

PART 1

OVERVIEW AND CONCEPTS

COPYRIGHTED MATERIAL

CHAPTER 1

THE COMPELLING NEED FOR DATA WAREHOUSING

CHAPTER OBJECTIVES

- Understand the desperate need for strategic information
- Recognize the information crisis at every enterprise
- Distinguish between operational and informational systems
- Learn why all past attempts to provide strategic information failed
- Clearly see why data warehousing is the viable solution
- Understand business intelligence for an enterprise

As an information technology (IT) professional, you have worked on computer applications as an analyst, programmer, designer, developer, database administrator, or project manager. You have been involved in the design, implementation, and maintenance of systems that support day-to-day business operations. Depending on the industries you have worked in, you must have been involved in applications such as order processing, general ledger, inventory, human resources, payroll, in-patient billing, checking accounts, insurance claims, and so on.

These applications are important systems that run businesses. They process orders, maintain inventory, keep the accounting books, service the clients, receive payments, and process claims. Without these computer systems, no modern business can survive. Companies started building and using these systems in the 1960s and have become completely dependent on them. As an enterprise grows larger, hundreds of computer applications are needed to support the various business processes. These applications are effective in what they are designed to do. They gather, store, and process all the data needed to successfully perform the daily routine operations. They provide online information and produce a variety of reports to monitor and run the business.

Organizations achieve competitive advantage:

◆ Retail	◆ Manufacturing
◆ Customer Loyalty	◆ Cost Reduction
◆ Market Planning	◆ Logistics Management
◆ Financial	◆ Utilities
◆ Risk Management	◆ Asset Management
◆ Fraud Detection	◆ Resource Management
◆ Airlines	◆ Government
◆ Route Profitability	◆ Manpower Planning
◆ Yield Management	◆ Cost Control

Figure 1-1 Organizations' use of data warehousing.

In the 1990s, as businesses grew more complex, corporations spread globally, and competition became fiercer, business executives became desperate for information to stay competitive and improve the bottom line. The operational computer systems did provide information to run the day-to-day operations but what the executives needed were different kinds of information that could be used readily to make strategic decisions. The decision makers wanted to know which geographic regions to focus on, which product lines to expand, and which markets to strengthen. They needed the type of information with proper content and format that could help them make such strategic decisions. We may call this type of information strategic information as different from operational information. The operational systems, important as they were, could not provide strategic information. Businesses, therefore, were compelled to turn to new ways of getting strategic information.

Data warehousing is a new paradigm specifically intended to provide vital strategic information. In the 1990s, organizations began to achieve competitive advantage by building data warehouse systems. Figure 1-1 shows a sample of strategic areas where data warehousing had already produced results in different industries.

At the outset, let us now examine the crucial question: why do enterprises really need data warehouses? This discussion is important because unless we grasp the significance of this critical need, our study of data warehousing will lack motivation. So, please pay close attention.

ESCALATING NEED FOR STRATEGIC INFORMATION

While we discuss the clamor by enterprises for strategic information, we need to look at the prevailing information crisis that was holding them back, as well as the technology trends of the past few years that are working in our favor, enabling us to provide strategic information. Our discussion of the need for strategic information will not be complete unless we study the opportunities provided by strategic information and the risks facing a company without such information.

Who needs strategic information in an enterprise? What exactly do we mean by strategic information? The executives and managers who are responsible for keeping the enterprise

competitive need information to make proper decisions. They need information to formulate the business strategies, establish goals, set objectives, and monitor results.

Here are some examples of business objectives:

- Retain the present customer base
- Increase the customer base by 15% over the next 5 years
- Improve product quality levels in the top five product groups
- Gain market share by 10% in the next 3 years
- Enhance customer service level in shipments
- Bring three new products to market in 2 years
- Increase sales by 15% in the North East Division

For making decisions about these objectives, executives and managers need information for the following purposes: to get in-depth knowledge of their company's operations, review and monitor key performance indicators and note how these affect one another, keep track of how business factors change over time, and compare their company's performance relative to the competition and to industry benchmarks. Executives and managers need to focus their attention on customers' needs and preferences, emerging technologies, sales and marketing results, and quality levels of products and services. The types of information needed to make decisions in the formulation and execution of business strategies and objectives are broad-based and encompass the entire organization. All these types of essential information may be combined under the broad classification called strategic information.

Strategic information is not for running the day-to-day operations of the business. It is not intended to produce an invoice, make a shipment, settle a claim, or post a withdrawal from a bank account. Strategic information is far more important for the continued health and survival of the corporation. Critical business decisions depend on the availability of proper strategic information in an enterprise. Figure 1-2 lists the desired characteristics of strategic information.

INTEGRATED	Must have a single, enterprise-wide view.
DATA INTEGRITY	Information must be accurate and must conform to business rules.
ACCESSIBLE	Easily accessible with intuitive access paths, and responsive for analysis.
CREDIBLE	Every business factor must have one and only one value.
TIMELY	Information must be available within the stipulated time frame.

Figure 1-2 Characteristics of strategic information.

