

The Compelling Need for Data Warehousing

Dr. Nidhi Khurana
Associate Professor

A decorative graphic in the top-left corner featuring several dice. There is a red die, a white die with black pips, a black die with white pips, and a small white die with black pips.

What all Business Managers Need to Understand about Business Intelligence

- ▶ 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



How various organizations use Data Warehousing

◆ Retail

- Customer Loyalty
- Market Planning

◆ Financial

- Risk Management
- Fraud Detection

◆ Airlines

- Route Profitability
- Yield Management

◆ Manufacturing

- Cost Reduction
- Logistics Management

◆ Utilities

- Asset Management
- Resource Management

◆ Government

- Manpower Planning
- Cost Control



Strategic 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.
- ▶ Let's see what these business objectives are:



Key Examples of Business Objectives

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



Things to Think about in your current workplace -

- 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?



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.



Information Overload!


▶ Some interesting facts:

- ▶ Organizations have a great deal of data
- ▶ information technology resources and systems are not effective at turning all that data into useful strategic information.
- ▶ Over the past 20 years companies generate tons of information about the operational side of their business
- ▶ Mountains of data exist
- ▶ Information is said to double every 18 months

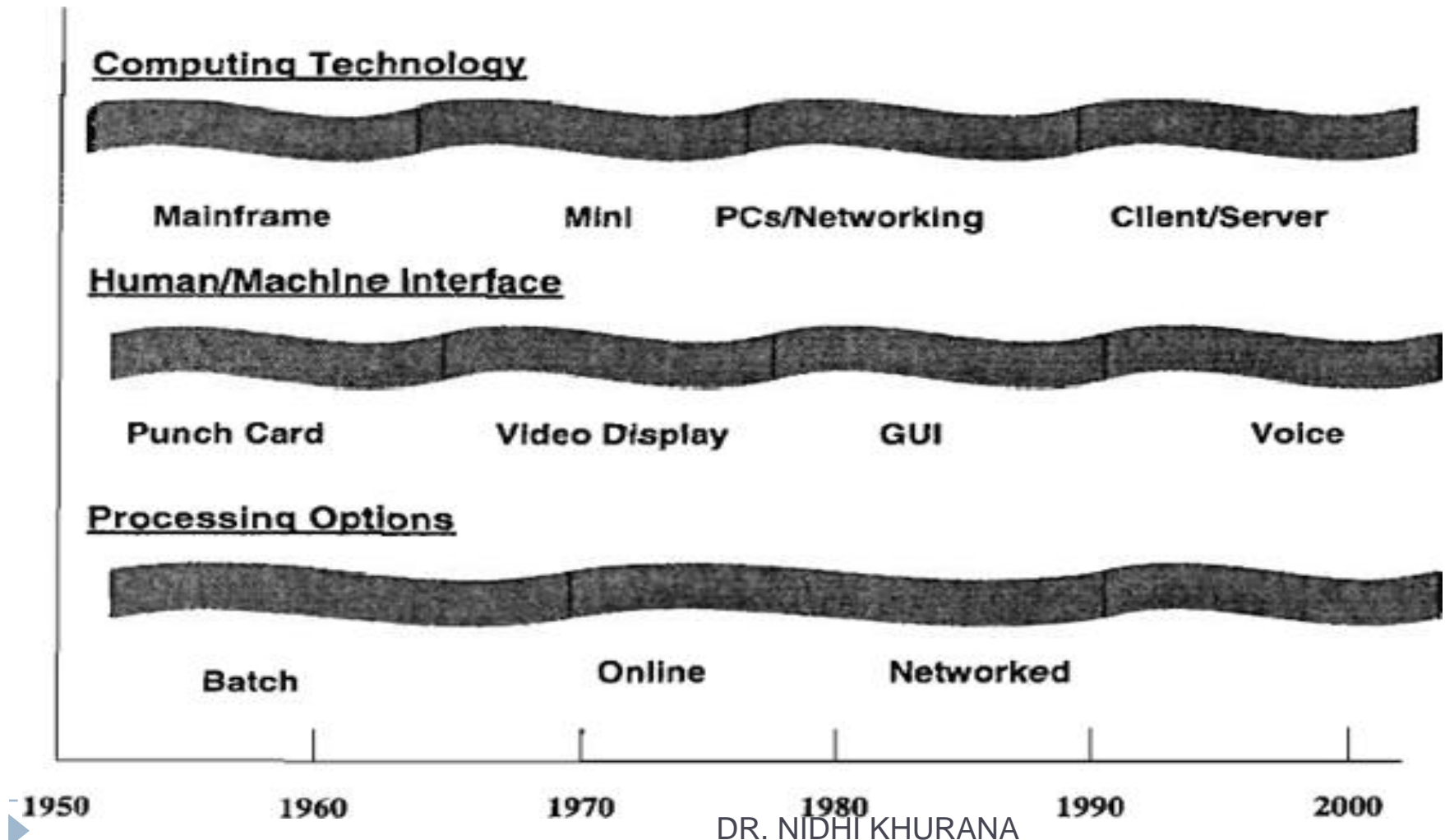


Processing Costs?

- ▶ MIPs is continuing to decline
- ▶ digital storage is costing less and less
- ▶ 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
- ▶ The next slide shows some of these growth areas



Explosive Growth of Information Technology






Examples of the opportunities made available 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.

Examples of the opportunities made available through the use of strategic information: (Cont'd)

- 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.

Examples of the opportunities made available through the use of strategic information: (Cont'd)

- 
- A cluster of several dice in various colors (red, white, black) and orientations, some showing different faces.
- ▶ 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.



happens without strategic information)

- 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.



The Other Side of the Coin (Cont'd)

- ▶ 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.



The Other Side of the Coin (Cont'd)

- ▶ 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 surviving itself depended on centralizing strategic information from various sources, streamlining data access, and facilitating analysis of the information by the business units.



