



The Compelling Need for Data Warehousing

Learning Outcomes

- 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

Topics Covered

- Escalating Need For Strategic Information
- Failures Of Past Decision-Support Systems
- Operational versus Decision-Support Systems
- Data Warehousing – The Only Viable Solution
- Data Warehouse Defined

- Applications that run the business: Order processing, general ledger, inventory, in-patient billing, checking accounts, insurance claims, etc.
 - Businesses grow more complex, corporations spread globally, and competition became fiercer.
- Kilobytes ■ As business grows, so do systems & data.
Megabytes ■ Business executives need (different kinds of)
Gigabytes information to make (strategic) decision.
Terabytes
Petabytes
Exabytes
Zetabytes
Yottabytes
Brontobytes
Geabyte

Organizations achieve competitive advantage:

- ◆ 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

Figure 1-1 Organizations' use of data warehousing.

Successful application of DW in the fields of ...

- Trade – sales and claims analyses, shipment and inventory control, customer care and public relations
- Craftsmanship – production cost control, supplier and order support
- Financial services – risk analysis and credit cards, fraud detection
- Transport industry – vehicle management
- Telecommunication services – call flow analysis and customer profile analysis
- Health care services – patient admission and discharge analysis and bookkeeping in accounts departments

ESCALATING NEED FOR STRATEGIC INFORMATION

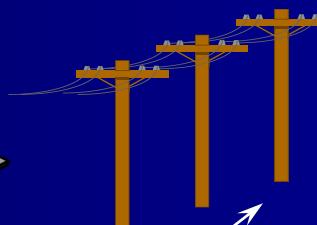
- Where's our next branch/warehouse? Which product line to expand? > strategic decision
- Increase customer base by 15 % over the next 5 years
- Increase sales by 15% in the North East Division
- Bring three new products to market in 2 years
- Enhance customer service level in shipments

Problem: Heterogeneous Information Sources

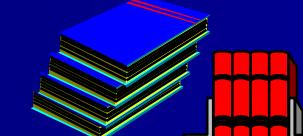
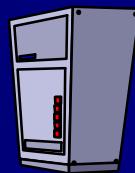
"Heterogeneities are everywhere"



