



Introduction to Information Technology
Turban, Rainer and Potter
John Wiley & Sons, Inc.



Chapter 3

Data and Knowledge Management

Chapter Outline

- Data Management: A Critical Success Factor
- Data Warehousing
- Information and Knowledge Discovery with Business Intelligence
- Data Mining Concepts and Applications
- Data Visualization Technologies
- Web-Based Data Management Systems
- Introduction to Knowledge Management
- Information Technology in Knowledge Management

Learning Objectives

- Recognize the importance of data, managerial issues, and life cycle .
- Describe the sources of data and their collection
- Describe document management systems.
- Explain the operation of data warehousing and its role in decision support
- Describe information and knowledge discovery and business intelligence
- Understand the power and benefits of data mining.
- Describe data presentation methods, and explain geographical information systems, visual simulations, and virtual reality as decision support tools
- Recognize the role of the web in data management.
- Define knowledge and describe the different types of knowledge.
- Describe the technologies that can be utilized in a knowledge management system

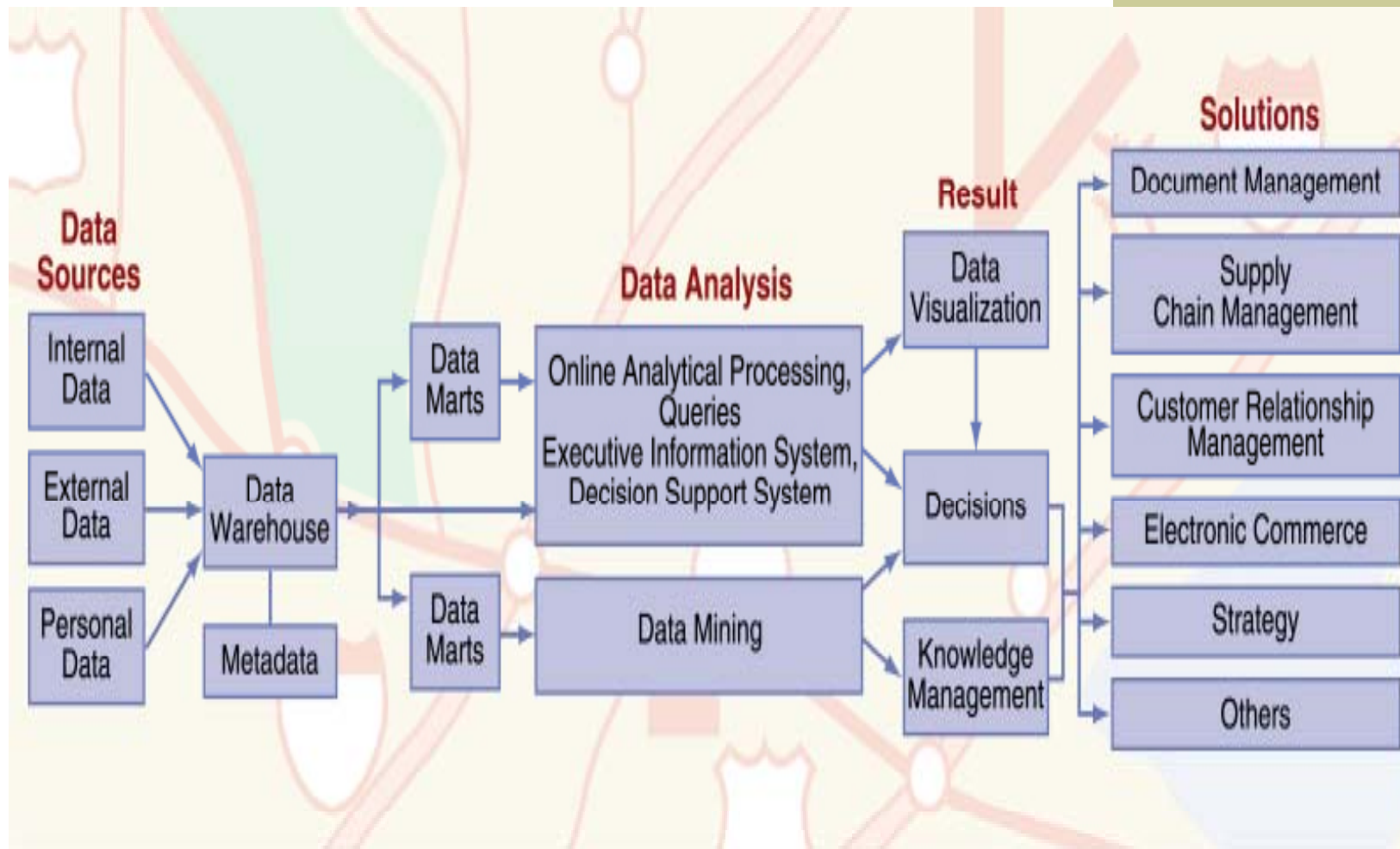
3.1 Data Management

- **A critical success factor:** IT applications cannot be done without using data. Data should be high-quality (accurate, complete, timely, consistent, accessible, relevant, and concise).
- **The Difficulties of managing Data:**
 - The amount of data increases exponentially with time
 - Data are scattered throughout organization and are collected by many individuals using several methods and devices.
 - An ever-increasing amount of external data needs to be considered in making organizational decisions.
 - Data security, quality, and integrity are critical, yet are easily jeopardized.

Critical Success Factors (CSF)

Those few things that must go right in to ensure an organization's survival and success

Data Life Cycle



Data Sources

- **Internal Data Sources:** data about people, products, services, and processes.
- **Personal Data:** IS users or other corporate employees may document their own expertise by creating personal data.
- **External Data Sources:** Data from commercial databases to sensors and satellites.

Document Management

- The automated control of electronic documents, page images, spreadsheets, word processing documents, and other complex documents through their entire life cycle within organization.
- The major tools of document management are workflow software, authoring tools, scanners, imaging systems, and database.
- **Document Management Systems (DMSs):** Computer systems that identify store, retrieve, track, and present information in an electronic format to decision makers