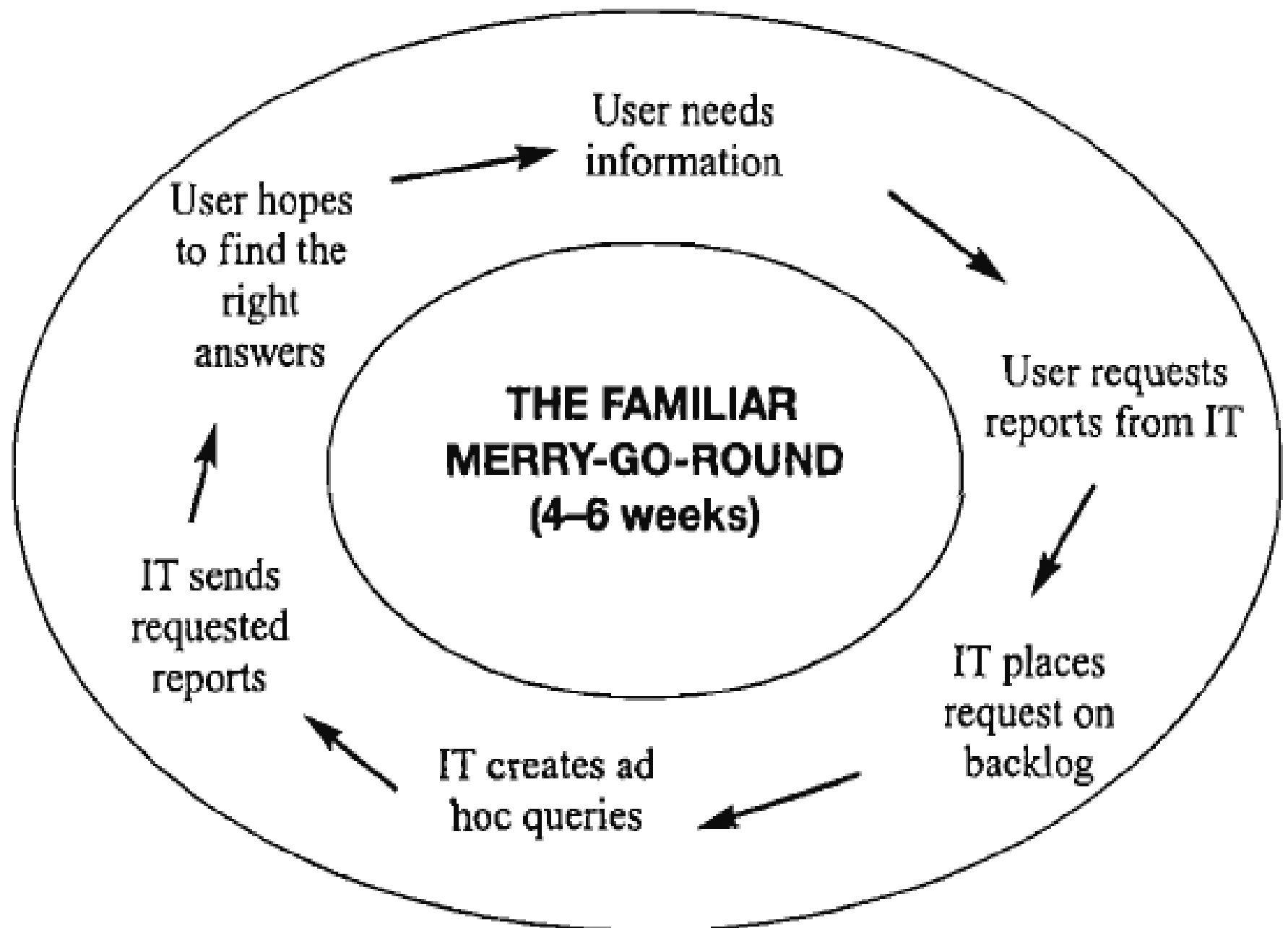
Past Decision-Support Systems

- ▶ Ad Hoc Reports
- ▶ Special Extract Programs
- ▶ Small Applications
- ▶ Information Centers
- ▶ Decision-Support Systems
- ▶ Executive Information Systems



Inability to Provide Information

- Every one of the past attempts at providing strategic information to decision makers was unsatisfactory
- Some of the factors relating to the inability to provide strategic information (Next Slide)
 - 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 not only 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 making 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

▶ Here are some good questions:

- ▶ What is a basic reason for the failure of all the previous attempts by IT to provide strategic information?
