and Hevner, 2007). On the other hand, these tools are still considered complex and somehow sophisticated to use from the end-user perspective.
5.2.7. DW Security & Privacy Risks
Most often, all enterprise data will be stored in its DW. Data in a warehouse is orderly, integrated, centrally-located and easily accessible thus offers an appealing target (Harmon, 1998). However, the data in a warehouse shouldn’t be threatened, lost, and/or manipulated. Moreover, the information and the knowledge extracted from a DW should be reliable; otherwise it can have disastrous effects on the enterprise (Finne, 1997). However, the quality of knowledge extracted from a DW is totally dependent on the quality of data stored in it. Hence, many techniques such as back-up, disaster recovery plans, strong password policy, Intrusion Detection Systems (IDSs), firewalls, encryption and anti-virus SW that reduce the risks of losing or manipulating the enterprise data should be implemented to mitigate the risks of losing this valuable asset. Furthermore, end-users’ awareness of these risks and security procedures is highly beneficial.

On the other hand, current privacy approaches such generalization, condensation, randomization, cloaking, etc. are all special cases to protect data by folding actual data values into patterns which in fact comes at the cost of some imprecision (Pedersen, 2007). Thus, more general significant approach is required.

5.2.8. Networks & Telecommunications
Since a DW serves most, if not all, departments in an enterprise, the ability to access it efficiently and effectively is a major concern. Sometimes, DW users are located in different regional areas and proper network connectivity should be available to allow them to access the DW effectively and reliably. Moreover, network bandwidth and its selection criteria is major concern when DW users are distributed among different sites (Jones, 1998).

6. Conclusions
Creating a DW has itself proved to be difficult and problematic (Grant, 2003) and it is highly perceived as high-risk/high-return initiatives (Waston et al., 2002). The researcher analysis reveals that despite agreement in the information system literature on the importance of data warehousing to an organization success through enhancing its decision-making quality, the attainment of such business intelligence based on it is still poorly proved empirically. Moreover and although the wide variety of motivators mentioned in the IS literature for developing DWs, overcoming the problems in the legacy heterogeneous systems as well as obeying to the governmental regulations are the most two conceivable reasons for which organizations are building their DWs.

DW is highly recognized as an infrastructure; many applications can run over it such as CRM and DSS systems. On the other hand, many techniques, such as data mining, OLAP and dashboards have been rising to prominence to extract business intelligence from DWs. Furthermore, DWs meant to be used by managers since they support decision-making process. Nevertheless, these techniques are still not very effective and are highly perceived as technically oriented by the end-users.

Nevertheless, DW has experienced relatively high failure rates and its spread and/or use has been to some extent limited (Quaddus and Intrapairot, 2001; Kotler, 2003). Perhaps due to the facts that designing and developing a DW is a risky, costly and complex process. It requires a huge amount of money as an investment, spans over years, and needs a wide variety of technical and managerial skills. Generally speaking, social aspects are shaping the technology. Hence, the interaction of technology and social context is the key determinant of DW success (Chenoweth et al., 2006). Nevertheless, despite the technical complexity of DW design and implementation, social/cultural and organizational factors are the most cited reasons behind DW failures.

From these insights and conclusions, each and every societal/cultural, organizational and/or technological issue and challenge mentioned previously represents a valuable area for research purposes. Interestingly, it is recommended to focus on areas such as improving business intelligence techniques in terms of user-friendliness and effectiveness in further researches. Other interesting and
challenging area for research purposes is the integration of structured, semi-structured and even unstructured data in DWs.

References