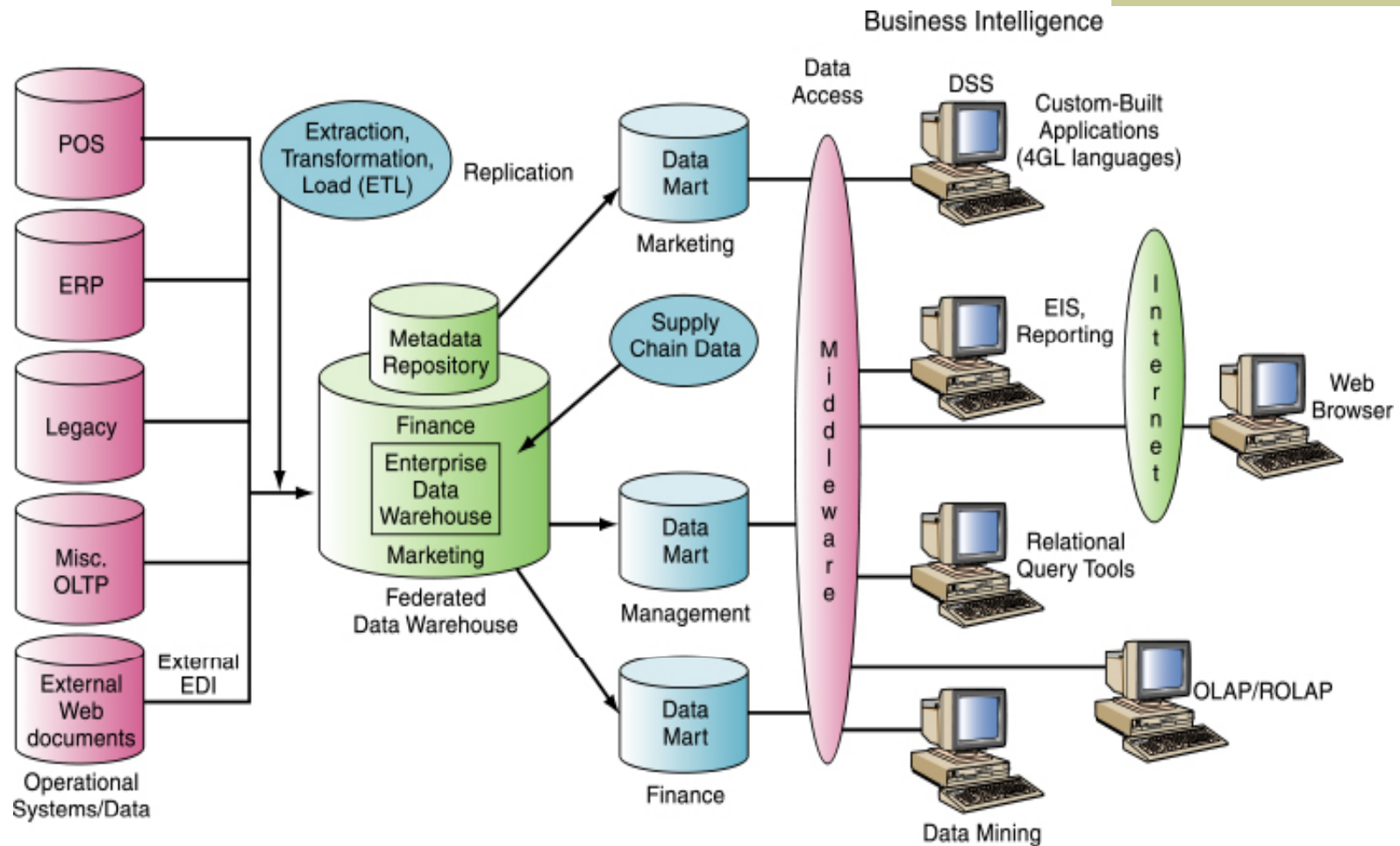
3.2 Data Warehousing

- **Transaction Processing:** The data are organized in hierarchical structure and centrally processed
- **Analytical Processing:** Analysis of accumulated data
- **Data Warehouse:** A repository of subject-oriented historical data that are organized to be accessible in a form readily acceptable for analytical processing.

Characteristics of a Data Warehouse

- **Organization.** Data are organized by subject and contain information relevant for decision support only .
- **Consistency.** Data in different operational databases may be encoded differently . In the data warehouse, though, they will be coded in a consistent manner.
- **Time variant.** The data are kept for many years so that they can be used for trends, forecasting, and comparisons over time.
- **Non-volatile.** Data are not updated once entered into the warehouse.
- **Multidimensional.** Typically the data warehouse uses a multidimensional structure .
- **Web-based.** Today's data warehouse are designed to provide an efficient computing environment for web-based applications.

Building a Data Warehouse



Relational and Multidimensional Database

- Relational databases store data in two – dimensional tables. Multidimensional databases typically store data in arrays, which consist of at least three business dimension.

