

HUGH J. WATSON *and*  
BARBARA J. HALEY

# MANAGERIAL CONSIDERATIONS

Organizational resistance on many fronts can derail the most promising systems, even those designed to address a specific organizational pain.

**T**he architectural advantage of having a separate, customized repository of data for decision-support applications was recognized as early as the 1970s, when these systems were first developed [6]. In the mid-1980s, large retail, banking, and telecommunications corporations began building data warehouses, even though the term was not coined until the late 1980s by Bill Inmon [4]. Driving this development were business needs arising from such changes as fragmentation of mass markets into microsegments with special needs and the introduction of specialized technology by IBM and Teradata [5]. In the early 1990s, more data warehousing tools became available, and data warehousing became one of the hottest developments in the corporate computing world.

Today, most large organizations either have built a data warehouse or are at least seriously thinking about doing so. And more than 900 hardware, software, and services vendors now specialize in data warehousing technology.

BRIDGET RILEY/TATE GALLERY, LONDON/ ART RESOURCE, NY

These systems are expensive undertakings, however, with the typical warehouse costing in excess of \$1 million [7]. They are also challenging projects because of the myriad related technical and organizational issues. The initial failure rate is estimated to be as high as 50%, although many organizations succeed with later efforts after experiencing a failure. [5].

Even though data warehousing is important in many organizations, few researchers have studied it. In order to bring a more academic perspective to this topic, we began a research program in data warehousing in 1995. Our agenda has included case studies, telephone interviews, and mail surveys of both an exploratory and a confirmatory nature. We've given special attention to data warehousing from a managerial perspective—how data warehousing projects are approved, what issues managers have to monitor, and what benefits can be realized from these initiatives. We describe some of our most interesting and important findings, discussing how organizations gain approval for data warehousing projects, how data warehouses are implemented, which benefits are possible, and what post-implementation issues should be addressed.

## Winning Approval

The fundamental reason for building a data warehouse is to improve the quality of information in the organization. This was borne out in a 1997 survey we conducted of 121 attendees at a data warehousing conference sponsored by The Data Warehousing Institute [7], the leading professional organization for data warehousing managers and professionals. When asked their company's motivations for data warehousing, 38% of the respondents indicated having better access to information, 21% better and more accurate information, and 20% a single source of data.

Data warehousing projects are either data- or application-centered [5]. With the data-centered approach, existing data is placed in the warehouse for the purpose of supporting a number of applications and a variety of users; the data model on which the warehouse is based is independent of any single application. With the application-cen-

tered approach, the data warehousing initiative is driven by a single application or a small set of applications, often in sales or marketing, in which the potential business benefits are significant. While the data model for the warehouse has to support the targeted applications, care should be taken to ensure it also supports future applications.

Some companies start with an independent data mart, which is a smaller, subject-area (such as customer data) data repository. This approach is especially appropriate with an application-centered project. And this data mart can be viewed as an attractive alternative to a full-scale data warehouse, because it can be done quickly, at minimal

cost, and generate high payoff. It also can serve as a limited proof-of-concept for data warehousing. However, the danger is that a quick-hit success can lead to proliferation of independent data marts perpetuating an organization's "silos-of-information" problems. Consequently, independent data marts should be part of an integrated plan for building an enterprise-

wide data warehouse.

Sales and marketing provide many high-payoff opportunities for data warehousing, explaining why many data warehousing projects start in these areas. A variety of developments and considerations interact to make sales and marketing attractive application areas, such as capturing data about customer purchases through bar scanning; advances in database technology, making it possible to store and process massive amounts of data; moving to 1:1 marketing; and realizing that keeping and increasing sales to existing customers is at least as lucrative as attracting new customers. As a result, telephone companies analyze individual customer calling patterns to create attractive rate programs; retailers perform market-basket analyses to learn which products are purchased together, so promotions, coupons, and store layouts are optimally designed; and banks identify the profitability of individual customers and products.

Sponsorship for a data warehouse can come from a variety of sources. In our survey, respondents indicated that 48% of their warehouses were sponsored by IT and 41% had a senior executive or

## Calculating data

warehouse ROI can be a  
challenging undertaking given  
the uncertainty of the benefits,  
many of which are intangible.

functional area manager as the project sponsor. This variety is also true of how the projects are funded; the survey found that they are funded by IT (24%), by a functional area(s) (17%), and by both IT and a functional area(s) (49%). The ideal situation is to have sponsorship include both IT and business-unit representation.

In the relationship between a project that is data- or application-centered and the source of its sponsorship, IT sponsorship is somewhat correlated with the data-centered approach, while functional-area sponsorship is more highly correlated with the application-centered approach. To illustrate, we studied a workers' compensation organization with a generic, organizationwide problem of being unable to access needed data. Several IT managers took the lead in proposing creation of a data warehouse. Conversely, a large energy company had different accounting systems at each organizational level, and information had to be summarized and aggregated as it was passed through the systems. At the highest level, there was no way to drill down into details without considerable time and effort. A top-level corporate controller who reported directly to the CFO saw the potential benefits of a warehouse and championed a project addressing specific accounting needs.