- ▶ What has IT been doing all along?
- ▶ IT has attempted to derive Strategic Information from Operational Systems
- ▶ Operational systems such as order processing, inventory control, claims processing, outpatient billing, and so on are not designed or intended to provide strategic information.



What are Operational Systems?

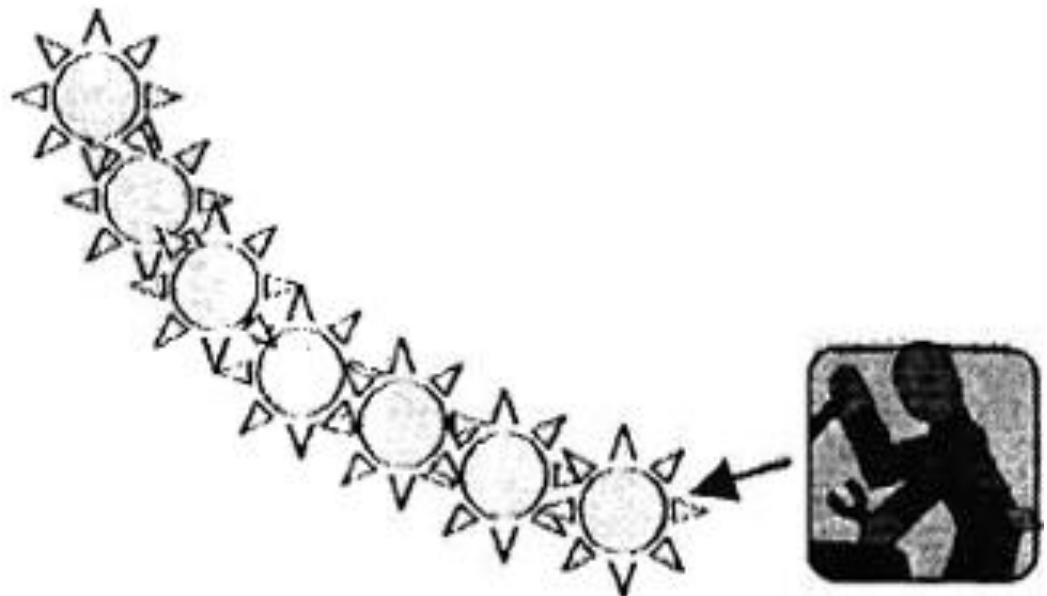
- ▶ Operational systems are online transaction processing (OLTP) systems
- ▶ They are the systems that are used to run the day-to-day core business of the company
- ▶ Operational systems make the wheels of business turn (see next slide)
- ▶ These systems typically get the data into the database
- ▶ Transactions process information about a single entity such as a single order, a single invoice, or a single customer


