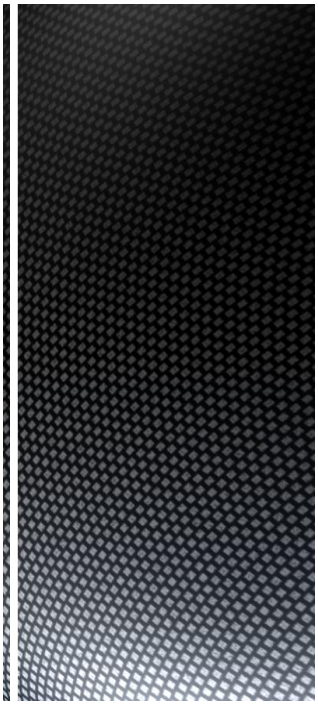
IT
(Information Technology)

Department of Industrial Engineering
Sharif University of Technology

Session# 6



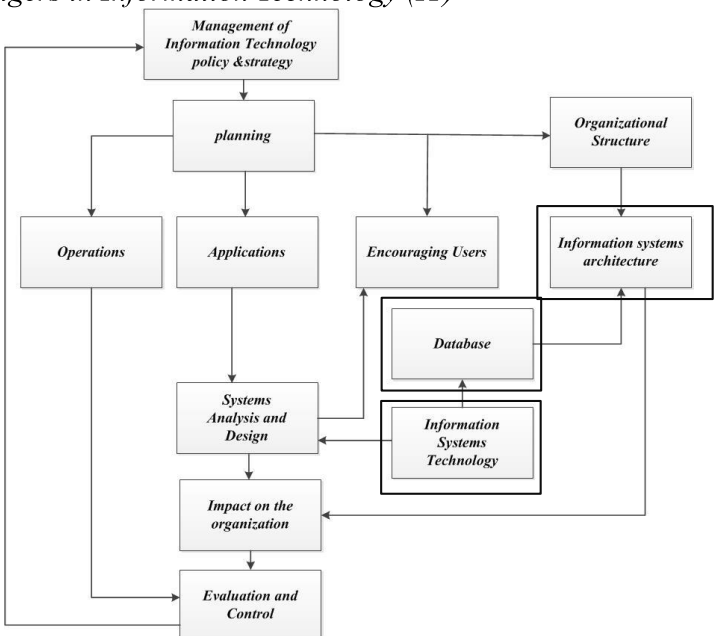
Course Description (Continued..)

- *Contents:*
  - *The role of managers in Information Technology (IT)* *(3 sessions)*
  - *Organizational Issues* *(3 sessions)*
  - *Information Technology* *(9 sessions)*
  - *Operational and enterprises systems* *(4 sessions)*
  - *Exciting directions in systems* *(3 sessions)*
  - *E-Business and E-Commerce* *(3 sessions)*
  - *Issues for senior management* *(2 sessions)*

Course Description *(Continued..)*

- *Contents:*
  - *Information Technology* *(9 sessions)*
    - *Fundamentals*
      - *The components of a personal computer*
    - *Software*
      - *Managerial concerns*
      - *The Contribution of Higher-Level languages*
      - *The Web Browser and Internet standards*
      - *The operating system*
    - *Database management*
      - *File elements*
      - *Enter database management software*
      - *Database in systems design*
      - *Data Warehouses, Data Marts, and Data Centers*
      - *Enterprise Content Management*

The role of managers in Information Technology (IT)

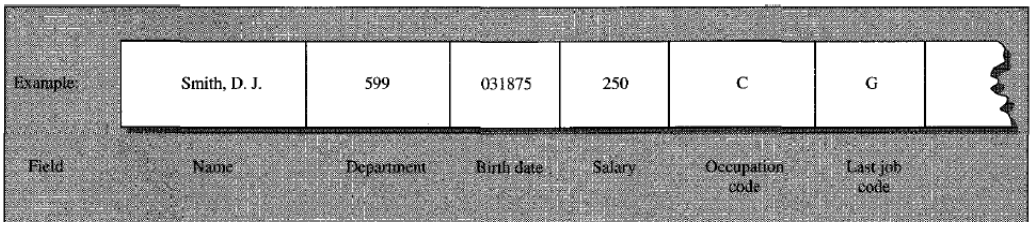


Information Technology

- *Database management*
  - *File Elements*
  - *Computers store data in a file, which can be defined simply as a collection of data.*
  - *A computer file is organized in a particular way with a well-defined structure for the information in the file.*
  - *A computer file consists of a collection of records, each of which is made up of fields. The various fields consist of groups of characters.*