Approval for a data warehousing project can be obtained in a variety of ways. Sometimes it is based on expected intangible benefits, like the ability to keep pace with a competitor. In other cases, there is the expectation by corporate management that future benefits should be quantified and integrated into a return on investment (ROI) analysis. Calculating ROI can be a challenging undertaking given the uncertainty of the benefits, many of which are intangible.

Our work has identified best practices in this difficult area. The starting point is to identify where "the pain" is in the organization. Pain refers to an opportunity, such as increasing sales, or a problem, such as high inventory levels, that might be addressed through data warehousing. Organizational pain can be identified through interviews with senior and functional area managers. The next

step is to baseline the current situation using performance metrics (including current sales levels). The next step is to work with management to identify the future state that is possible through data warehousing. A prototype is useful to show what can be done. The gap between the current and the future state is where data warehousing can contribute and should be quantified, including dollar benefits. This analysis becomes the core of the data warehousing business case.

Opposition to a data warehousing project can arise on many fronts within the organization. Among our respondents, 25% indicated that some opposition occurred in their organizations, from sources including IT (41%), functional area management (44%), and executive management (19%). Opposition from executive management was mainly due to concerns over the project's high cost. Functional area management expressed cost concerns and questioned the need for a data warehouse. IT opposition was particularly interesting, sometimes caused by concerns about potential loss of control, inadequate corporate data and in-house technical skills, and lack of an appropriate methodology.

## Opposition from

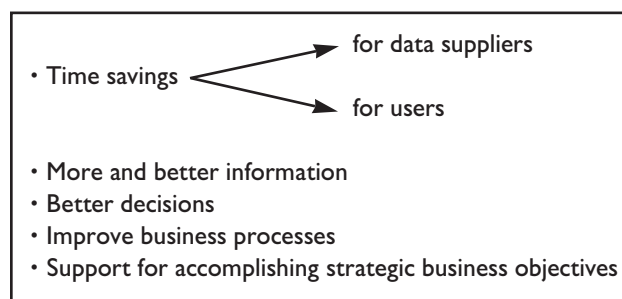
IT can be due to concerns about potential loss of control, inadequate corporate data and in-house technical skills, and lack of an appropriate methodology.

## Implementation Challenges

Data warehousing implementation challenges include a complex array of hardware and software components with highly specialized capabilities, such as symmetric multiprocessing, multidimensional databases, and data extraction and cleansing tools. Team members on these implementations need to address issues foreign to the operational world, such as the design of multidimensional views of data, the architecture of terabyte-size data stores, and the management of complex, ad hoc queries. The data used in data warehousing is often drawn from internal and external data sources containing nonstandardized data. All these challenges may be the primary reason why most organizations (59%) use consultants to implement warehousing [6].

In a survey of 111 data warehousing organizations that consider their initiative either an up-and-coming system or a runaway success, the most important factors for successful implementations

were organizational in nature. Strong management support and adequate corporate resources were found to have the strongest relationships with success, because these factors worked to overcome political resistance in the organization, address change-management issues, and increase organizationwide support for data warehousing. However, successful data warehousing initiatives also have to overcome some operational-level hurdles. One issue that has received a tremendous amount of attention recently is metadata management, which is a great challenge for most warehousing projects.



**Figure 1.** Business benefits from data warehousing

Managing metadata involves shielding the technical infrastructure of data warehousing from users while keeping the makeup of the data perfectly clear. Users are asking questions they previously could not ask, and they are thinking in new ways; therefore, they rely on data as a powerful tool to support their knowledge work. Metadata is the key to understanding the tool. But information extends much further—past simple meanings and relationships available to users in the operational world. To make decisions, users need to know such information as when the data was last refreshed, what sources provided the data, and how the data has been combined or manipulated. We have found that users without this information refrain from using the warehouse, spend inordinate amounts of time developing and testing queries, or ask someone more skilled to write their queries.

Creating and maintaining metadata are difficult tasks. Entering and maintaining sufficient metadata to satisfy users and developers take time and skill. Updates to metadata have to occur regularly to reflect changes in data sources, the underlying technical infrastructure, and transformation processes [5]. Additionally, integrating metadata throughout the technical environment can be nearly impossible. Each software tool in the data warehousing environment has proprietary metadata requirements, with unique standards and unique subsets of metadata.

The Metadata Coalition was established by leading vendors in 1995 to improve the exchange of metadata among software tools. Its mission involves creating a common metadata standard to facilitate the sharing of metadata among all data warehousing software tools. However, little progress has apparently been made to date, and skeptics expect that few vendors will reengineer their products to conform to the standards. Until progress is made, the warehouse development team needs to devote significant attention to providing users with information that helps them manipulate, manage, and navigate through the data in the data warehouse.