The Information Crisis

You may be working in the IT department of a large conglomerate or you may be part of a medium-sized company. Whatever may be the size of your company, think of all the various computer applications in your company. Think of all the databases and the quantities of data that support the operations of your company. How many years' worth of customer data is saved and available? How many years' worth of financial data is kept in storage? Ten years? Fifteen years? Where is all this data? On one platform? In legacy systems? In client/server applications?

We are faced with two startling facts: (1) organizations have lots of data, (2) information technology resources and systems are not effective at turning all that data into useful strategic information. Over the past two decades, companies have accumulated tons and tons of data about their operations. Mountains of data exist. Information is said to double every 18 months.

If we have such huge quantities of data in our organizations, why can't our executives and managers use this data for making strategic decisions? Lots and lots of information exists. Why then do we talk about an information crisis? Most companies are faced with an information crisis not because of lack of sufficient data, but because the available data is not readily usable for strategic decision making. These large quantities of data are very useful and good for running the business operations but hardly amenable for use in making decisions about business strategies and objectives.

Why is this so? First, the data of an enterprise is spread across many types of incompatible structures and systems. Your order processing system might have been developed 25 years ago and is still running on an old mainframe. Possibly, some of the data may still be on VSAM files. Your later credit assignment and verification system might be on a client/server platform and the data for this application might be in relational tables. The data in a corporation resides in various disparate systems, multiple platforms, and diverse structures. The more technology your company has used in the past, the more disparate the data of your company will be. But, for proper decision making on overall corporate strategies and objectives, we need information integrated from all systems.

Data needed for strategic decision making must be in a format suitable for easy analysis to spot trends. Executives and managers need to look at trends over time and steer their companies in the proper direction. The tons of available operational data cannot be readily used to discern trends. Operational data is event-driven. You get snapshots of transactions that happen at specific times. You have data about units of sale of a single product in a specific order on a given date to a certain customer. In the operational systems, you do not readily have the trends of a single product over the period of a month, a quarter, or a year.

For strategic decision making, executives and managers must be able to review data from different business viewpoints. For example, they must be able to review and analyze sales quantities by product, salesperson, district, region, and customer groups. Can you think of operational data being readily available for such analysis? Operational data is not directly suitable for review from different viewpoints.

Technology Trends

Those of us who have worked in the information technology field for two or three decades have witnessed the breathtaking changes that have taken place. First, the name of the computer department in an enterprise went from "data processing" to "management information

systems,” then to “information systems,” and more recently to “information technology.” The entire spectrum of computing has undergone tremendous changes. The computing focus itself has changed over the years. Old practices could not meet new needs. Screens and preformatted reports are no longer adequate to meet user requirements.

Over the years, the price of MIPS (million instructions per second) is continuing to decline, digital storage is costing less and less, and network bandwidth is increasing as its price decreases. Specifically, we have seen explosive changes in these critical areas:

- Computing technology
- Human–machine interface
- Processing options

Figure 1-3 illustrates these waves of explosive growth.

What is our current position in the technology revolution? Hardware economics and miniaturization allow a workstation on every desk and provide increasing power at reducing costs. New software provides easy-to-use systems. Open systems architecture creates cooperation and enables the use of multivendor software. Improved connectivity, networking, and the Internet open up interaction with an enormous number of systems and databases.

All of these improvements in technology are meritorious. These have made computing faster, cheaper, and widely available. But what is their relevance to the escalating need for strategic information? Let us understand how the current state of the technology is conducive to providing strategic information.

Providing strategic information requires collection of large volumes of corporate data and storing it in suitable formats. Technology advances in data storage and reduction in storage costs readily accommodate data storage needs for strategic decision-support systems. Analysts, executives, and managers use strategic information interactively to analyze and spot business trends. The user will ask a question and get the results, then ask another question, look at the results, and ask yet another question. This interactive process continues. Tremendous advances in interface software make such interactive analysis possible.

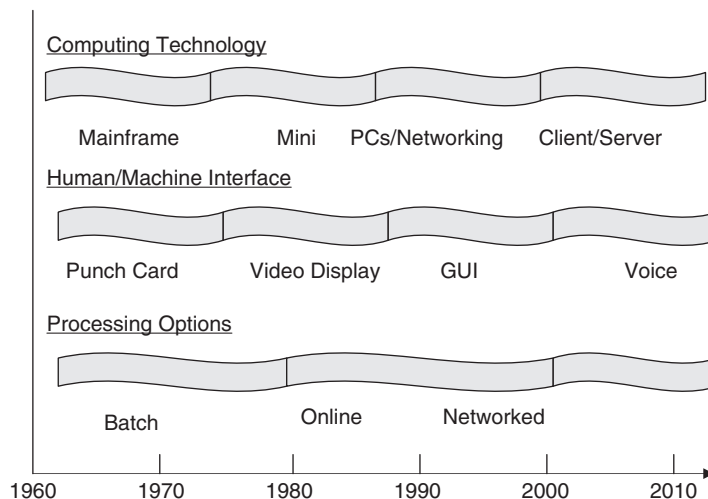


Figure 1-3 Explosive growth of information technology.

Processing large volumes of data and providing interactive analysis requires extra computing power. The explosive increase in computing power and its lower costs make provision of strategic information feasible. What we could not accomplish a few years earlier for providing strategic information is now possible with the current advanced stage of information technology.