Operational Systems

Get the data in

Making the wheels of business turn

- ◆ Take an order
- ◆ Process a claim
- ◆ Make a shipment
- ◆ Generate an invoice
- ◆ Receive cash
- ◆ Reserve an airline seat



A decorative image in the top-left corner featuring several dice of different colors (red, white, black) and sizes, some showing different faces.

If not Operational, what are Decision Support Systems?

- 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
- They give management the ability to make strategic decisions to improve the business (see next slide).
- 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
- Again, Decision-support systems are developed to provide strategic information.

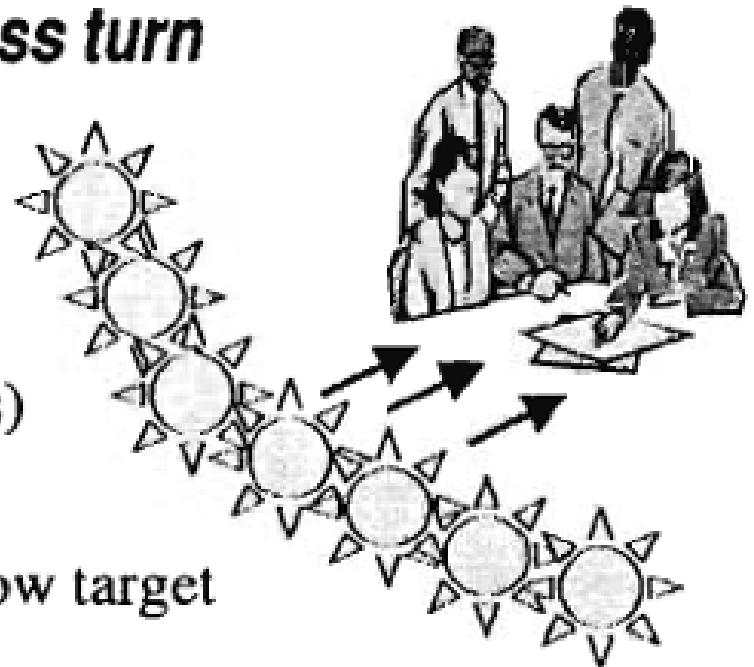
A collection of several dice in various colors (red, yellow, black, white) and sizes, some showing different faces.

Decision Support Systems

Get the information out

Watching the wheels of business turn

- ◆ Show me the top-selling products
- ◆ Show me the problem regions
- ◆ Tell me why (drill down)
- ◆ Let me see other data (drill across)
- ◆ Show the highest margins
- ◆ Alert me when a district sells below target



A decorative image in the top-left corner featuring several dice of different colors (red, white, black) and sizes, some showing different faces.

The Type of Informational Systems that need to be Designed

- ▶ 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



Let's Look at the Difference between Operational and Informational

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

How r they different ?

Operational DATA	Informational DATA
Holds current data	Holds historic data
Data is dynamic	Data is largely static
Read/Write accesses	Read only accesses
Repetitive processing	Adhoc complex queries
Transaction driven	Analysis driven
Application oriented	Subject oriented
Used by clerical staff for day-to-day operations	Used by top managers for analysis
Normalized data model (ER model)	Denormalized data model (Dimensional model)
Must be optimized for writes and small queries.	Must be optimized for queries involving a large portion of the warehouse.

A decorative graphic in the top-left corner featuring several dice: a red one, a white one with black pips, a black one with white pips, and a small white one with black pips.

DATA WAREHOUSING-THE ONLY VIABLE SOLUTION

- The desired features of this 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

A decorative image in the top-left corner featuring several dice of different colors (red, white, black) and sizes, some showing different faces.

Processing Requirements in this new Environment

▶ Most of the processing in the new environment for strategic information will have to be analytical. There are 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 Spot historical trends and apply them for future results

A decorative graphic in the top-left corner featuring several dice. There is a large white die with black pips, a smaller black die, a red die, and a small white die.