Personal
Databases



Scientific Databases



Digital Libraries

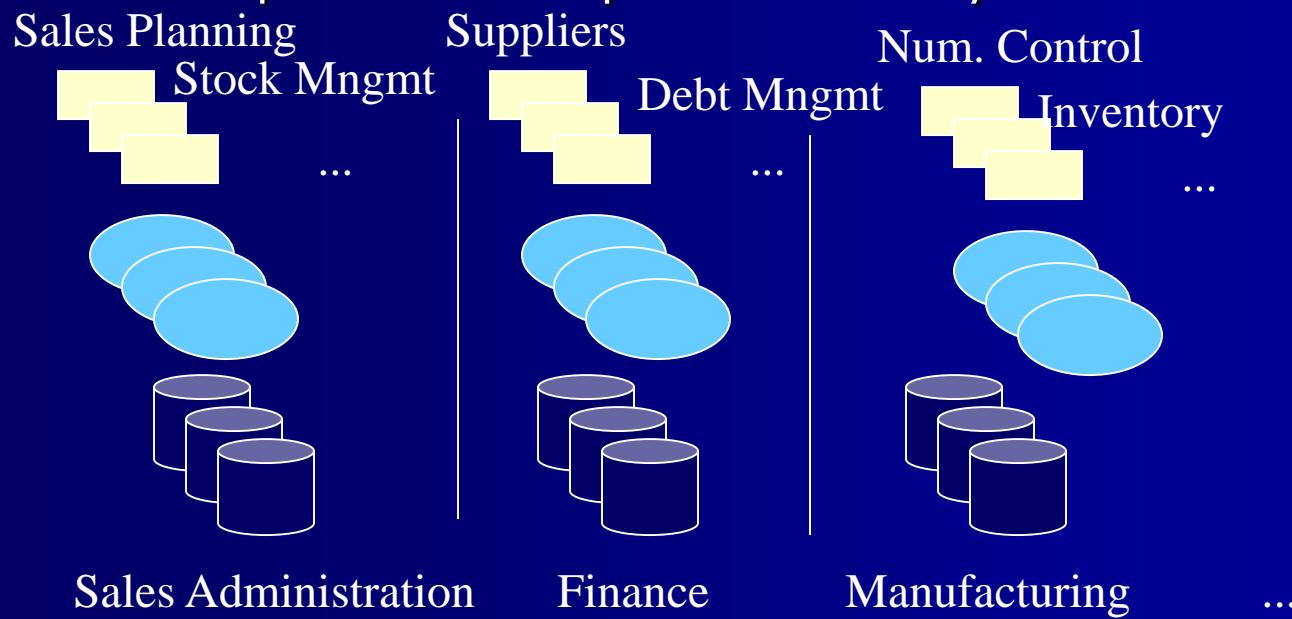


World
Wide
Web

- Different interfaces
- Different data representations
- Duplicate and inconsistent information

Problem: Data Management in Large Enterprises

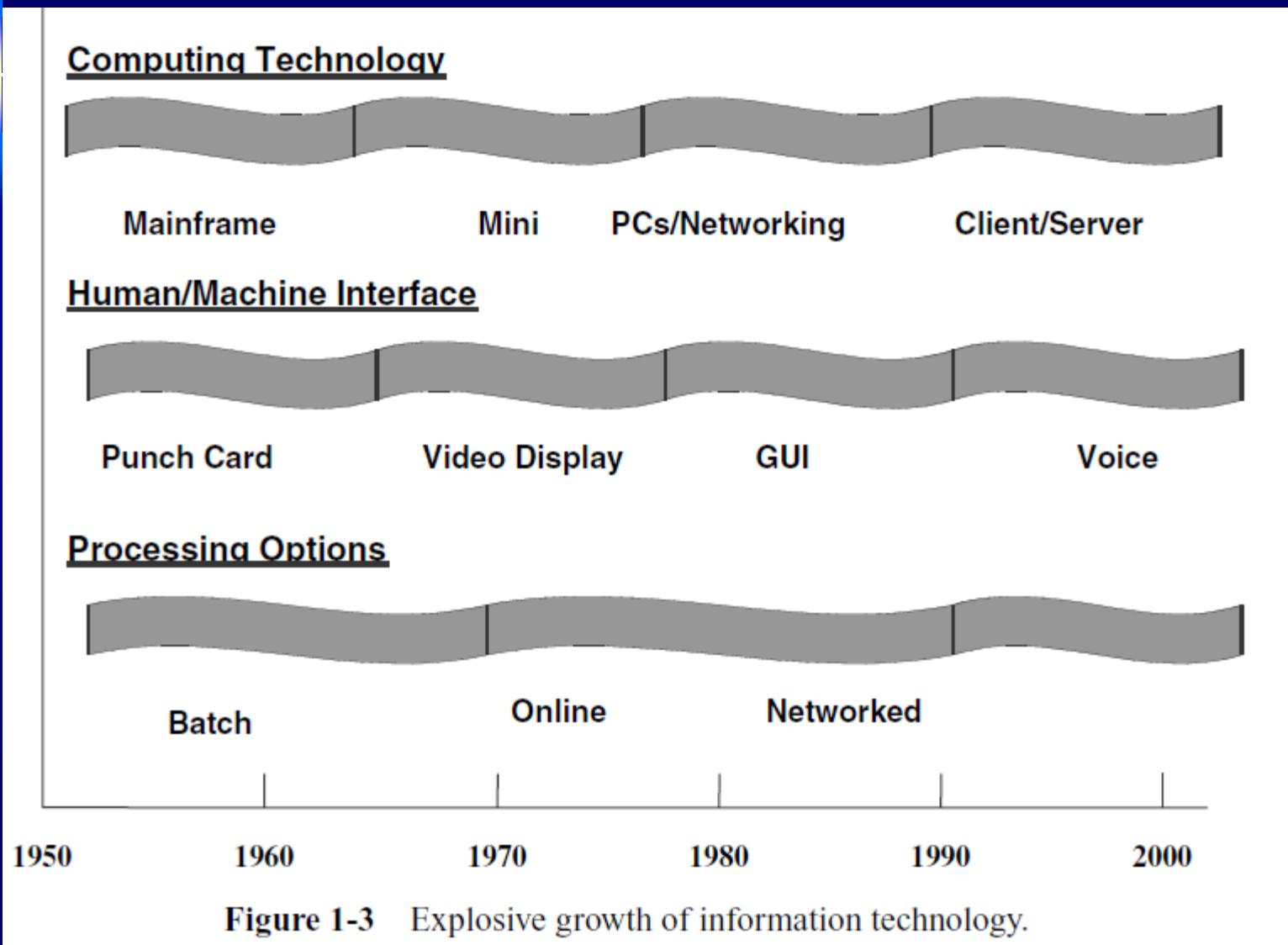
- Vertical fragmentation of informational systems (vertical stove pipes)
- Result of application (user)-driven development of operational systems