Opportunities and Risks

We have looked at the information crisis that exists in every enterprise and grasped that in spite of lots of operational data in the enterprise, data suitable for strategic decision making is not available. Yet, the current state of the technology can make it possible to provide strategic information. While we are still discussing the escalating need for strategic information by companies, let us ask some basic questions. What are the opportunities available to companies resulting from the possible use of strategic information? What are the threats and risks resulting from the lack of strategic information available in companies?

Here are some examples of the opportunities made available to companies through the use of strategic information:

- A business unit of a leading long-distance telephone carrier empowers its sales personnel to make better business decisions and thereby capture more business in a highly competitive, multibillion-dollar market. A Web-accessible solution gathers internal and external data to provide strategic information.
- Availability of strategic information at one of the largest banks in the United States with assets in the \$250 billion range allows users to make quick decisions to retain their valued customers.
- In the case of a large health management organization, significant improvements in health care programs are realized, resulting in a 22% decrease in emergency room visits, 29% decrease in hospital admissions for asthmatic children, potentially sight-saving screenings for hundreds of diabetics, improved vaccination rates, and more than 100,000 performance reports created annually for physicians and pharmacists.
- At one of the top five U.S. retailers, strategic information combined with Web-enabled analysis tools enables merchants to gain insights into their customer base, manage inventories more tightly, and keep the right products in front of the right people at the right place at the right time.
- A community-based pharmacy that competes on a national scale with more than 800 franchised pharmacies coast to coast gains in-depth understanding of what customers buy, resulting in reduced inventory levels, improved effectiveness of promotions and marketing campaigns, and improved profitability for the company.
- A large electronics company saves millions of dollars a year because of better management of inventory.

On the other hand, consider the following cases where risks and threats of failures existed before strategic information was made available for analysis and decision making:

- With an average fleet of about 150,000 vehicles, a nationwide car rental company can easily get into the red at the bottom line if fleet management is not effective. The fleet is the biggest cost in that business. With intensified competition, the potential for failure is immense if the fleet is not managed effectively. Car idle time must be kept to an

absolute minimum. In attempting to accomplish this, failure to have the right class of car available in the right place at the right time, all washed and ready, can lead to serious loss of business.

- For a world-leading supplier of systems and components to automobile and light truck equipment manufacturers, serious challenges faced included inconsistent data computations across nearly 100 plants, inability to benchmark quality metrics, and time-consuming manual collection of data. Reports needed to support decision making took weeks. It was never easy to get company-wide integrated information.
- For a large utility company that provided electricity to about 25 million consumers in five mid-Atlantic states in the United States, deregulation could result in a few winners and lots of losers. Remaining competitive and perhaps even just surviving depended on centralizing strategic information from various sources, streamlining data access, and facilitating analysis of the information by the business units.

FAILURES OF PAST DECISION-SUPPORT SYSTEMS

Assume a specific scenario. The marketing department in your company has been concerned about the performance of the West Coast region and the sales numbers from the monthly report this month are drastically low. The marketing vice president is agitated and wants to get some reports from the IT department to analyze the performance over the past two years, product by product, and compared to monthly targets. He wants to make quick strategic decisions to rectify the situation. The CIO wants your boss to deliver the reports as soon as possible. Your boss runs to you and asks you to stop everything and work on the reports. There are no regular reports from any system to give the marketing department what they want. You have to gather the data from multiple applications and start from scratch. Does this sound familiar?

At one time or another in your career in information technology, you must have been exposed to situations like this. Sometimes, you may be able to get the information required for such ad hoc reports from the databases or files of one application. Usually this is not so. You may have to go to several applications, perhaps running on different platforms in your company environment, to get the information. What happens next? The marketing department likes the ad hoc reports you have produced. But now they would like reports in a different format, containing more information that they did not think of originally. After the second round, they find that the contents of the reports are still not exactly what they wanted. They may also find inconsistencies among the data obtained from different applications.

The fact is that for nearly two decades or more, IT departments have been attempting to provide information to key personnel in their companies for making strategic decisions. Sometimes an IT department could produce ad hoc reports from a single application. In most cases, the reports would need data from multiple systems, requiring the writing of extract programs to create intermediary files that could be used to produce the ad hoc reports.

Most of these attempts by IT in the past ended in failure. The users could not clearly define what they wanted in the first place. Once they saw the first set of reports, they wanted more data in different formats. The chain continued. This was mainly because of the very nature of the process of making strategic decisions. Information needed for strategic decision making has to be available in an interactive manner. The user must be able to query online, get results, and query some more. The information must be in a format suitable for analysis.

In order to appreciate the reasons for the failure of IT to provide strategic information in the past, we need to consider how IT was attempting to do this all these years. Let us, therefore, quickly run through a brief history of decision support systems.

History of Decision-Support Systems

Depending on the size and nature of the business, most companies have gone through the following stages of attempts to provide strategic information for decision making.

Ad hoc Reports. This was the earliest stage. Users, especially from marketing and finance, would send requests to IT for special reports. IT would write special programs, typically one for each request, and produce the ad hoc reports.

