Developing a Data Warehouse: Some Guidelines and Suggestions

Mary B. Curtis, (E-mail: curtism@unt.edu), University of North Texas
Kailash Joshi, (E-mail: joshi@umsl.edu), University of Missouri – St. Louis

Abstract

Although accounting information systems (AIS) are an important source of information for strategic development, the systems possess overriding limitations. Data Warehousing (DW) has emerged as an important tool to overcome these limitations by making use of existing data storage and database management technologies to organize the storage of relevant corporate data spanning several sources and time periods into one central location. Because DW projects are risky and expensive undertakings, however, care must be taken in their planning and development. This paper includes recommendations for the successful implementation of a DW project, addressing effectiveness, efficiency, and security issues.

Introduction

Data Warehousing (DW) is the combination of detailed and summarized historical data into an integrated database. This database provides an efficient means of managing corporate knowledge (Amidon and Skyrme 1998), which can be analyzed for strategic decision-making and mined for knowledge discovery (Fayyad et al. 1996). DW is one of the most recent information technology (IT) innovations, representing a highly significant conceptual development in improving access to information for decision-makers. DW makes creative use of existing data storage and database management technologies to organize the storage of relevant corporate data spanning several sources and time periods into one central location. DW, along with newer technologies such as tools for on-line analytical processing (OLAP) and data mining, also permits an organization to obtain the full benefits of pertinent historical data, organized for efficient management support. DW facilitates the creation of an effective,

user friendly environment to help executives identify new opportunities and make tactical and strategic choices based on facts, both internal and external. A recent survey conducted by The META Group (a consulting firm in Stamford, CT) reports that 95% of corporations it surveyed have plans for building DWs -- up from only 10 to 15% a couple of years ago.

Today's accounting information systems (AIS) are designed primarily for the support of operations and the preparation of current-year financial statements. Although AIS are an important source of information for strategic development, the systems possess two overriding limitations: data is retained for only the short-term, and the data that is retained is only quantitative and internal. Strategic decision-makers take a more long-term perspective requiring various types of data covering greater periods of time. In traditional transaction processing systems, detailed data is typically archived offline after the current year's processing. Consequently, strategic decision-makers are limited to previously printed summary

Readers with comments or questions are encouraged to contact the authors via e-mail.

reports, aggregated on preplanned attribute(s), such as by month, for their information source. Such reports severely curtail users' flexibility to aggregate the data by other attributes or to conduct any additional analysis. To perform further analysis, users of historical data may be forced to manually re-enter data from the printed reports into PC tools such as spreadsheets. Such analyses are limited by the aggregation choices made in those historical reports. Additionally, much of the information relevant to strategic decision-making is qualitative and/or created externally.

An alternative is to keep the data online in AIS databases. However, several factors inhibited users from accessing data stored in such systems for their decision-making needs. First, the data is often dispersed in several disjoint databases that can not be easily integrated. Second, the definitions and formats of data across the databases are often inconsistent, making it difficult to integrate and interpret the data. Third, high transaction volume and fast response time requirements in transaction processing databases does not permit users the luxury of data analysis and exploration which would tie up the system resources and downgrade performance. Summarization and aggregation of data again and again for different users or even the same users would result in very poor response time and consume processing capacity, which would further limit decision analysis as well as operational support. Finally, AIS generally do not provide support for external data or qualitative data that may be of interest to management in their analysis and is essential to the maintenance of corporate knowledge. All of these factors combined suggest that users requiring decision support and historical analysis are badly served by decision-support systems based on transaction-processing databases.

The data warehouse provides a natural solution to many of the problems discussed above. Development of an analysis-oriented DW results in organizations maintaining two different data storage facilities: a traditional AIS for transaction processing and a separate database containing historical data from diverse internal sources and some

external sources, with online historical data for query and analysis by users. This separation of informational sources permits the architecture of each facility -- AIS and DW -- the flexibility to focus on its mission and develop its own specialized data storage and access structures.

A distinct advantage of the DW is its ability to combine data from multiple sources into an integrated, logical data-set. This is of particular importance to multinational firms whose various locations operate different accounting systems with little consistency of contents and format. Along with this integration of corporate information, DW facilitate the complex decision support requirements of today's competitive business environments, such as environmental scanning, strategy formulation, and tactical planning and control (Nelson 1998). DWs also promote the use of exploratory data mining tools to help uncover patterns in large data sets, enhance organizational knowledge and lead to better decision making. Thus, a well-designed DW forms the backbone of contemporary decision support and executive information systems.