Business Intelligence at the Data Warehouse

- ▶ This new system environment that users desperately need to obtain strategic information happens to be the new paradigm of data warehousing
- ▶ This new environment is kept separate from the system environment supporting the day-to-day operations
- ▶ The data warehouse essentially holds the business intelligence for the enterprise to enable strategic decision making.
- ▶ The data warehouse is the only viable solution

A decorative image in the top-left corner featuring several dice of different colors (red, white, black) and sizes, some showing different faces.

Where does the Data Warehouse get It's Data?

- ▶ Might contain units of sales, by product, day, customer group, sales district, sales region, and promotion
- ▶ The data is derived from the operational systems that support the basic business processes of the organization
- ▶ 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

- Let's look at the Author's 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 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 decorative graphic in the top-left corner featuring several dice of different colors (red, white, black) and sizes, some showing different faces.

Data Ware Housing is 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

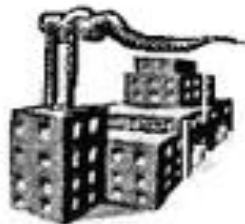


Data Ware Housing is an Environment – Not a Product (Cont'd)

- In Summary:
 - An ideal environment for data analysis and decision support
 - Fluid, flexible, and interactive
 - 100 percent user-driven
 - Very responsive and conducive to the ask-answer-ask-again pattern
 - Provides the ability to discover answers to complex, unpredictable questions
- Although a simple concept, it involves different functions: data extraction, the function of loading the data, transforming the data, storing the data, and providing user interfaces. (see next slide)

The data warehouse: a blend of technologies

OPERATIONAL SYSTEMS



Basic business processes

Extraction, cleansing, aggregation

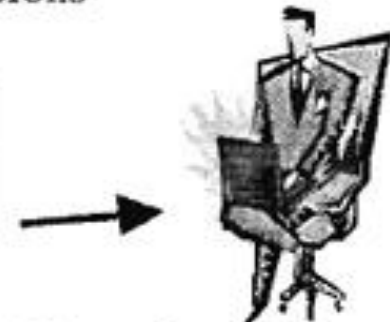


Data Transformation

Key measurements, business dimensions

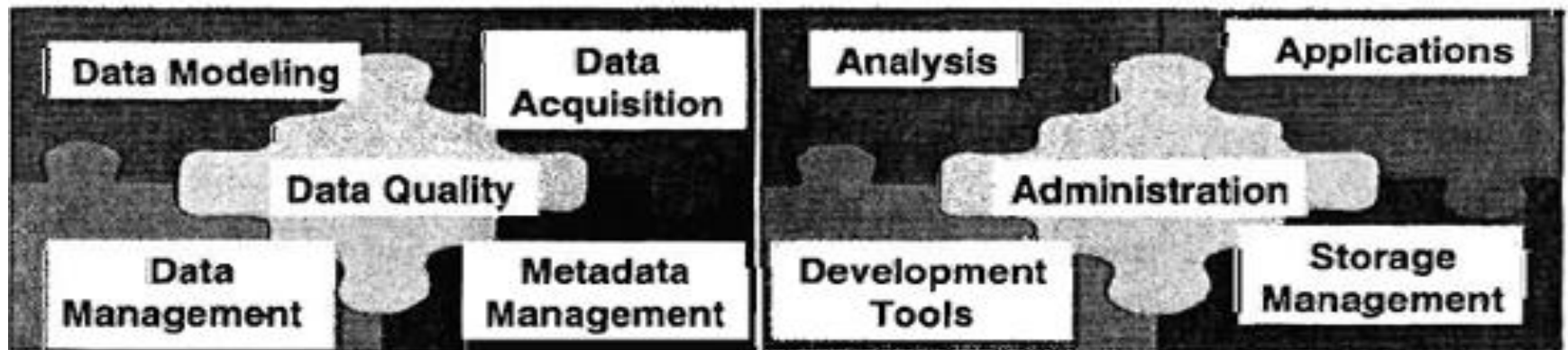


DATA WAREHOUSE



Executives/Managers/Analysts

BLEND OF TECHNOLOGIES





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 for every enterprise.



QUESTIONS??????