Special Extract Programs. This stage was an attempt by IT to anticipate somewhat the types of reports that would be requested from time to time. IT would write a suite of programs and run the programs periodically to extract data from the various applications. IT would create and keep the extract files to fulfill any requests for special reports. For any reports that could not be run off the extracted files, IT would write individual special programs.

Small Applications. In this stage, IT formalized the extract process. IT would create simple applications based on the extracted files. The users could stipulate the parameters for each special report. The report printing programs would print the information based on user-specific parameters. Some advanced applications would also allow users to view information through online screens.

Information Centers. In the early 1970s, some major corporations created what were called information centers. The information center typically was a place where users could go to request ad hoc reports or view special information on screens. These were predetermined reports or screens. IT personnel were present at these information centers to help the users to obtain the desired information.

Decision-Support Systems. In this stage, companies began to build more sophisticated systems intended to provide some semblance of strategic information. Again, similar to the earlier attempts, these systems were supported by extracted files. The systems were menu-driven and provided online information and also the ability to print special reports. Many such decision-support systems were for marketing.

Executive Information Systems. This was an attempt to bring strategic information to the executive desktop. The main criteria were simplicity and ease of use. The system would display key information every day and provide the ability to request simple, straightforward reports. However, only preprogrammed screens and reports were available. After seeing the total countrywide sales, if the executive wanted to see the analysis by region, by product, or by another dimension, it was not possible unless such breakdowns were already preprogrammed. This limitation caused frustration and executive information systems did not last long in many companies.

Inability to Provide Information

Every one of the past attempts at providing strategic information to decision makers was unsatisfactory. Figure 1-4 depicts the inadequate attempts by IT to provide strategic information. As IT professionals, we are all familiar with the situation.

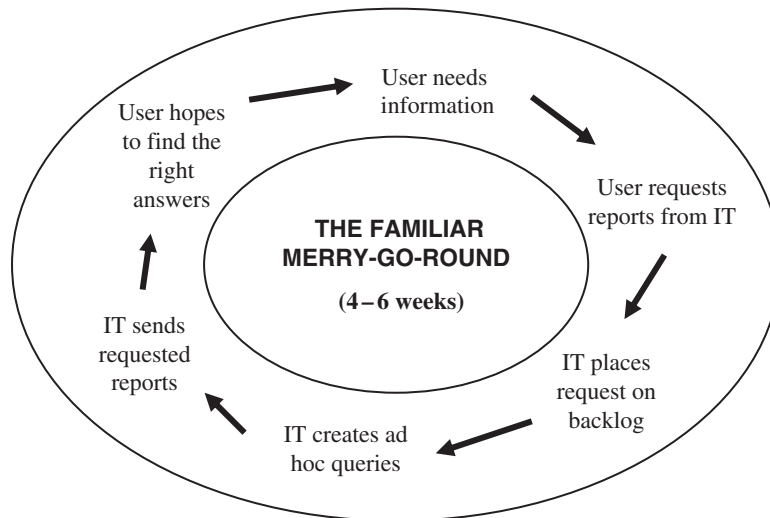


Figure 1-4 Inadequate attempts by IT to provide strategic information.

Here are some of the factors relating to the inability to provide strategic information:

- IT receives too many ad hoc requests, resulting in a large overload. With limited resources, IT is unable to respond to the numerous requests in a timely fashion.
- Requests are too numerous; they also keep changing all the time. The users need more reports to expand and understand the earlier reports.
- The users find that they get into the spiral of asking for more and more supplementary reports, so they sometimes adapt by asking for every possible combination, which only increases the IT load even further.
- The users have to depend on IT to provide the information. They are not able to access the information themselves interactively.
- The information environment ideally suited for strategic decision making has to be very flexible and conducive for analysis. IT has been unable to provide such an environment.

OPERATIONAL VERSUS DECISION-SUPPORT SYSTEMS

Is there an underlying reason for the failure of all the previous attempts by IT to provide strategic information? What has IT been doing all along? The fundamental reason for the inability to provide strategic information is that we have been trying all along to provide strategic information from the operational systems. These operational systems such as order processing, inventory control, claims processing, outpatient billing, and so on are not designed or intended to provide strategic information. If we need the ability to provide strategic information, we must get the information from altogether different types of systems. Only specially designed decision support systems or informational systems can provide strategic information. Let us understand why.

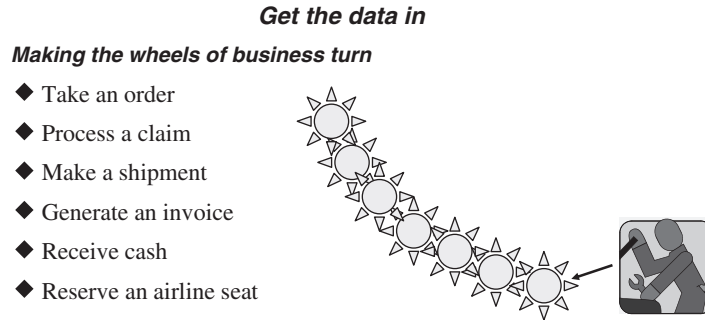


Figure 1-5 Operational systems.