Information Technology

- *Database management*
  - *Data*
  - *The smallest unit of storage is the byte, which consists of 8 bits.*
  - *This byte can represent numbers, characters, or parts of an image. The unit of interest in processing business data is the character, for example, the number 9 or the letter A.*
  - *A key to a record is a specific field of interest that will be used as a basis for storing and retrieving data.*

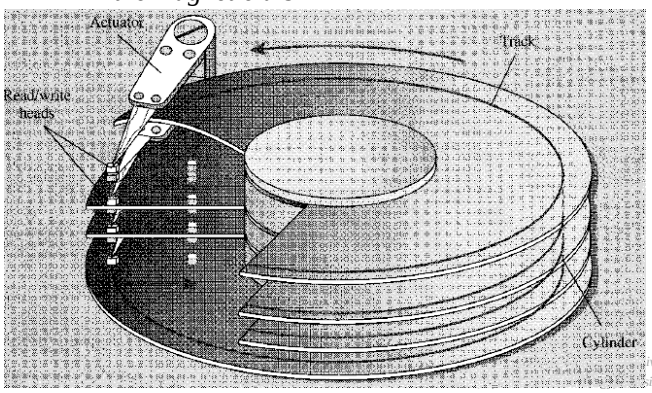


Information Technology

- *Database management*
  - *Direct-Access Files*
  - *There are two major types of files:*
    - *Sequential and*
    - *Direct access.*
  - *Sequential files were the first type of secondary storage. Records in this type of file are located one after another according to a given sequence.*
  - *On the average, if there are n records in the file, you will read n/2 records to find the one you are seeking.*
  - *A direct-access file uses a physical medium and programming, which facilitate the storage and retrieval of specific records.*

Information Technology

- *Database management*
  - *Storage Media*
  - *The most common device for storing direct-access files is the magnetic disk*

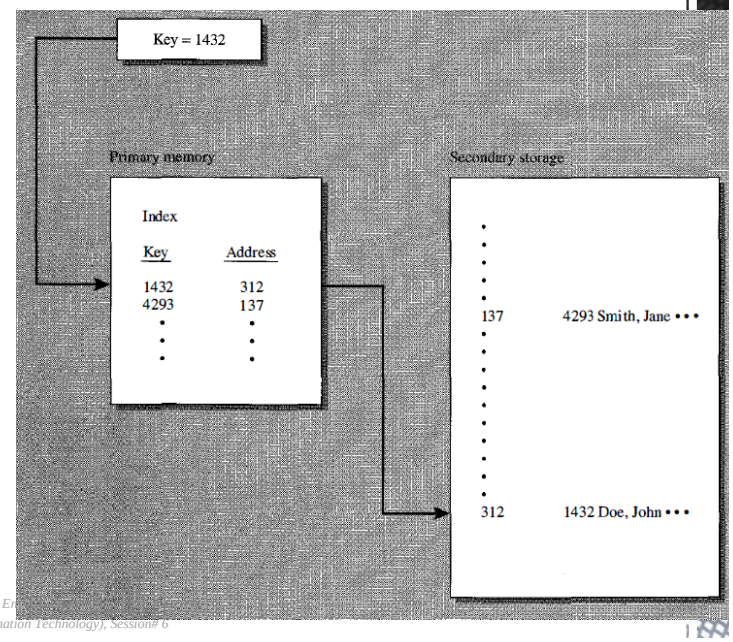


Information Technology

- *Database management*
  - *Storage Media*
  - *The total access time to read or write is made up of two components:*
    - *Seek time and*
    - *Rotational-delay time.*
  - *Seek time is the time needed to move the read-write heads from one position to another.*
  - *Rotational delay occurs because the data we want may not be directly under the read-write heads, even though they are located over the correct track.*
  - *The total time for seek and rotational delay adds to the average access time for the disk, usually 10-20 milliseconds.*