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.

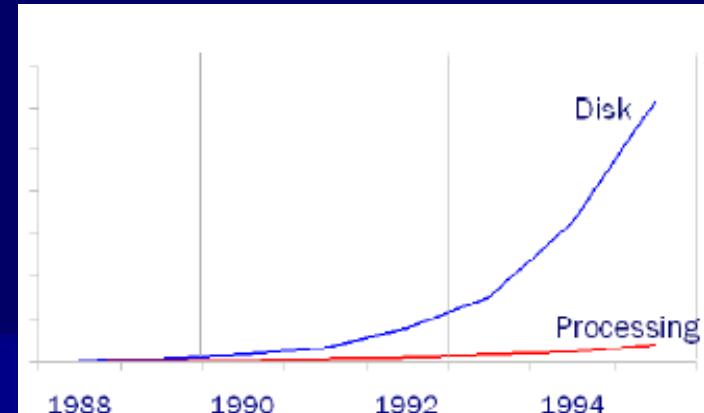
Technology Trends



Examples

- Europe's Very Long Baseline Interferometry (VLBI) has 16 telescopes, each of which produces 1 Gigabit/second of astronomical data over a 25-day observation session
 - storage and analysis a big problem
- Walmart reported to have 24 Tera-byte DB
- AT&T handles billions of calls per day
 - data cannot be stored -- analysis is done on the fly

Growth Trends



- Moore's law
 - Computer Speed doubles every 18 months
- Storage law
 - total storage doubles every 9 months
- Consequence
 - very little data will ever be looked at by a human
- Knowledge Discovery is **NEEDED** to make sense and use of data.

- "We have mountains of data in this company, but we can't access it."
- "We need to slice and dice the data every which way."
- "You've got to make it easy for business people to get at the data directly."
- "Just show me what is important."
- "It drives me crazy to have two people present the same business metrics at a meeting, but with different numbers."
- "We want people to use information to support more fact-based decision making."

Opportunities and Risks

- 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.

Opportunities and Risks

- 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.