Making the Wheels of Business Turn

Operational systems are online transaction processing (OLTP) systems. These are the systems that are used to run the day-to-day core business of the company. They are the so-called bread-and-butter systems. Operational systems make the wheels of business turn (see Fig. 1-5). They support the basic business processes of the company. These systems typically get the data into the database. Each transaction processes information about a single entity such as a single order, a single invoice, or a single customer.

Watching the Wheels of Business Turn

On the other hand, specially designed and built decision-support systems are not meant to run the core business processes. They are used to watch how the business runs, and then make strategic decisions to improve the business (see Fig. 1-6).

Decision-support systems are developed to get strategic information *out of* the database, as opposed to OLTP systems that are designed to put the data *into* the database. Decision-support systems are developed to provide strategic information.

Different Scope, Different Purposes

Therefore, we find that in order to provide strategic information we need to build informational systems that are different from the operational systems we have been building to run the basic business. It will be worthless to continue to dip into the operational systems for

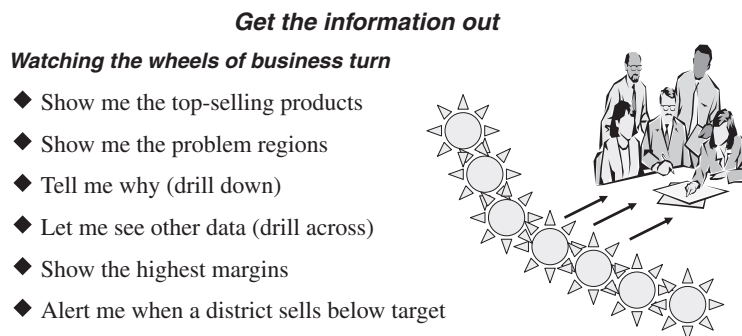


Figure 1-6 Decision-support systems.

How are they different?

	OPERATIONAL	INFORMATIONAL
Data Content	Current values	Archived, derived, summarized
Data Structure	Optimized for transactions	Optimized for complex queries
Access Frequency	High	Medium to low
Access Type	Read, update, delete	Read
Usage	Predictable, repetitive	Ad hoc, random, heuristic
Response Time	Sub-seconds	Several seconds to minutes
Users	Large number	Relatively small number

Figure 1-7 Operational and informational systems.

strategic information as we have been doing in the past. As companies face fiercer competition and businesses become more complex, continuing the past practices will only lead to disaster.

We need to design and build informational systems

- That serve different purposes
- Whose scopes are different
- Whose data content is different
- Where the data usage patterns are different
- Where the data access types are different

Figure 1-7 summarizes the differences between the traditional operational systems and the newer informational systems that need to be built.

DATA WAREHOUSING—THE ONLY VIABLE SOLUTION

At this stage of our discussion, we now realize that we do need different types of decision-support systems to provide strategic information. The type of information needed for strategic decision making is different from that available from operational systems. We need a new type of system environment for the purpose of providing strategic information for analysis, discerning trends, and monitoring performance.

Let us examine the desirable features and processing requirements of this new type of system environment. Let us also consider the advantages of this type of system environment designed for strategic information.

A New Type of System Environment

The desired features of the new type of system environment are:

- Database designed for analytical tasks
- Data from multiple applications
- Easy to use and conducive to long interactive sessions by users

- Read-intensive data usage
- Direct interaction with the system by the users without IT assistance
- Content updated periodically and stable
- Content to include current and historical data
- Ability for users to run queries and get results online
- Ability for users to initiate reports

Processing Requirements in the New Environment

Most of the processing in the new environment for strategic information will have to be analytical. There are at least four levels of analytical processing requirements:

1. Running of simple queries and reports against current and historical data.
2. Ability to perform “what if” analysis in many different ways.
3. Ability to query, step back, analyze, and then continue the process to any desired length.
4. Ability to spot historical trends and apply them in future interactive processes.

Strategic Information from the Data Warehouse

This new system environment that users desperately need to obtain strategic information happens to be the new paradigm of data warehousing. Beginning with the late 1980s and into the 1990s enterprises began building such system environments. This new environment, known as the data warehouse environment, is kept separate from the system environment that supports the routine day-to-day operations. The data warehouse essentially has become the source of strategic information for the enterprise to enable strategic decision making. The data warehouse has proved to be the only viable solution. We have clearly seen that solutions based on the data extracted from operational systems have all been totally unsatisfactory. Figure 1-8 shows the general overview of the data warehouse as the source of strategic information for the enterprise.

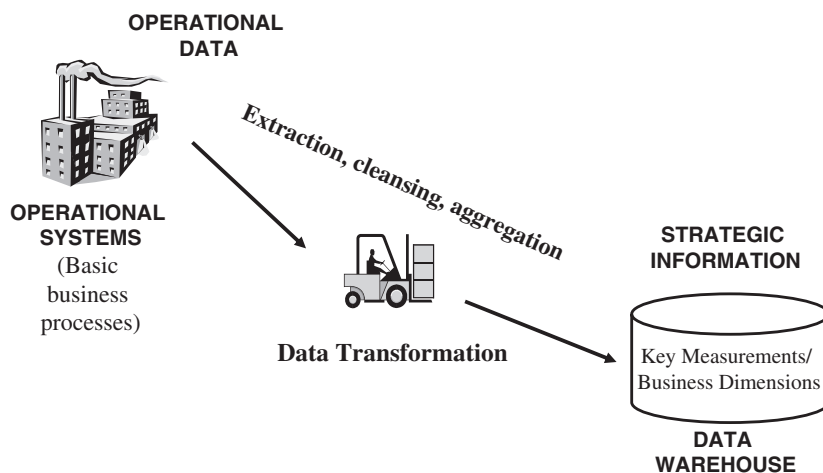


Figure 1-8 General overview of the data warehouse.