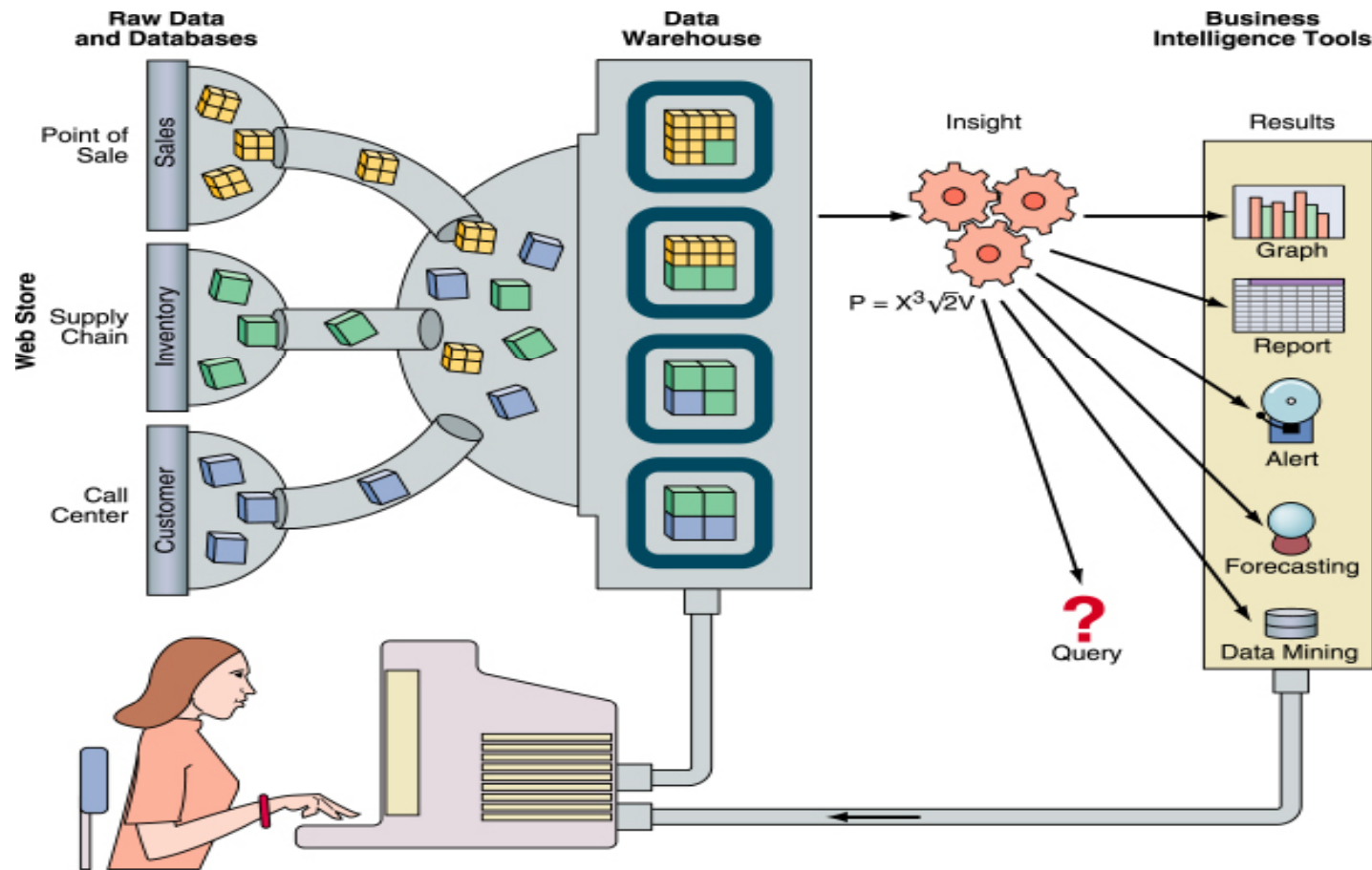
Data Marts

- ✿ **Data Mart:** A small data warehouse designed for a strategic business unit (SBU) or a department
- ✿ **The advantage of data marts include:** low cost (Prices under \$100,000 versus \$1million or more for data warehouses); significantly shorter lead time for implementation (often less than 90 days), local rather than central control (conferring power on the using group), More rapid response and more easily understood and navigated than an enterprise wide data warehouse .

3.3 Information & Knowledge Discovery with Business Intelligence

- **Business Intelligence:** A broad category of applications and techniques for gathering, storing, analyzing , and providing access to data to help enterprise users make better business and strategic decisions.