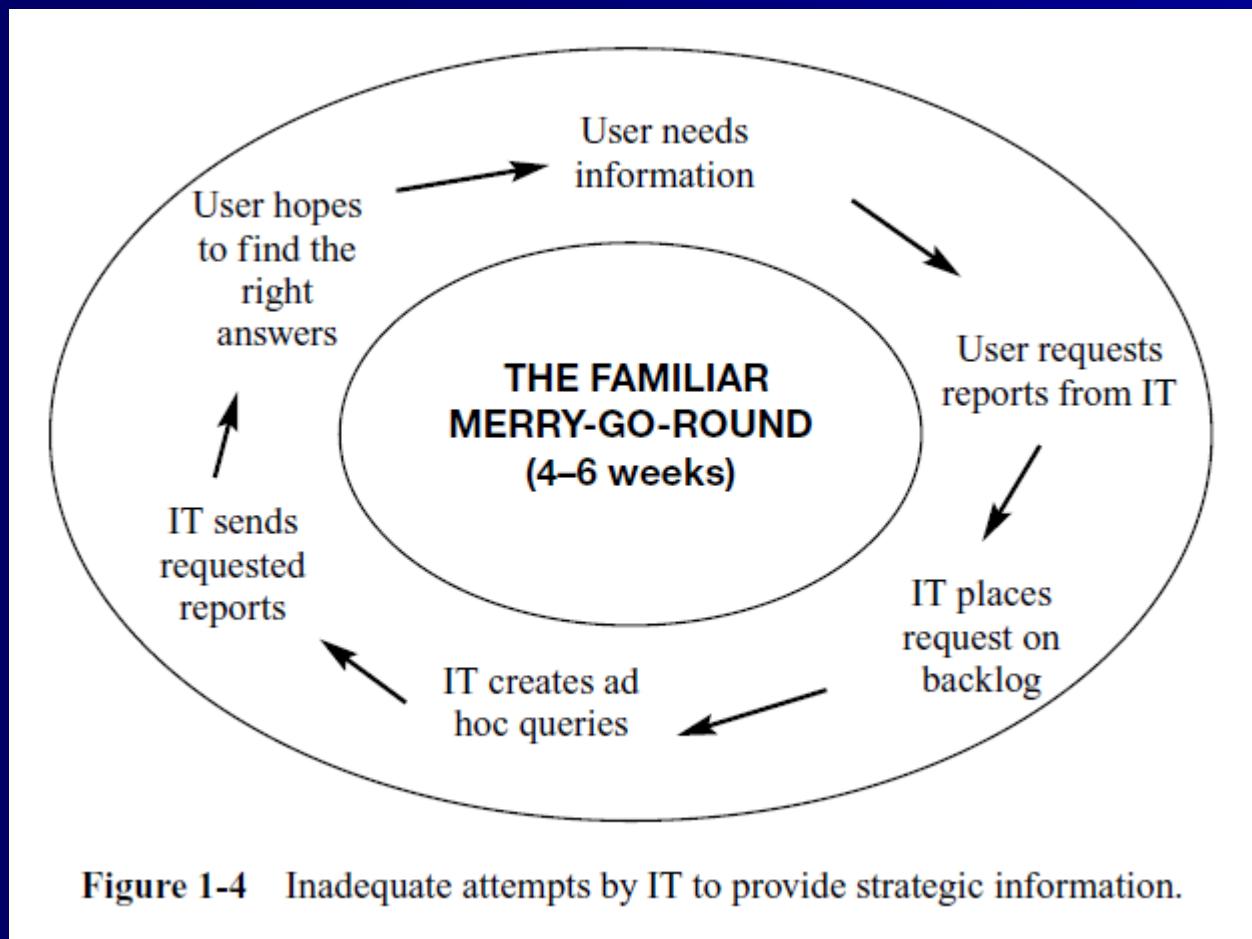
FAILURES OF PAST DECISION-SUPPORT SYSTEMS

- Ad-hoc reports -> special extract programs -> small applications -> DSS -> EIS
- 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.

FAILURES OF PAST DECISION-SUPPORT SYSTEMS..2

- 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.

FAILURES OF PAST DECISION-SUPPORT SYSTEMS..3



OPERATIONAL VERSUS DECISION-SUPPORT 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