At a high level of interpretation, the data warehouse contains critical metrics of the business processes stored along business dimensions. For example, a data warehouse might contain units of sales, by product, day, customer group, sales district, sales region, and promotion. Here the business dimensions are product, day, customer group, sales district, sales region, and sales promotion type. Unit sales represent the metrics being measured across products, days, customer groups, sales districts, sales regions, and sales promotion types.

From where does the data warehouse get its data? The data is derived from the operational systems that support the basic business processes of the organization. In between the operational systems and the data warehouse, there is a data staging area. In this staging area, the operational data is cleansed and transformed into a form suitable for placement in the data warehouse for easy retrieval.

DATA WAREHOUSE DEFINED

We have reached the strong conclusion that data warehousing is the only viable solution for providing strategic information. We arrived at this conclusion based on the functions of the new system environment called the data warehouse. So, let us try to come up with a functional definition of the data warehouse.

The data warehouse is an informational environment that:

- Provides an integrated and total view of the enterprise.
- Makes the enterprise's current and historical information easily available for strategic decision making.
- Makes decision-support transactions possible without hindering operational systems.
- Renders the organization's information consistent.
- Presents a flexible and interactive source of strategic information.

A Simple Concept for Information Delivery

In the final analysis, data warehousing is a simple concept. It is born out of the need for strategic information and is the result of the search for a new way to provide such information. The methods of the previous decades using the operational computing environment were unsatisfactory. The new concept is not to generate fresh data, but to make use of the large volumes of existing data and to transform it into forms suitable for providing strategic information.

The data warehouse exists to answer questions users have about the business, the performance of the various operations, the business trends, and about what can be done to improve the business. The data warehouse exists to provide business users with direct access to data, to provide a single unified version of the key performance indicators, to record the past accurately, and to provide the ability for viewing the data from many different perspectives. In short, the data warehouse is there to support decisional processes.

Data warehousing is really a simple concept: Take all the data you already have in the organization, clean and transform it, and then provide useful strategic information. What could be simpler than that?

An Environment, Not a Product

A data warehouse is not a single software or hardware product you purchase to provide strategic information. It is, rather, a computing environment where users can find strategic

information, an environment where users are put directly in touch with the data they need to make better decisions. It is a user-centric environment.

Let us summarize the characteristics of this new computing environment called the data warehouse:

- An ideal environment for data analysis and decision support
- Fluid, flexible, and interactive
- 100% user-driven
- Very responsive and conducive to the ask–answer–ask again pattern
- Provides the ability to discover answers to complex, unpredictable questions

A Blend of Many Technologies

Let us reexamine the basic concept of data warehousing. The basic concept of data warehousing is:

- Take all the data from the operational systems.
- Where necessary, include relevant data from outside, such as industry benchmark indicators.
- Integrate all the data from the various sources.
- Remove inconsistencies and transform the data.
- Store the data in formats suitable for easy access for decision making.

Different technologies are, therefore, needed to support these functions. Figure 1-9 shows how a data warehouse is a blend of the many technologies needed for the various functions.

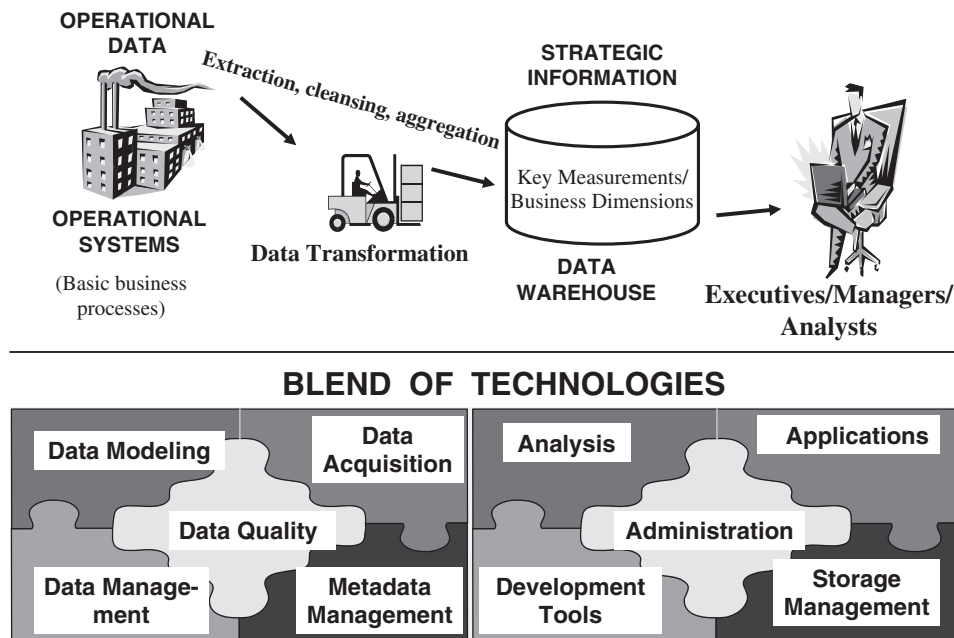


Figure 1-9 The data warehouse: a blend of technologies.