### Many Possible Benefits

A data warehouse by itself does not create value; value comes from the use of the data in the warehouse. A number of applications are potential users of warehouse data: queries, decision-support systems, executive information systems, and data mining [2]. See Figure 1 for the taxonomy of potential benefits. Prior to data warehousing, IT personnel and users spent inordinate amounts of time getting information, creating extract files or running ad hoc queries for users. Users also spent considerable effort accessing needed data and processing it in meaningful ways. According to at least one estimate, users spend three to four times as much time locating and gathering data as they do actually using it [1]. In one of our case studies, a quality engineer in a large manufacturing firm used to read more than 20,000 service call tickets a month looking for and investigating problems. Now, with the service call data in a warehouse, the engineer can check about 30 to 40 times as many service tickets. If a problem has a service incident rate as small as 0.1%, the company saves more than \$35,000 in service calls per month. Similarly, a bank we studied can now analyze the profitability of the different types of accounts it offers.

Improved decision making usually results from the better information available from a data warehouse. Returning to a quality example from our case studies, manufacturing firms are constantly looking for ways to produce higher-quality, lower-cost products. When a potentially better component part is identified, it is placed in a sample of products and monitored using service-call data in the data warehouse. This information now makes it possible to make decisions six months sooner than before the data warehouse was online and to make the decisions based on hard facts as opposed

to a decision maker's intuition.

Such benefits are ubiquitous; nearly every user of a data warehouse can provide multiple examples of them.

The greatest potential benefits of data warehousing are when the warehouse is used in the redesign of business processes and to support strategic business objectives. These are also the most difficult benefits to achieve, because of the amount of top management support, commitment, and involvement and the amount of organizational change required. As a result, many organizations don't realize these higher-order benefits.

Most organizations' legacy systems were developed many years ago in a relatively independent, nonintegrated manner, often focusing on functional areas, such as marketing systems and production systems. Today, these systems often have compatibility problems because of differing data models, definitions, updating cycles, and technical factors. As a result, it is difficult to access the kinds of data needed for a cross-functional view of the organization and to redesign systems that emphasize business processes.

Wal-Mart, the giant U.S. retailer, provides a leading example of how a company redesigned its business process around a data warehouse and how the warehouse became a key to corporate strategy [7]. Wal-Mart collects sales data at its nearly 2,800 stores, supercenters, and wholesale clubs and maintains this data in its more than 24TB-size warehouse. The company's 4,000 suppliers have access to the warehouse and are jointly responsible for managing Wal-Mart's inventory and shelf stock—down to the individual store level. This warehouse-based cooperation has greatly reduced inventory and made Wal-Mart more responsive to changing market conditions.

## Post-Implementation

Although companies may reap benefits at first, the direction of data warehousing can veer off course over time, and momentum can lag without the continued investment of time from the business side. This investment should focus on providing support to users to help them change the way they do the organization's business.

Our case studies show data warehousing supports decision making when business units get involved. These local units understand the questions users ask and the ways data warehousing can best support changing ways of doing business. The ideal profile of the person providing training and support includes functional knowledge, technical

know-how, and the ability to communicate with users. One public organization we studied identified both functional and technical people who collaboratively provide training and support to its data warehouse users. Users received tool training heavily augmented with applicable, real-task examples. Furthermore, this initial training was followed by training that explained the new ways users were expected to do their jobs. Topics included working with relational data, understanding the logic used to create SQL queries, and running statistics appropriate for sample data—all topics extending beyond manipulation of a software tool.

Organizations that fail to provide sufficient business-unit support begin to see informal user support contacts. This arrangement leads to a lot of inefficiency, because power users do less and less of their own work as they spend time solving the problems of their less technically savvy colleagues. A more formalized process needs to be put in place whereby support is recognized and help is better managed. **C**

## REFERENCES

1. Devlin, B. *Data Warehouse: From Architecture to Implementation*. Addison-Wesley, Reading, Mass., 1997.
2. Gray, P., and Watson, H. *Decision Support in the Data Warehouse*. Prentice-Hall, Upper Saddle River, N.J., 1998.
3. Haley, B. Implementing the decision support infrastructure: Key success factors in data warehousing. Ph.D. Dissertation, Univ. of Georgia, Athens, 1998.
4. Inmon, W. *Building the Data Warehouse*. Wiley, New York, 1992.
5. Kelly, S. *Data Warehousing in Action*. Wiley, New York, 1997.
6. Sprague, R., and Watson, H. *Decision Support for Management*. Prentice-Hall, Upper Saddle River, N.J., 1996.
7. Watson, H., and Haley, B. Data warehousing: A framework and survey of current practices. *J. Data Warehousing* 2, 1 (1997), 10–17.
8. Wilder, C. Chief of the year. *Inf. Week*, 662 (Dec. 22–29, 1997), 43–48.

**HUGH J. WATSON** (hwatson@arches.uga.edu) is a professor of MIS and holder of the C. Herman and Mary Virginia Terry Chair of Business Administration in the Terry College of Business at the University of Georgia.

**BARBARA J. HALEY** (bhaley@mindspring.com) is an assistant professor of commerce at the University of Virginia McIntire School of Commerce.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.