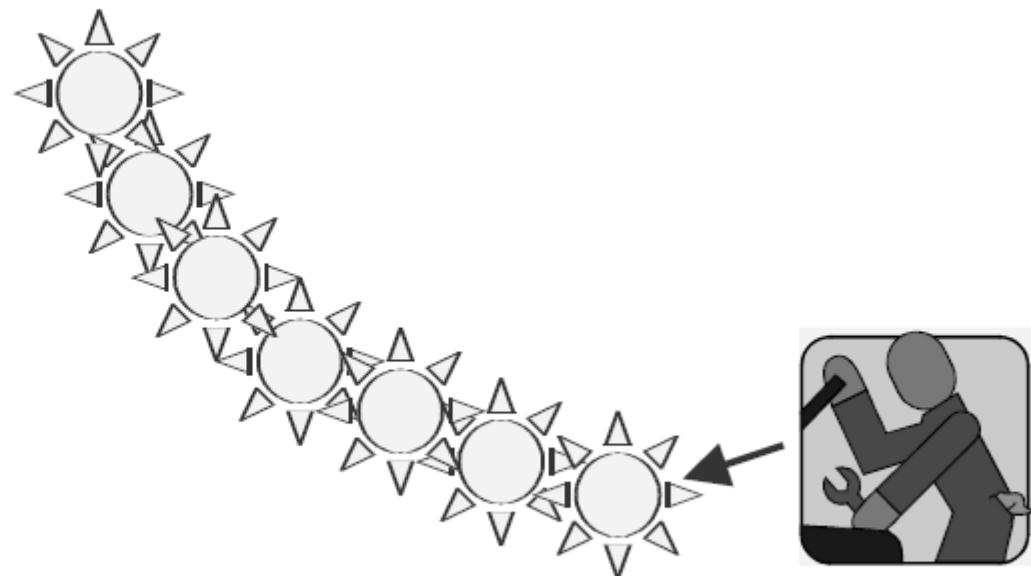


Figure 1-5 Operational systems.

OPERATIONAL VERSUS 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

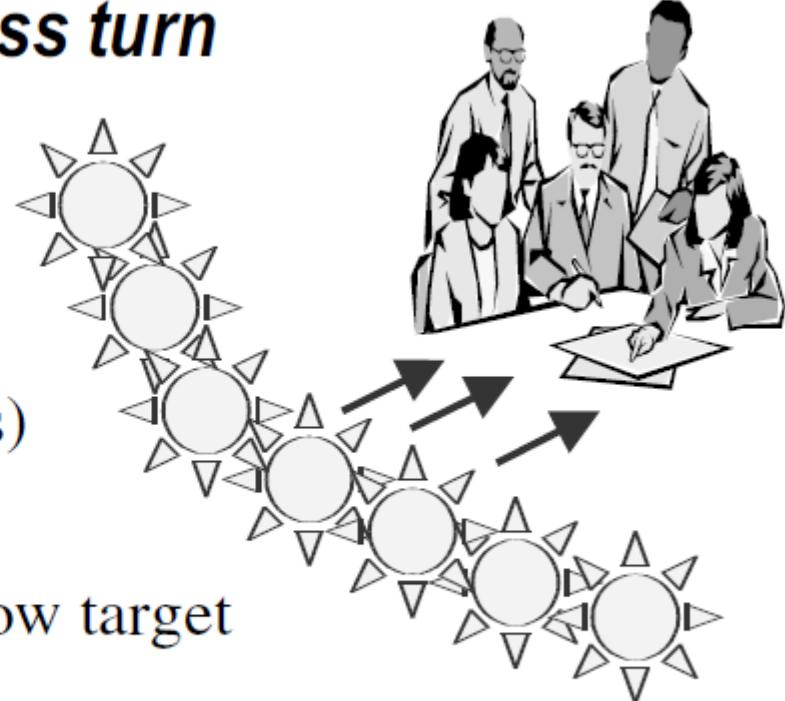


Figure 1-6 Decision-support systems.

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.

DATA WAREHOUSING—THE ONLY VIABLE SOLUTION

Desired features (system environment):

- 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

DATA WAREHOUSING—THE ONLY VIABLE SOLUTION..2

Processing requirements:

- Running of simple queries and reports against current and historical data
- Ability to perform “what if ” analysis in many different ways
- Ability to query, step back, analyze, and then continue the process to any desired length
- Spot historical trends and apply them for future results

OPERATIONAL SYSTEMS



Basic
business
processes

Extraction,
cleansing,
aggregation



Data Transformation

Key measurements,
business dimensions

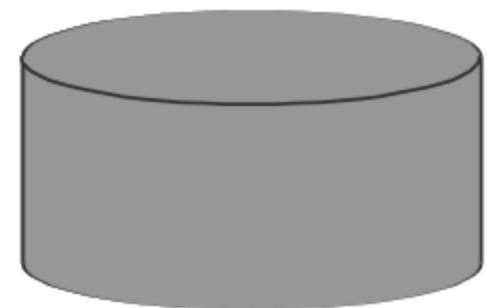


Figure 1-8 Business intelligence at the data warehouse.

DATA WAREHOUSE DEFINED

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

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.

CHAPTER SUMMARY

- 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.