Although many technologies are in use, they all work together in a data warehouse. The end result is the creation of a new computing environment for the purpose of providing the strategic information every enterprise needs desperately.

THE DATA WAREHOUSING MOVEMENT

As enterprises began to realize the effectiveness of data warehousing, more and more organizations jumped on the bandwagon and data warehousing began to spread at a rapid rate. First the large companies that were able to quickly afford the outlay of resources began to launch data warehousing projects. Medium-sized companies also entered the data warehousing arena. Soon several businesses began to reap the benefits provided by data warehousing. Much research began to be focused on this new phenomenon. Many vendors began to offer hardware and software products to support the different functions within the data warehouse.

Prior to the data warehousing concept with an architectural model for the movement of data from operational systems to decision support environments, companies attempted multiple decision-support environments within their organizations. This had to be done with enormous costs fraught with large amounts of data redundancies and inconsistencies. Each decision-support environment was intended to serve specific groups of users for limited purposes. The adoption of data warehousing changed all of this. Similar to industrial warehouses, data warehouses are intended for large-scale collection and storage of corporate data to provide strategic information for the overall needs. Just as products stored in industrial warehouses are distributed to retail stores or marts, data stored in data warehouses may be channeled to data marts for specific users.

Data Warehousing Milestones

As data warehousing gained acceptance during the 1980s and 1990s, we may trace some of the highlights of the movement. Listed below are the significant milestones during the initial phase of the movement:

- 1983—Teradata introduces a database management system (DBMS) designed for decision-support systems.
- 1988—The article *An Architecture for a Business and Information Systems* introducing the term “business data warehouse” is published by Barry Devlin and Paul Murphy in the *IBM Systems Journal*.
- 1990—Red Brick Systems introduces Red Brick Warehouse, a DBMS specifically for data warehousing.
- 1991—Bill Inmon publishes his book *Building the Data Warehouse* (he is generally considered the father of data warehousing).
- 1991—Prism Solutions introduces Prism Warehouse Manager software for developing a data warehouse.
- 1995—The Data Warehousing Institute, a premier institution that promotes data warehousing is founded. (This institution has since emerged as the leading voice in the data warehousing and business intelligence arena providing education, research, and support.)

- 1996—Ralph Kimball publishes a seminal book *The Data Warehousing Toolkit*. (He is among the top authorities in the field of data warehousing and decision support systems.)
- 1997—Oracle 8, with support for STAR schema queries, is released.

Initial Challenges

As the adoption of data warehousing by organizations continued, even those companies that implemented data warehouses faced significant challenges that promoted a moderate shift from the original implementations. Here is a list of the key challenges that had to be overcome:

- Customers had become more sophisticated and savvy, pressing for greater service, improved quality, and innovative customization.
- Government deregulation of industries exposed companies to fiercer competition and the need for leaner operation.
- Expansion of globalization opened the arena for competitors, more in number and greater in power.
- New privacy regulations created the need to revise methods of collection and use of information.
- Improper architecture of some initial data warehousing systems produced fragmented views of corporate data and tended to produce disparate information silos.
- Query, reporting, and analysis tools provided to the users in the early data warehousing environments for self-service proved to be too complex and overwhelming for use by the users themselves.
- The promises of early data warehouse environments to provide user-friendly tools for the masses remained unfulfilled.

EVOLUTION OF BUSINESS INTELLIGENCE

The initial challenges following the adoption of early data warehousing systems forced companies to take a second look at providing decision support. Companies began to perceive that the goal of decision-support systems is twofold: transformation of data to information; derivation of knowledge from information. Each of these two aspects needs to be emphasized and strengthened appropriately to provide the necessary results. Business intelligence for an organization requires two environments, one to concentrate on transformation of data into information and the other to deal with transformation of information into knowledge.

Business intelligence (BI), therefore, is a broad group of applications and technologies. First, the term refers to the systems and technologies for gathering, cleansing, consolidating, and storing corporate data. Next, business intelligence relates to the tools, techniques, and applications for analyzing the stored data. The Gartner Group popularized BI as an umbrella term to include concepts and methods to improve business decision making by fact-based support systems. The Data Warehousing Institute compares BI to a data refinery. Similar to an oil refinery, a BI setting takes data as the raw material, collects it, refines it, and processes it into several information products.

BI: Two Environments

When you consider all that BI encompasses, you may view BI for an enterprise as composed of two environments:

Data to Information. In this environment data from multiple operational systems are extracted, integrated, cleansed, transformed and stored as information in specially designed repositories.

Information to Knowledge. In this environment analytical tools are made available to users to access and analyze the information content in the specially designed repositories and turn information into knowledge.