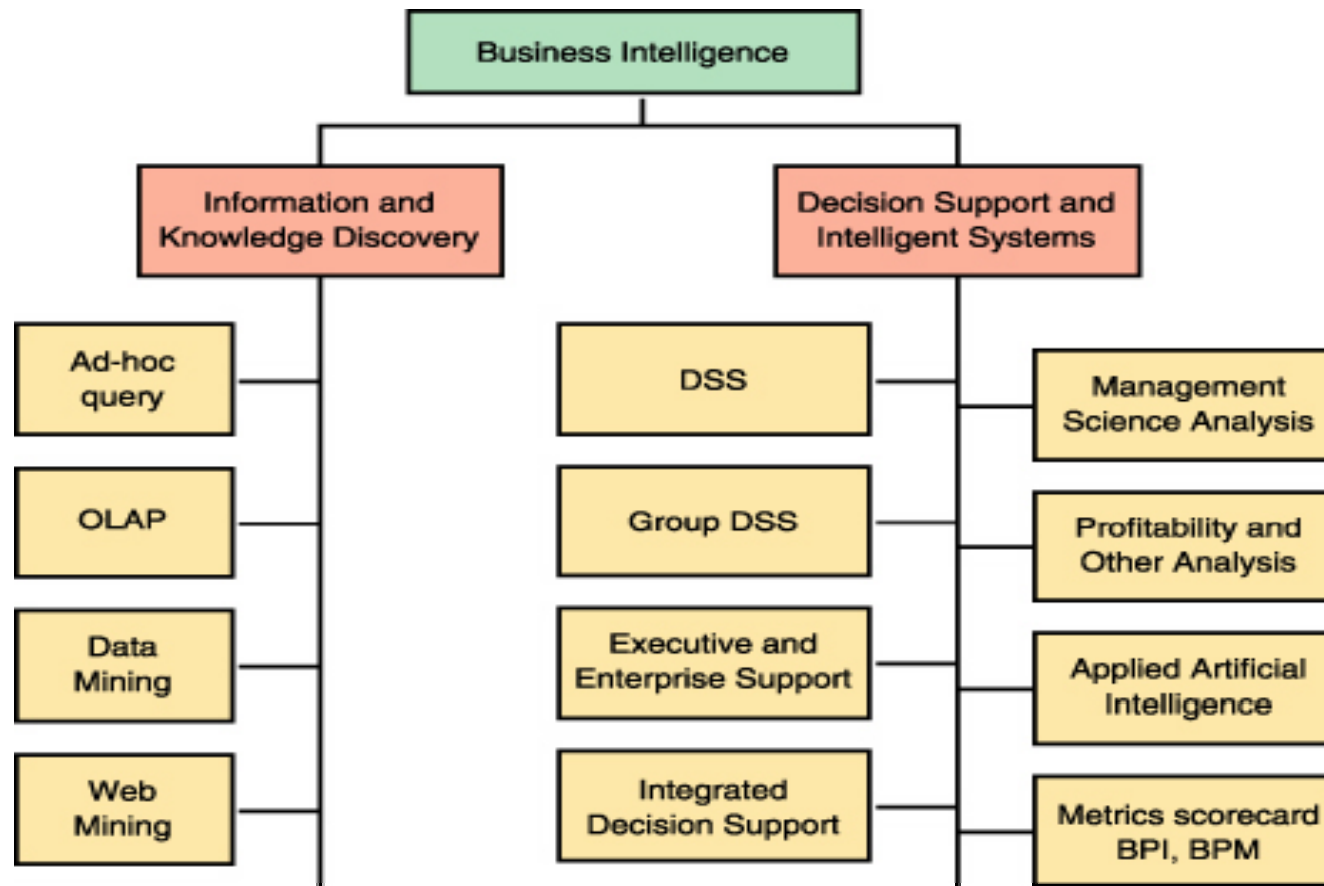
How Business Intelligence works?



The Tools and techniques of business intelligence

- The major application include the activities of query and reporting, online analytical processing, decision support , data mining, forecasting, and statistical analysis.
- BI tools are divided into two major categories:
 - (1) information and knowledge discovery
 - (2) decision support and intelligent analysis.

Categories of business intelligence



Knowledge Discovery (KD)

- The process of extracting knowledge from volumes of data; includes data mining .

Stage in the evolution of knowledge discovery

Evolutionary stage	Business question enabling technologies characteristic		
Data collection(1980s)	What was my total revenue in the last 5 years?	Computers ,tapes , disks	Retrospective , static data delivery
Data access (1980s)	What were unit sales in new England last March ?	Relational databases (RDBMS), structured query language (SQL)	Retrospective , dynamic data delivery at record level
Data warehousing and decision support (early 1990s)	What were the sales in region A by product , by salesperson?	OLAP, multidimensional databases, data warehouses	Retrospective , proactive data delivery at multiple level
Intelligent data mining (late 1990s)	What's likely to happen to the tBoston unit's sales next month ? Why?	Advanced algorithms, multiprocessor computers, massive databases	Prospective , proactive information delivery
Advanced intelligent systems; complete integration(2000-2004)	What is the best plan to follow? how did we perform compared to metrics?	Neural computing advanced al models, complex optimization, web services	Proactive , integrative ; multiple business partners

3.4 Data Mining Concepts

- ▶ **Data mining:** The process of searching for valuable business information in a large database, data warehouse, or data mart.
- ▶ **Data mining capabilities include:**
 - 1) Automated prediction of trends and behaviours, and
 - 2) Automated discovery of previously unknown patterns.

Data Mining Application

- Retailing and sales
- Banking
- Manufacturing and production
- Insurance
- Police work
- Health care
- Marketing

Text Mining

- The application of data mining to non-structured or less-structured text files.
- Text mining helps organizations to do the following (1) find the “hidden” content of documents, including additional useful relationship and (2) group documents by common themes (e.g., identify all the customers of an insurance firm who have similar complaints).