Since accountants and auditors are integral to the development and control of DW, it is incumbent upon accounting professionals to gain an understanding of how and why DW are developed and maintained. In this paper we analyze and discuss some issues in the design of DWs with the objective of providing information that may be useful to developers, auditors, and users of data warehouses. We present a brief overview of DW concepts in the next section. The following section discusses some key issues in DW design. The final section presents a summary and conclusions.

Data Warehouse Architecture

DW has been defined as a subject-oriented, integrated, time variant, non-volatile collection of data in support of management's decision making processes (Inman and Kelley 1994). Providing a single window of the "business reality" to management (Hackathorn 1995) facilitates access and analysis of data to support decision

making and strategy formulation (Kimball 1996).

Subject orientation refers to organization and summarization of data pertaining to specific business activities. For example, a chemical company may wish to focus on sales data, a manufacturer may focus on manufacturing data relating to work-in-process, rejections, and manufacturing cycle times, while an insurer may focus on claims submitted and processed. Integration refers to structured procedures for the collection of data from diverse sources after developing common data formats, definitions, validation, and integrity rules. The sources of data may include separate internal AIS databases on different platforms and locations, as well as external databases. Such integration of information from varied origins permits different users to adopt a unified, common view of data to support their analyses.

Time variant refers to data that spans a number of years, as opposed to the current-year information maintained by AIS. A time dimension permits aggregation of data over time spans ranging from daily to monthly and yearly aggregates, supporting time-series analysis. Given the historical nature of data, there are generally no changes or updates to the stored data, although additional data is regularly added to the DW. Thus, the data is basically non-volatile in a DW as opposed to transaction system databases where the current status is constantly changing.

Data stored in a DW is typically extracted from the various AIS through specially designed procedures, which help populate the warehouse and make periodic additions to keep it current. These extraction programs can filter and scrub data, transform the data to standardized formats, integrate from different tables, classify according to given class membership rules, and aggregate or summarize. Figure 1 provides an overview of this process.

The potential of a DW for user analysis and decision support can not be fully exploited, however, without the development of On-Line Analytical Processing (OLAP) capabilities to pro-

vide an environment for decision support analysis, business modeling, and operational research support (Rob and Coronel 1997). OLAP systems provide easy to use interfaces (GUI) and support multidimensional data analysis techniques. Multidimensional analysis permits data to be analyzed and cross-tabulated along different dimensions. For example, sales transactions can be analyzed along the different dimensions of time, product, sales person, customer, and location. A three dimensional analysis may examine sales by month by product class for each customer.

Relational databases have emerged as the familiar, cost effective technology for managing a DW's data store (Francett 1994). Several vendors provide relational databases for DW application, among which Oracle has captured a large share of the market.

Issues in Data Warehouse Design

Given the specific decision support focus and the need to store and process large volumes of historical data, design requirements of a DW are different from a typical AIS database (for an overview see Table 1). In transaction processing systems, data is generally stored in structures that are optimized for frequent and fast updates, inserts, and deletes. However, a data warehouse represents a different environment where there are almost no updates or deletes. Additional data is added only in a batch mode on a periodic basis.

The primary requirement of a DW is to optimize frequent reading, selection, and aggregation of data along different dimensions for different user analytical requirements. One way to support such user requirements in a timely manner is to also store pre-summarized data that has been aggregated on different attributes for each dimension. For example, sales data can be aggregated by store, city, state, and region for each product for each day, week, or month. An alternative to designing DW table structures to support data aggregation is the use of star schema in the design of warehouse data storage structures. Figure 2 presents an example of the star schema for sales data.