Information Technology

- *Database management*
  - *Finding Data on the File*
  - *In a sequential file, finding the data you want is not too difficult, though it may be time-consuming. Each record is in a sequence, so you simply read the file until you get to the location of the record of interest*
  - *The major advantage of the direct-access file is, as its name implies, that you can locate any record in the file in roughly the same short (milliseconds) period.*



Information Technology

- *Database management*
  - *More Complex Access*

FILE EXAMPLE

Record	Part no.	Assembly	On hand	Vendor
1	4326	103	27	ACME
2	6742	607	51	JOHNSON
3	8137	12	100	DAWES
4	3218	103	13	FRAZIER
5	3762	607	43	ARMOR

- *To avoid this excess read time, we use a pointer-a piece of data whose value points to another record.*
- *This type of file structure is known as a linked list or a chained file.*

Information Technology

- *Database management*
  - *More Complex Access*

FILE EXAMPLE

Record	Part no.	Assembly	On hand	Vendor	Pointer
1	4326	103	27	ACME	4
2	6742	607	51	JOHNSON	5
3	8137	12	100	DAWES	13
4	3218	103	13	FRAZIER	42
5	3762	607	43	ARMOR	106

INDEX TO ASSEMBLIES

Assembly	Record
12	3
25	212
103	1
104	62
607	2

Information Technology

- *Database management*
  - *Database management software*
  - *In the 1960s, software vendors developed products called database management systems (DBMSs).*
  - *A DBMS has to provide:*
    - *A method for defining the contents of the database.*
    - *A way to describe relationships among data elements and records .*
    - *A mechanism to set up the database in the first place.*
    - *Ways to manipulate the data including:*
      - *Updating (adding, modifying, and/or deleting information).*
      - *Using complex criteria to retrieve selected data.*

Information Technology

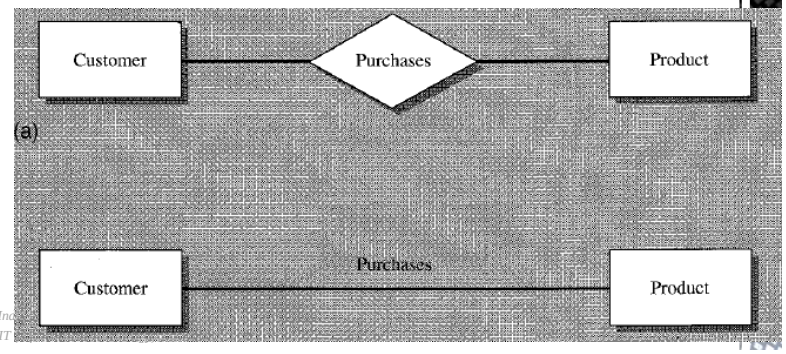
- *Database management*
  - *Relational Model*
  - *The relational model is the dominant structure for vendors writing DBMSs.*
  - *The underlying concept of a relational file system is very simple: Data are organized in two-dimensional tables.*
  - *The name of the model is derived from the fact that each table represents a relation.*
  - *Because different users see different sets of data and different relationships among them, it is necessary to extract subsets of the table columns for some users and to join tables together to form larger tables for others. The mathematics provides the basis for extracting some columns from the tables and for joining various columns.*

Information Technology

- *Database management*
  - *Normalization*
  - *One of the major tasks in designing a relational database is normalization.*
  - *The process of normalization ensures that there will not be any problems in updating the database and that operations on the various relations will not lead to inconsistent and incorrect data*
  - *First normal form requires that all occurrences of a record type contain the same number of fields.*
  - *Second and third normal forms require the examination of the relationship between key fields and other fields in the record.*
  - *In general, normalization creates a database in which there is minimum redundancy of data, and risks of damaging the database through updating are minimized.*