Web Mining

- The application of data mining techniques to discover actionable and meaningful patterns, profiles , and trends form web resources.
- Web mining is used in the following areas:
information filtering, surveillance, mining of web-access logs for analyzing usage, assisted browsing, and services that fight crime on the internet .
- Web mining can perform the following function :
 - Resource discovery
 - Information extraction
 - Generalization

3.5 Data Visualization Technologies

- Data Visualization: Visual presentation of data by technologies such as graphics, multidimensional tables and graphs, videos and animation, and other multimedia formats.

Geographical Information System (GIS)

- A computer- based system for capturing, storing, checking, integrating, manipulating, and displaying data using digitized maps.

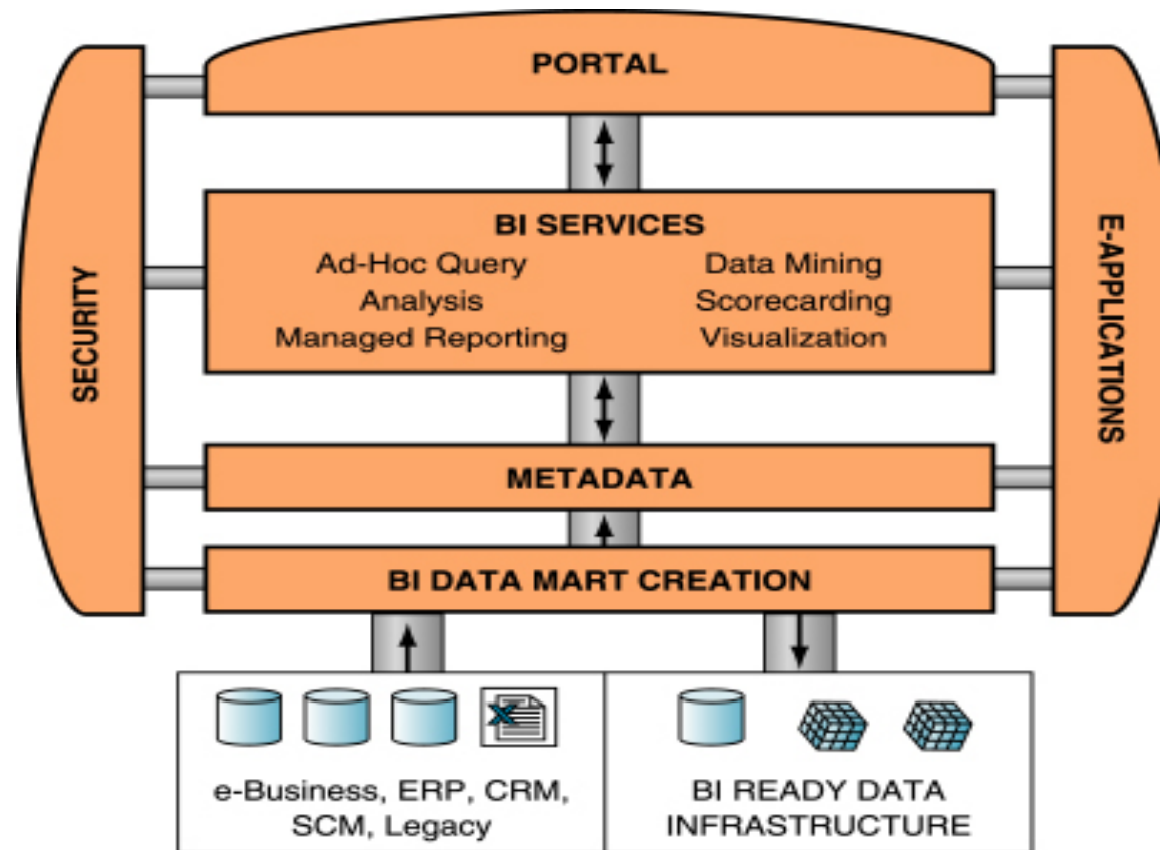
Visual Interactive Model and Simulation

- **Visual Interactive Modeling (VIM):** The use of computer graphic displays to represent the impact of different management or operational decisions on goals such as profit or market share.
- **Visual Interactive Simulation (VIS):** A visual interactive modeling method in which the end user watches the progress of the simulation model in an animated form, using graphics terminals.