Table 1
Salient Differences Between Transaction Based Systems and Warehouse Based Systems

Transaction Based Systems	Data Warehouse Based Systems
On-line, real time update	Batch update
Detailed data	Summarized and integrated data
Oriented for high transaction throughput	Relaxed performance requirements
Relatively small amounts of data processed	Large amount of data being processed
Day-to-day operations	Decision support and management use
Redundancy is avoided	Redundancy and summarization
May have been developed without modeling	Modeling is essential
Most current data values	Values selected are time dependent
Generate activity for downstream systems	Do not generate data for other applications
Processing requirements understood before development	Processing requirements remain ad-hoc and flexible, keep evolving after development

Adopted from: Paul Arrowwood, 1997

The sales fact table is at the center of the star, while tables for dimensions such as time, location, product, and person are on the spokes. Each dimension contains additional qualifying attributes for analysis, such as the time dimension's day, month, quarter, and year fields. The data stored in a fact table can be aggregated along different attributes of each dimension and stored in summarized fact tables for faster response to user queries. Figure 3 shows the layout of summarized fact tables for sales data along the different attributes (e.g., region, state, city, and store location) of the location dimension.

Alignment of the Data Warehouse Project to Business Needs

Building a DW represents a significant investment of resources and effort. It also presents significant business risk if long-term strategic decisions are made based on inappropriate information. Therefore, before undertaking such a massive effort, it is necessary to clearly define the scope and goals of the overall project, as well as the business needs the data warehouse would be expected to meet. Since a DW is designed to provide information relevant to performance and trends within the organization, it is important to identify the overall strategy and critical success factors for the busi-

ness. This focus will help identify subject areas that should be given priority in the development of a DW. Such prioritizing is necessary for data warehousing projects in view of the large, multi-year efforts involved in such projects. Further, while the goal of storing all information in the DW appears promising, this may be economically infeasible and unwise for most organizations (except probably the largest corporations), given the investments needed. Therefore, the issue of what to include and what to exclude should be carefully considered. Such analysis should form the basis for identifying the scope of the project and subject areas to be included in the warehouse.

Planning for the Data Warehouse Project

A successful DW project requires architectural-level design (Lambert 1997). Thus, it may be useful to learn from the experience of other organizations and DW developers in planning for a DW project. Many organizations have spent liberally in an attempt to build massive DW projects only to find that they do not meet requirements or, in worst cases, do not work at all (Bird 1996). To guard against failure, developers should make a realistic assessment and attempt to limited the scope of the initial project. To build a scalable solution, some developers recommend a series of six

Figure 2 : An Example of Star Schema for Sales Data
(Daily sales in the fact table and time, location, product and salesperson attributes in dimension tables)

Adopted from: Rob and Coronel, 1997

month projects, rather than a large one-shot project (Robinson 1996). Alternatively, the solution may be to start with a data mart, a DW that focuses only on one or more strategic subject areas or divisions. In either case, however, the effort should be a part of an overall design to develop the potential full-scope DW. In the absence of an overall architecture, integrating the different pieces developed separately may be too expensive and complex (DePompa 1996). Additionally, an overall design approach should include development of business process models in addition to data models. This will allow the designers to gain a better understanding of the business and ensure that DW structure conforms to that model.

Given the enormous resources and commitment that development of a DW requires, the involvement of senior user executives and top management is usually essential. It is wise to find a key corporate sponsor among strategic users, who's participation and support may be critical in initially educating senior management and maintaining enthusiasm for the project as resource demands increase and target dates slip. In addition to providing resources for the project, involvement of management at all levels is also helpful in the analysis of scope and subject areas to insure that the information needs of the highest level strategic decision-makers are met.

A disadvantage of this diverse and high-level involvement by management is that DW projects frequently suffer from project creep. As more and more users and managers develop an understanding of the DW approach, they begin to scale up their demands and expectations. Developers should guard against such project creep and keep the scope of each sub-project limited, to ensure timely phased deliveries.

Span and Detail of Data

Another choice confronting designers is the extent of detailed data that should be stored in the DW. While large volumes of detailed data can reduce the performance of the DW, the lack of detailed data could limit users' ability to perform

drill down analysis. The volume of data would also be determined by the span, i.e., the number of years of historical data kept in the warehouse.

Both span and degree of detail require careful analysis and consideration. Span should be determined through interaction with users and identification of specific business needs. Given the rapidly changing business environment, the utility of very old data should be questioned and justified. For example, one credit card company keeps only thirteen months of data in their warehouse. Once the span is determined, arrangements should be made for regular downflow and archiving of data out of the warehouse, as well as the inflow of new data from operational databases and its periodic summarization (Hackathorn 1995b).