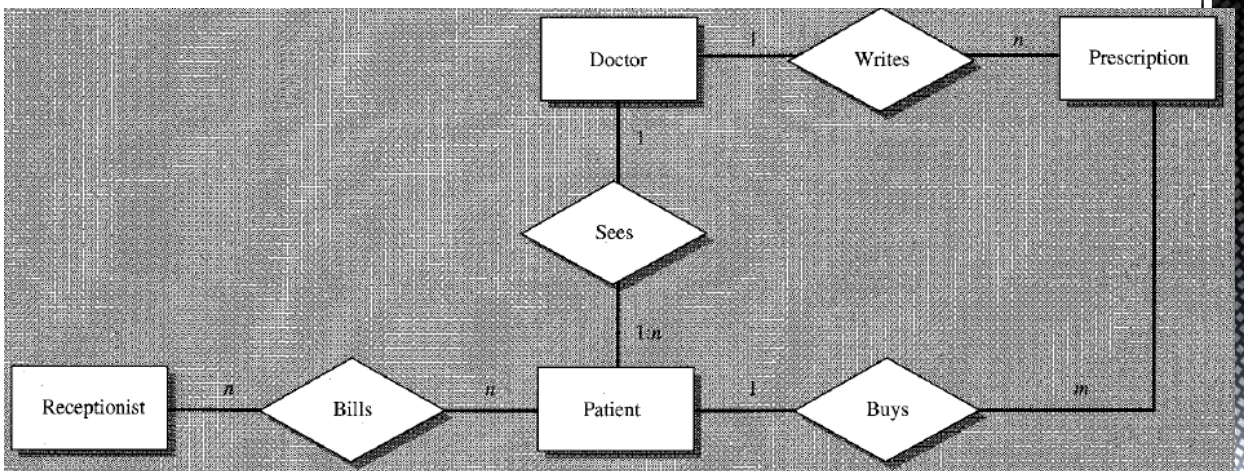
Information Technology

- *Database management*
  - *Data Modeling*
  - *A data model is useful for a number of reasons.*
    - *It helps us understand the relationships among different components in a systems design.*
    - *Data models show users more clearly how a system will function.*
  - *The most common type of data model is the entity-relationship (ER) diagram.*



Information Technology

- *Database management*
  - *Data Modeling ::ERD*



Information Technology

- *Database management*
  - *The Role of the Database Administrator*
  - *Many organizations using database software have created a new position known as the database administrator (DBA).*
  - *This individual is responsible for working with systems analysts and programmers to define the physical and logical views of the data to be manipulated by computers.*
  - *DBMSs in Building Systems*
  - *Database management systems are very popular packages for personal computers.*
  - *These packages feature friendly interfaces that make it easy for users to define the structure of relations and enter data.*

Information Technology

- *Database management*
  - *Distributed Databases*
  - *Organizations are building more distributed databases in which different parts of the database are located on different computers in a network.*
  - *This type of database raises a number of issues for the organization:*
    - *Will data be replicated across computers, or will there be only one copy?*
    - *If data are replicated, how frequently must different versions be updated to reflect changes?*
    - *How will updates to the database be coordinated so that its integrity is maintained?*
    - *Who "owns" distributed data, and who has access to it?*
    - *Distributed databases offer users easier access to data at the cost of overall higher complexity of the system*

Information Technology

- *Database management*

Billions, Even Trillions of Bytes of Data

Bank of America regularly creates a database that consolidates 35 million records processed by separate computers handling checking, savings, and other routine transactions. The consolidated database has 800 billion characters of data. How does the bank use this information? Every day about 100,000 customers call the bank to check a balance, challenge a charge on a credit card, or ask about interest rates. The bank decided to try to sell them something when they call.

The way to accomplish this cross selling was to tailor the product to each customer's needs. For example, if you have been accidentally bouncing checks, maybe you would pay for overdraft protection. The consolidated database provides bank employees with incredible insights into customer behavior and preferences. Some companies call these applications "data warehouses" for obvious reasons.

Burlington Coat Factory depends on a 1.5 trillion byte database that runs on a cluster of eight superminicomputers from Sequent computer systems. Company managers use the warehouse to determine a variety of information, for example, what styles are selling best, how are different stores performing, where to open the next store and so on.

These data warehouses are good candidates for parallel computing—multiple processors working in parallel are powerful and cheap enough to perform analyses on billions of bytes of data. John Alden Life Insurance company has a warehouse with four years of detailed medical claims with extensive cross indexing, comprising some 150 billion bytes of data. The company figures that asking a question to compare hospital networks in Illinois and New Jersey on hip replacements would tie up a mainframe all night. A 24-processor IBM SP2 does this job in the "tens of minutes."