REVIEW QUESTIONS

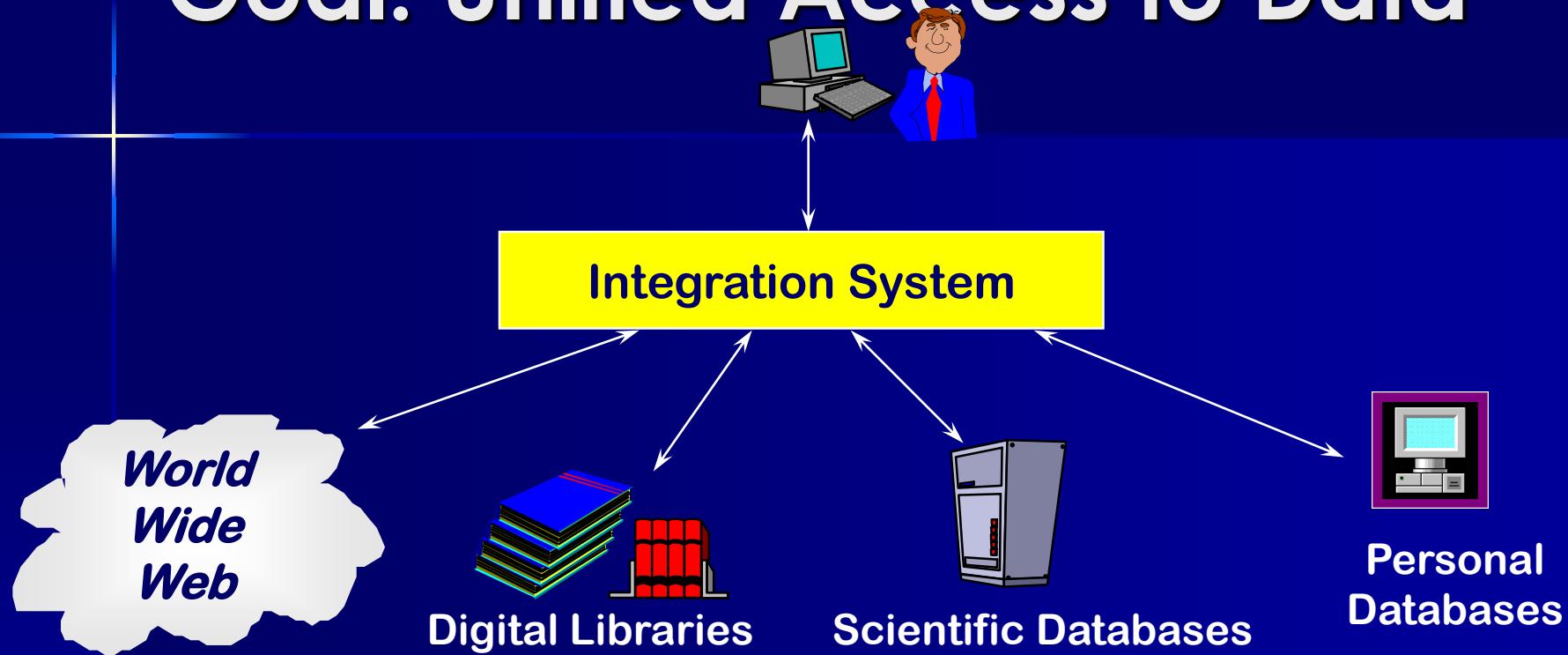
1. What do we mean by strategic information? For a commercial bank, name five strategic objectives.
2. Do you agree that a typical retail store collects huge volumes of data through its operational systems? Name three types of transaction data likely to be collected by a retail store in large volumes 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. Describe the types of processing that take place in a data warehouse.

- 
4. Describe the types of processing take place in a data warehouse?