The degree of detail poses another significant dilemma. Most DW projects aim to keep a varying degree of summarization along different dimension and attributes. As previously discussed, star schema provides a useful method for organizing these different degrees of summarization. Such summarization can significantly aid the performance for queries along planned paths. For example, an insurer summarizes claim data in thirty different ways. However, one important aim of data warehousing projects is to support ad-hoc querying and analysis. Such flexibility to analyze along unplanned dimensions requires access to detailed data. Detailed data is also needed to support drill downs by users from the summarized data. To support this requirement, analysts must identify the appropriate level of detail based on performance issues and user needs. For example, in a grocery store, daily summaries of sales transactions by product, store, promotion, payment method, and other transaction details may be adequate, rather than storing complete detail records.

DW administrators need to constantly monitor data volume and usage characteristics to refine the structure of the DW. Stored summarizations of the data along different dimensions may be created to match users' frequent ad-hoc summarization requests. A key resource to support monitoring of the DW and to manage its structures

Adapted from: Rob and Coronel, 1997

Figure 3: Multiple Fact Tables and Dimension Tables for Drill Down Support Along Store Locations
(Summarization of daily sales data at different levels of aggregation along the location dimension)

is the establishment of a metadata encyclopedia and directory.

Metadata

One of the key success factors in a DW design is the management and organization of metadata. Metadata is important for users as well as for systems professionals who are responsible for the development, maintenance, and refinement of DWs. Metadata typically provides a well-organized and accessible directory of all data available in the database, including details about formats, sources, definitions, and table layouts. Such details are vital for users to manage their ad-hoc querying and analysis, which is an important goal of data warehousing in terms of putting power in the hands of users. For developers, metadata is critical for accomplishing the goals of managing and organizing the data warehouse. For this purpose, Metadata should also include the organization's business models, business rules, data models, data definitions and formats, data usage model, report dictionary, data element descriptions, data conversion rules, and table layouts.

Metadata should be time stamped to allow changes in data definition over time. It is expected that when historical data spanning a number of years is being stored, there may be changes in metadata such as in data definitions, formats, business rules, and other data related specifications. Therefore, it is important to provide for a facility that would interpret the historical data with the appropriate formats or business rules. Finally, Metadata can help in managing refinements and upgrades to a DW by storing information about usage and access patterns, data volume and analysis patterns. Well-designed metadata will aid a project team's response to needed changes in DW structure or applications.

Knowledge Management

While operational databases provide support for day-to-day management, the focus of a DW is on strategic and longer term issues, such as knowledge management and assessing the status of

a company's operations. These longer-term objectives require access to external information. For example, in the absence of industry data, it is impossible to assess market share and evaluate trends.

Perhaps an insurance executive learns that application or claim processing lead time at her company is 18 days. Without industry performance and benchmarking data, she will be unable to evaluate the relative timeliness of her company's performance. Such data is frequently available from external sources such as marketing research organizations, industry associations, published sources, external databases, and firms specializing in disseminating benchmarking information.

Many DW projects focus only on quantitative data. However, management may need support in environmental scanning for identifying industry trends, monitoring competitors' activities, maintaining an historical record of corporate actions, and locating new opportunities (Madnani 1998). Frequently, information pertinent to such support is found in external qualitative reports. Therefore, locating and organizing such data would be very helpful for decision-makers. Ready access to qualitative information can be particularly helpful for management in formulating new strategies and in taking corrective actions. A DW provides a natural location and context for organizing such information. In one chemical company, four analysts are assigned to constantly scan the information sources (including the Internet and information services such as LEXIS/NEXIS) to identify any qualitative or quantitative information pertinent to agricultural chemicals for inclusion in the data warehouse. There may also be a need for storing images, sound, pictures, and full motion video data.

Internal sources of qualitative data are particularly important for managing corporate knowledge. Often, different units or individuals in an organization perform analyses and make reports or notes of their findings and conclusions. Frequently, dissemination of such information is limited to only one department or few executives. Organizing such reports in a DW would make them universally available to all decision-makers. For

example, results of a marketing research effort for a product may not be available to the R&D function. Even if a copy of such a report is sent to the R&D department, it may not be known to the relevant staff in that department. Further, access and awareness of the report is likely to diminish as time passes. However, if the report is available in the DW and properly indexed for search, all interested users can access it. This discussion brings us to the next important issue: Indexing.