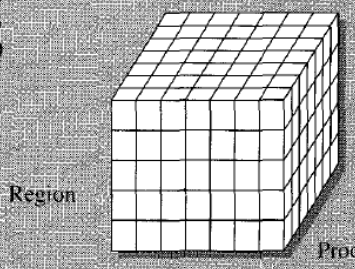
The computer can scan for information users request, or it can look for interesting relationships and patterns, a process called data mining. This kind of technology provides you with the ability to understand your customers and the nature of your business far better than in the past.

Information Technology

- *Database management*
  - *Data Warehouse*
  - *Businesses collect a operations*
  - *Usually, Instead of t multidimensional ar*
  - *To accommodate thi processing, firms of*
  - *You must define the based on those dime*

Rep	Date	Part	Quantity
Jones	September	A629	100
Smith	October	B206	50
Doe	September	A109	75
Jones	October	B656	150

Relational Table



Multidimensional Representation

Information Technology

- *Database management*
  - *Data Warehouse*
  - *One strategy for creating the "data cube" associated with a multidimensional database is to create a "fact cube" through an n-way crossing of all the dimensions specified when defining the database.*
  - *One objective of a data warehouse is to help you understand your business better. This kind of technology, then, helps create a "learning organization," an organization that is able to better understand its market, customers, and itself*

Information Technology

- *Database management*
  - *Data Mining*
  - *One of the reasons for building a data warehouse is to undertake data mining.*
  - *The idea is to look for interesting and important patterns in a huge database*
  - *Data mining is associated with **knowledge discovery** systems, applications that try to make sense out of data.*

Information Technology

- *Database management*

Virtual Databases

It has been estimated that about 90 percent of the data in the world is not in relational databases; it is scattered across Internet Web sites, legacy applications, and nonrelational databases. Virtual database technology makes external data sources a part of your corporate, relational database system. With virtual database software, you can make queries of data that are scattered over a variety of locations and stored in different ways.

To find a job for a marketing manager position in a company at a certain location where the company's sales are growing at 25 percent a year would require an extensive search of Internet job directories. A virtual database allows the user to make one query. The database consolidates job listings from many sources and develops indexes for job title, category, and other attributes of the listings. The programmer accesses the virtual database using standard SQL commands while the end user might work with a query form on the Web.

There are currently a number of virtual database applications on high-traffic Web sites such as Yahoo! and the Wall Street Journal Online. It is also possible to combine external data with a company's internal data warehouse to expand the range of data-mining and knowledge discovery programs. Virtual database technology is also very useful for electronic commerce. Junglee Corporation, a subsidiary of Amazon.com, uses a virtual database to integrate data from multiple merchants to give shoppers comparison data. One example is a virtual database that combines the contents of Amazon.com and Powell's Books with the *New York Times Book Review;* the result is a unified, relational schema with two tables: books and reviews. Junglee has also applied this technology for classified employment ads, real estate, and apartment listings.

Virtual database technology extends databases beyond the confines of the organization; it provides consistency in accessing data that originate in many different places and that is stored in different ways. The result is an extremely powerful technology for accessing the tremendous amount of data that exists in automated databases.

The role of managers in Information Technology (IT)

- *HW#*
  - *Data base Assignment*
  - *Study your reference book through pages 255 to 257. The book talks about the two case studies for database management systems.*
  - *Try to answer the questions in the case studies based on our described context in this session.*

- *The Home work should be sent to FValilai@sharif.edu*

Course Description *(Continued..)*

- *Contents:*
  - *Information Technology* *(9 sessions)*
    - *Fundamentals*
      - *The components of a personal computer*
    - *Software*
      - *Managerial concerns*
      - *The Contribution of Higher-Level languages*
      - *The Web Browser and Internet standards*
      - *The operating system*
    - *Database management*
      - *File elements*
      - *Enter database management software*
      - *Database in systems design*
      - *Data Warehouses, Data Marts, and Data Centers*
      - *Enterprise Content Management*