REFERENCES

- Ponniah, P. (2001) Data Warehousing Fundamentals. New York: John Wiley & Sons.
- Galfarelli, M. & Rizzi, S. (2009) Data Warehouse Design: Modern Principles and Methodologies. McGraw-Hill.
- SIMS 257: Database Management, School of Information Management and Systems, UC Berkeley. Fall 2005

Goal: Unified Access to Data

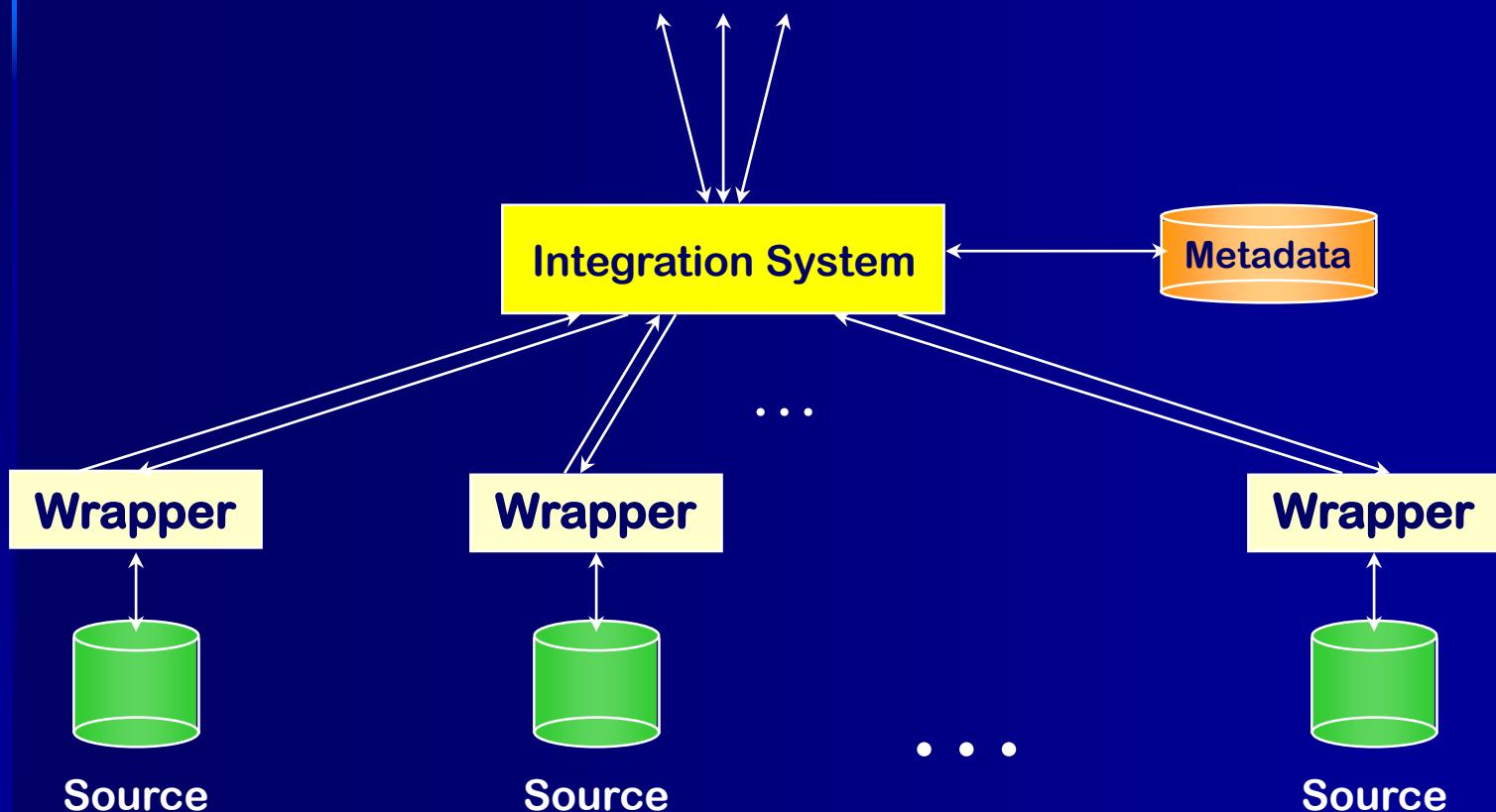


- Collects and combines information
- Provides integrated view, uniform user interface
- Supports sharing

The Traditional Approach

Query-driven (lazy, on-demand)

Clients



The Warehousing Approach

Information integrated in advance

Stored in WH for direct querying and analysis