Performance Issues and Indexing

Data warehousing projects can become ambitious in terms of scope and volume of data. Data storage requirements can exceed many tera bytes. In such environments, performance of access becomes an important issue. There are several solutions to this problem (Richards 1997). One solution is to install bigger and faster processing hardware and storage facilities. Parallel processors, specialized multidimensional database servers, and distributed data warehouses have been suggested as possible solutions. However, a hardware approach represents a very costly alternative that requires new technologies and skills and, in many instances, may not solve the problem. When alternatives exist, a hardware approach should be the last option, since such solutions can only alleviate the problem temporarily, until the volume of data again outgrows capacity.

Another set of solutions is software based. One alternative is the use of tools such as OLAP (on-line analytical processing), where the queries are tuned for multidimensional analysis along preplanned dimensions. Another approach is partitioning, where the data is segmented into logical areas. Such partitioning can reduce search efforts. An unpopular and unlikely option is to regulate the kind of queries that users can put to the DW through query trapping. Unfortunately, these software solutions may not solve performance and access problems in ad-hoc querying.

An attractive solution to data-access performance problem is to build inverted indexes in addition to the traditional B-tree indexes found in

relational databases. Well-designed indexes can reduce input/output operations and speed up the retrieval and join of required data. While a B-tree index is good for traditional keyed access, it is not very effective when a user is looking for trends or generalities. Inverted indexes can provide fast, ad-hoc access for previously undefined queries. Inverted indexes contain pointers to the database as data. If a table contains employee data with records 2, 13, 25, 36, 59 as Managers, an inverted index would hold Manager as data with pointers to records 2, 13, 25, 36, 59. An inverted index can look up Smith in a name field, irrespective of how it is stored in the string: "A.K. Smith", "Smith, John, Jr.", or "Smith Watson." Thus inverted indexes can be effective for even unstructured and qualitative data. Overall, a well designed set of indexes provides an efficient solution to the problems of data access in the DW environment without requiring special hardware or major restructuring of the DW model.

Design of User Interface and Access

There are three issues related to user access to data: design of the user interface, training those who will use the DW, and security over the data access methodologies. First, to facilitate utilization of the DW resource, design of a friendly and easy-to-use interface is critical. Several approaches are feasible for this purpose. Approaches to user interface can be broadly classified as those based on pre-planned analysis to provide standardized access and analysis facilities for identified requirements, and those composed of exploratory facilities that permit power users to explore and analyze available data on an ad-hoc basis without restrictions. Because users' abilities and needs may vary widely and individual user requirements may move from one to another category in time, data warehouse designers should provide users the greatest flexibility in the choice of access methods and strategies.

Tools are available for designing user interface and providing users with analytical capabilities along planned dimensions, such as OLAP tools. Some of the available OLAP packages in-

clude Trinzic Corporation's Forest and Trees, Business Object's Mercury, Cognos, Micro Strategy Inc.'s DSS Agent, Oracle Express, Oracle Discoverer, Oracle Analyzer, Intersolver's Data Direct Explorer, Axyss Information Advantage, and Expressway Technologies's Expressway. Programs can also be written to provide customized reports on demand at a regular frequency.

Other tools are available which permit developer customization by defining the structure of the DW, allowing users to conveniently perform their own exploratory analysis. When the data is stored in relational databases, users can also use SQL to directly retrieve the needed data for review and further analysis. A more advanced development involves data visualization techniques, which will make it easy for users to analyze and spot patterns in data. The newest access tools provide interfaces to corporate DWs based on web browsers such as Netscape or Microsoft's Explorer. The use of intranets, Internet, and web browsers to provide access to their DWs to dispersed users and customers (Bird 1996; Meyer and Cannon 1998) will revolutionize access to data while also simplifying it.

Regardless of the types of user access provided, user training is essential. IS management should pay serious attention to training users on data access and analysis tools. Users should also be trained in the layout and structure of the current DW as well as in the use of metadata resources. Users should be encouraged to learn SQL, spreadsheets, and other statistical data analysis packages to help them meet their specific data needs through the data warehouse. Such training programs should be regularly and repeatedly conducted at time slots convenient to users. Organizations that provide visibility and rewards such as presentation of letters of recognition, certificates, cash awards, salary raises, and mention in the in-house journal would find it easier to motivate users to learn the new tools. A high degree of interaction between IS and users in training programs can lay the foundation for obtaining regular user feedback and suggestions for future refinements.