Virtual Reality (VR)

- Interactive, computer-generated, three-dimensional graphics delivered to the user through a head-mounted display.
- **Virtual reality and the web.** A platform-independent standard for VR called virtual reality mark up language (VRML) makes navigation through online supermarkets, museums, and stores as interacting with textual information.

3.6 Web-based Data Management System



3.7 Knowledge Management

- **Knowledge**: Information that is contextual, relevant, and actionable .
- Intellectual capital (**intellectual assets**): other terms for knowledge.

Knowledge Management (KM)

- A process that helps organizations identify, select, organize, disseminate, transfer, and apply information and expertise that are part of the organization's memory and that typically reside within the organization in an unstructured manner

Chief Knowledge Officer (CKO)

- Executive whose objectives are to maximize the firm's knowledge assets, design and implement knowledge management strategies, and effectively exchange knowledge asset internally and externally .
- **Community of practice:** A group of people in an organization with a common professional interest.

Knowledge Management cont...

- **Explicit Knowledge:** The more objective, rational, and technical types of knowledge
- **Tacit knowledge:** The cumulative store of subjective or experiential learning; it is highly personal and hard to formalize.
- **Knowledge management systems (KMSs):** Information technologies used to systematize, enhance, and expedite intra- and interfirm knowledge management.

The Knowledge Management System Cycle

- ✦ **Create knowledge** . Knowledge is created as people determine new ways of doing thing or develop know-how. Sometimes external knowledge is brought in.
- ✦ **Capture knowledge**. New knowledge must be identified as valuable and be represented in a reasonable way.
- ✦ **Refine knowledge**. New knowledge must be placed in context so that it is actionable . This is where human insight (tacit qualities) must be captured along with explicit facts.
- ✦ **Store knowledge**. Useful knowledge must then be stored in a repository so that others in the organization can access it.
- ✦ **Manage knowledge**. Like a library, the knowledge must be kept current. It must be reviewed to verify that it is relevant and accurate.
- ✦ **Disseminate knowledge**. Knowledge must be made available in a useful format to anyone in the organization who needs it, anywhere and any time.

3.8 IT in Knowledge Management

- **Communication technologies:** allow users to access needed knowledge, and to communicate with each other- especially with experts .E-mail, the Internet, corporate intranets, and other web based tools provide communication capabilities.
- **Collaboration technologies:** provide the means to perform group work. Collaborative computing capabilities such as electronic brainstorming enhance group work especially for knowledge contribution .
- **Storage and retrieval technologies:** originally meant using a database management system to store and manage explicit knowledge . Electronic document management system and specialized storage system that are part of collaborative computing system are the tools used to capture, store, and manage tacit knowledge .

Technologies Supporting Knowledge Management

- **Artificial intelligence.** The study of human thought processes and the representation of those processes in machines.
- **Intelligent Agents.** Work and provide assistance in their daily tasks.
- **Knowledge Discovery in Databases.** A process used to search for and extract useful information from volumes of documents and data.

Seven Knowledge Management Tools

Tool	Description	Vendor/Product Examples
Collaboration computing	Groupware products; used to enhance tacit knowledge transfer within an organization	Group systems; Lotus Notes / Domino
Knowledge server	Contain the main knowledge management software, including the knowledge repository; provides access to other knowledge information , and data.	Hummingbird knowledge server; Autonomy's intelligent data operating layer (IDOL)
Enterprise knowledge portal	Presents a single access point into a knowledge management system ' organizes the sources of unstructured information in an organization .	Plum tree; Hyper wave
Electronic document management	Allows users to access needed documents over a corporate intranet; allows electronic collaboration on document creation and revision.	Doc Share; Lotus Notes
Knowledge –harvesting tools	Capture organizational knowledge unobtrusively; may be embedded in a knowledge management system.	Knowledge Mail ; Active Knowledge
Search engines	Locate and retrieve documents from vast collections in corporate repositories .	Google: Verity ; Inktomi
Knowledge management suites	Integrate communications, collaboration, and storage technologies in one complete, out-of- the- box solution	Web Sphere; knowledge X