BI: Data Warehousing and Analytics

As some of the early challenges indicated, sufficient and separate attention needs to be given to the two environments that BI encompasses. In today's businesses, extraction, consolidation, transformation, and storing of data as strategic information is a formidable task. Again, using this information with sophisticated tools for proper decision making is equally challenging. Therefore, the trend is to consider these as two distinct environments for corporate BI. Vendors also tend to specialize in tools appropriate for these two distinct environments.

However, the two environments are complementary and need to work together. Figure 1-10 shows the two complementary environments, the data warehousing environment, which transforms data into information, and the analytical environment, which produces knowledge from information. As we proceed from chapter to chapter, we will keep expanding and intensifying our discussion of these two environments.

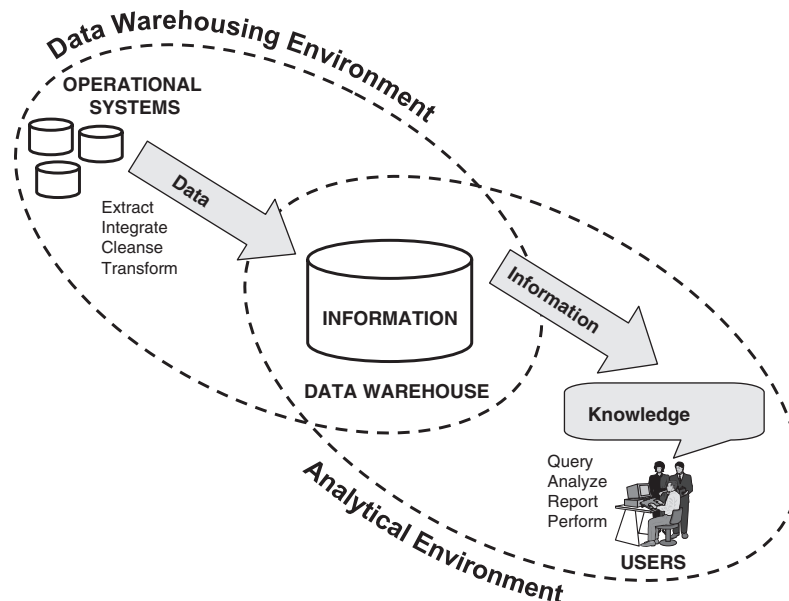


Figure 1-10 BI: data warehousing and analytical environments.

CHAPTER SUMMARY

- Companies are desperate for strategic information to counter fiercer competition, extend market share, and improve profitability.
- In spite of tons of data accumulated by enterprises over the past decades, every enterprise is caught in the middle of an information crisis. Information needed for strategic decision making is not readily available.
- All the past attempts by IT to provide strategic information have been failures. This was mainly because IT has been trying to provide strategic information from operational systems.
- Informational systems are different from the traditional operational systems. Operational systems are not designed for strategic information.
- We need a new type of computing environment to provide strategic information. The data warehouse promises to be this new computing environment.
- Data warehousing is the viable solution. There is a compelling need for data warehousing in every enterprise.
- The challenges faced in early data warehouse implementations led the movement towards maturity.
- The notion of business intelligence for an enterprise has evolved as an umbrella concept embracing data warehousing and analytics to transform data into information and information into knowledge.

REVIEW QUESTIONS

1. What do we mean by strategic information? For a commercial bank, name five types of strategic objectives.
2. Do you agree that a typical retail store collects huge volumes of data through its operational systems? Name three types of transaction data likely to be collected by a retail store in large volumes during its daily operations.
3. Examine the opportunities that can be provided by strategic information for a medical center. Can you list five such opportunities?
4. Why were all the past attempts by IT to provide strategic information failures? List three concrete reasons and explain.
5. Describe five differences between operational systems and informational systems.
6. Why are operational systems not suitable for providing strategic information? Give three specific reasons and explain.
7. Name six characteristics of the computing environment needed to provide strategic information.
8. What types of processing take place in a data warehouse? Describe.
9. A data warehouse is an environment, not a product. Discuss.

10. Data warehousing is the only viable means to resolve the information crisis and to provide strategic information. List four reasons to support this assertion and explain them.

EXERCISES

1. Match the columns:

- | | |
|----------------------------|---------------------------------|
| 1. information crisis | A. OLTP applications |
| 2. strategic information | B. produce ad hoc reports |
| 3. operational systems | C. explosive growth |
| 4. information center | D. despite lots of data |
| 5. data warehouse | E. data cleaned and transformed |
| 6. order processing | F. users go to get information |
| 7. executive information | G. used for decision making |
| 8. data staging area | H. environment, not product |
| 9. extract programs | I. for day-to-day operations |
| 10. information technology | J. simple, easy to use |

2. The current trends in hardware/software technology make data warehousing feasible. Explain with some examples how exactly technology trends do help.
3. You are the IT director of a nationwide insurance company. Write a memo to the executive vice president explaining the types of opportunities that can be realized with readily available strategic information.
4. For an airline company, how can strategic information increase the number of frequent flyers? Discuss giving specific details.
5. You are a senior analyst in the IT department of a company manufacturing automobile parts. The marketing VP is complaining about the poor response by IT in providing strategic information. Draft a proposal to him introducing the concept of business intelligence and how data warehousing and analytics as part of business intelligence for your company would be the optimal solution.