Unfortunately, with increased access comes increased risk of unauthorized access. Security within the application can be strengthened through the use of defined user groups with unique access privileges and defined views for each user (Curtis and Joshi 1997). Since DW contain a valuable store of information about all aspects of the organization's business – information which could be even more valuable to competitors, security external to the application can be particularly important when access is provided over the internet. However, typical Internet security such as firewalls and system-based security, supplemented by encryption, will generally provide the needed level of security from external threat (Inmon 1997).

Integrity and Reliability of Data

Finally, the issue that may be of the highest concern to the accounting community is the integrity and reliability of data stored in the DW. Because of the importance of decisions based on DW information, high quality data is essential. Unfortunately, in many organizations, data in the various transaction-processing databases is not consistent in regard to formats, definitions, and usage. Therefore, control over the integration of data gathered from various locations into the warehouse is paramount.

During the design phase, important steps include a data audit of data sources to identify potential problems and the establishment of standards for data definitions and usage. From this, an extraction process can be refined to filter, scrub, and clean the data before loading into the warehouse.

Besides traditional data validation, such efforts may require code and format conversions, joining of data in different tables on the same platform or across platforms, and summarization of data to various levels. Another important requirement is to validate the data against business rules prior to loading. Such rules should be stored in the metadata storage facility, both in user understandable form as well as in the database language so that they can be applied consistently to all data upload or manipulation instances.

Controlling the accuracy and reliability of external or qualitative data may be particularly difficult. It may be impossible to devise rules to be stored in metadata to address this issue, since every case is different. The variety and complexity of external and qualitative data which makes it so useful to decision-makers, also makes it impossible to validate with pre-set rules.

Two new positions within the company may be created to aid data integrity efforts (Curtis 1998). A data auditor may be established within the IS function to provide day-to-day control over the validation of data. Data proponents, established within user groups, may be assigned responsibility for identifying and correcting errors within subsets of DW data fields and for scanning external and qualitative data for reasonableness.

Summary and Conclusions

A DW represents significant new opportunities for an organization to utilize internal and external information in an effective manner to help sustain competitive advantage. The use of DW technology directly relates to the strategic use of information to help focus and refine business tactics and strategy for improving performance on critical success factors. In the past, management was generally been limited to the analysis of operational data due to the lack of organization and accessibility of historical data. DW helps overcome these difficulties and limitations.

While there is a growing recognition of the important role DWs can play in supporting decision-makers, DW projects are risky and expensive undertakings. Stories describing DW development failures abound. In this paper we highlighted some of the issues to be considered for a successful DW project. Our recommendations for the successful implementation of a DW project include starting with the identification of actual business objectives to be met through a DW; designing the overall architecture while defining a narrow focus and choosing limited subject area(s) to implement in the first phase to gain experience; and finding a senior user sponsor for the project to help

obtain management commitment and required resources. Other factors to consider include careful study of subject areas to be included, level of granularity in the fact data, span of past data to be kept online, and development of a suitable optimization strategy through storage of pre-aggregated data as well proper indexing of data. Data integrity and quality are vital for user reliance on and use of data stored in a DW. Storage of external benchmarking data along with internal and external qualitative data is also very critical in helping users to access all relevant information to evaluate current status and to identify new opportunities. Finally, given the uncertain and unstructured nature of data analysis requirements of strategic and many operational users, designers should provide the broadest range of possible user access and analysis tools, while maintaining a high degree of security. Such support should include structured pre-programmed reports and screens, customized data exploration and drill down tools (e.g., OLAP), structured decision support facilities, and facilities for free form ad-hoc analysis of data for power users without any constraints, through native data access languages such as SQL. Future enhancements may include data visualization and web based access for ease of use. The usefulness and quality of a DW and its contribution in enhancing business performance will ultimately depend on developers' ability to make the DW environment user friendly.

Even when initial attempts to develop a data warehouse are unsuccessful, the organization will benefit in many ways. For example, consideration of a DW provides an organization the opportunity to review and formalize its business objectives, goals, plans, and strategy, thereby gaining organizational knowledge. Such efforts can also help identify, generate discussion on, and refine critical success factors to help organize and focus management efforts throughout the business. Such an exercise can also have beneficial outcomes for IS and accounting by helping these functional areas gain membership to the inner management team, recognized as part of the business team rather than outside specialists.

Many organizations have begun to exploit the potential of DW technology for gaining competitive advantage. Those waiting on the sidelines may not have a choice if they are to remain competitive. It is wise to begin immediately, since the successful DW project is not always the first undertaken by a company. DW projects can become successful showcases of accounting and IS departments, and help enhance those departments' organizational standings and alignment to business needs, when such projects are properly designed and successfully implemented. Our final caution is that efforts to develop a DW should be viewed as an ongoing process requiring continued financial commitment, rather than a one-time effort.

References

1. Amidon, Debra M. and David J. Skyrme, "Managing the Corporate Mind," *Internal Auditor*, December, pp. 19-20, 1998.
2. Arrowwood, Paul (Ernst and Young, LLP), "Data Warehousing," a presentation made at the University of Missouri, St. Louis, 1997.
3. Bird, Jane, "Data in Knowledge Out," *Management Today*, September, pp. 70-72, 1996.
4. Curtis, Mary B., "Control Issues in Data Warehousing," a presentation made at the Information Systems Audit and Control Association's 1998 national conference (CACS) in Chicago, April 1998.
5. Curtis, Mary B. and Kailash Joshi, "Internal Control Issues for Data Warehousing," *Information Systems Audit and Control Journal*, Vol. 4, pp. 40-45, 1997.
6. DePompa, Barbara, "Companies That Start with Data Marts Find it Difficult to Move to a Data Warehouse," *Information Week*, No. 605, November 11, pp. 68-79, 1996.
7. Fayyad, Usama, Gragory Piarersky-Shapiro, and Padhraic Smyth, "The KDD Process for Extracting Useful Knowledge from Volumes of Data," *Communications of the ACM*, Vol. 39, No. 11, November, pp. 27-34, 1996.
8. Francett, Barbara, "Decisions, Decisions: Users Take Stock of Data Warehouse Shelves," *Software Magazine*, Vol. 14, No. 8, August, pp. 63-70, 1994.
9. Hackathorn, Dick, "Reinventing Enterprise Systems Via Data Warehousing," The Data Warehousing Institute Annual Conference, Washington, DC, 1995.
10. Hackathorn, Dick, "Data Warehousing energizes Your Enterprise," *Datamation*, Vol. 41, No. 2, February 1, p. 38, 1995.
11. Inman, William H., "Security in the Data Warehouse / Internet Environment," *Information Systems Audit and Control Journal*, Vol. 4, pp. 8-11, 1997.
12. Inman, William H. and Chuck Kelley, "The 12 Rules of Data Warehouse for a Client/Server World," *Data Management Review*, Vol. 4, May, pp. 6-16, 1994.
13. Kimball, Ralph, *The Data Warehouse Tool Kit*, John Wiley and Sons, Inc., New York, 1996.
14. Lambert, Bob, "Data Warehousing Fundamentals: What You Need to Know to Succeed," a paper available at <http://www.data-warehouse.com/articles/lamb8.htm>, 12 pages, 1997.
15. Madnani, Kay (Perot Systems), "The Data Warehouse and Data Mining," a presentation made to the Information Systems Audit and Control Association, North Texas chapter, Dallas, Texas, June 11, 1998.
16. Meyer, Don and Casey Cannon, *Building a Better Data Warehouse*, Prentice Hall, Upper Saddle River, NJ, 1998.
17. Nelson, Emily, "Why Wal-Mart Sings, 'Yes, We Have Bananas,'" *Wall Street Journal*, October 6, 1998.
18. Richards, Louis, "Bringing Performance to Your Data Warehouse," a white paper from <http://www.disc.com/DISC12.html>, 16 pages, 1997.
19. Rob, Peter, and Carlos Coronel, *Database Systems: Design, Implementation, and Management*, Course Technology, Cambridge, Massachusetts, 1997.
20. Robinson, Teri, "To Avoid Pitfalls, Walk Before You Run," *Software Magazine*, Vol. 16, No. 10, October, pp. S8-S12, 1996.

21. Steinmann, Michael (Regional Technology Manager for Informix), "Delivering on the Future: Data Warehousing Solutions," presentation at the University of Missouri-St. Louis, December, St. Louis, MO, 1998.