Database Management

Database Management Systems

- A database management system (DBMS), or database program, is software that allows you to create, access, and manage a database.

- DBMSs are available for many sizes and types of computers.
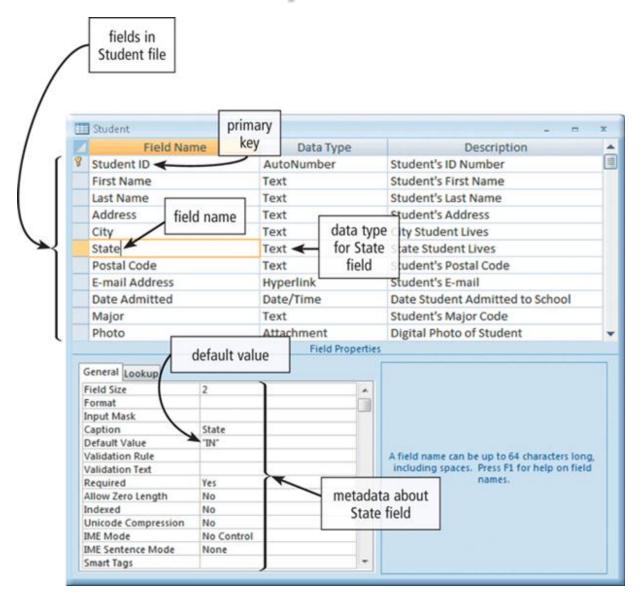
Database Management Systems

Popular Database Management Systems		
Database	**Manufacturer**	**Computer Type**
Access	Microsoft Corporation	Personal computer, server, mobile devices
Adabas	Software AG	Server, mainframe
D^3	Raining Data	Personal computer, server
DB2	IBM Corporation	Personal computer, server, mainframe
Essbase	Oracle Corporation	Personal computer, server, mobile devices
FastObjects	Versant Corporation	Personal computer, server
FileMaker	FileMaker, Inc.	Personal computer, server
GemFire	GemStone Systems	Server
Informix	IBM Corporation	Personal computer, server, mainframe
Ingres	Ingres Corporation	Personal computer, server, mainframe
InterBaseSMP	Embarcadero Technologies	Personal computer, server
KE Texpress	KE Software, Inc.	Personal computer, server
MySQL	MySQL AB	Personal computer, server
ObjectStore	Progress Software Corporation	Personal computer, server
Oracle Database	Oracle Corporation	Personal computer, server, mainframe, mobile devices
SQL Server	Microsoft Corporation	Server, personal computer
SQL Server Compact Edition	Microsoft Corporation	Mobile devices
Sybase	Sybase Inc.	Personal computer, server, mobile devices
Teradata Database	Teradata	Server
Versant	Versant Corporation	Personal computer, server
Visual FoxPro	Microsoft Corporation	Personal computer, server

Data Dictionary

- A **data dictionary**, sometimes called a *repository*, contains data about each field in those files.
- For each file it stores details such as the file name, description, the file's relationship to other files, and the number of records in the file.
- For each field, it stores the field name, description, field type, field size, default value, validation rules, and the field's relationship to other fields.

Data Dictionary

- Because the data dictionary contains details about data, some call it *metadata*, meta meaning more comprehensive.

- It is a crucial backbone to a DBMS so only skilled professionals should update its contents.

- It is used to perform validation checks and limit the type of data that can be entered.

Data Dictionary

- A data dictionary allows users to specify a *default value* for a field, which is a value that the DBMS initially displays in a field.
  - Ex. If most students live in Indiana, the default value could read 'IN'.
- The user doesn't have to type the default value which helps in reducing errors.

Data Dictionary

fields in Student file

primary key

Field Name	Data Type	Description
Student ID	AutoNumber	Student's ID Number
First Name	Text	Student's First Name
Last Name	Text	Student's Last Name
Address	Text	Student's Address
City	Text	City Student Lives
State	Text	State Student Lives
Postal Code	Text	Student's Postal Code
E-mail Address	Hyperlink	Student's E-mail
Date Admitted	Date/Time	Date Student Admitted to School
Major	Text	Student's Major Code
Photo	Attachment	Digital Photo of Student

field name

data type for State field

Field Properties

General Lookup

Field Size	2
Format	
Input Mask	
Caption	State
Default Value	"IN"
Validation Rule	
Validation Text	
Required	Yes
Allow Zero Length	No
Indexed	No
Unicode Compression	No
IME Mode	No Control
IME Sentence Mode	None
Smart Tags	

default value

A field name can be up to 64 characters long, including spaces. Press F1 for help on field names.

metadata about State field

File Retrieval and Maintenance

- A DBMS provides several tools that allow users and programs to retrieve and maintain data in the database (adding, modifying, and deleting).
- To retrieve or select data in a database, you **query** it, which is a request for specific data from the database.
- Users can instruct the DBMS to display, print, or store the results of a query, making it one of the more powerful database features.

Query Language

- A **query language** consists of simple, English-like statements that allow users to specify the data to display, print, or store.
- Each has its own grammar and vocabulary.
- A person without programming experience can learn a query language in a short time.
- Most queries are used to retrieve data.
- Some DBMSs provide wizards to guide users through the steps of creating a query.

Query by Example

- Most DBMSs include **query by example (QBE)**, a feature that has a graphical user interface to assist users with retrieving data.

Form

- A **form**, sometimes called a *data entry form*, is a window on the screen that provides areas for entering or modifying data in a database.
- Well-designed forms should validate data as it is entered.
- A form that sends entered data across a network or the Internet is called an *e-form*, short for *electronic form*.
- E-forms generally use a means to secure the data while it is transported across the network.

Form

Report Generator

- A **report generator**, also called a *report writer*, allows users to design a report on the screen, retrieve data into the report design, and then display or print the report.

- Unlike a form, report generators can only retrieve data.

- Report generators allow you to format the page and some allow you to create a report as a Web page.

Report Generator

Student List by Major

Major	Last Name	Student ID	First Name	Address	City	Date Admitted
BIO						
	Drake	3876	Louella	33 Timmons Place	Bonner	8/9/2009
CT						
	Ruiz	3928	Adelbert	99 Tenth Street	Sheldon	10/8/2009
GEN						
	Tu	2928	Benjamin	2204 Elm Court	Rowley	9/4/2009
SOC						
	Brewer	2295	Milton	54 Lucy Court	Charlestown	6/10/2009
	Gupta	4872	Elena	76 Ash Street	Rowley	9/3/2010

Data Security

- A DBMS provides means to ensure that only authorized users can access data at permitted times.

- Most DBMSs allow different levels of access privileges to be identified for each field in the database, defining the actions that a specific user or group can perform.

Data Security

- *Access privileges* for data involve establishing who can enter new data, modify existing data, delete unwanted data, and view data.
  - Ex. A student would have read-only privileges: allowing them to view the list of offered classes but not change them.
  - Ex. A department head would be able to modify the data.
  - Ex. Other users would have no access privileges to the data.

Data Security

- Many organizations adopt a *principle of least privilege* policy, where users' access privileges are limited to the lowest level necessary to perform required tasks.

Backup and Recovery

- Occasionally a database is damaged or destroyed, so a DBMS provides a variety of techniques to restore the database to a usable form.

- A **backup**, or copy, of the entire database should be made on a regular basis.

- Some DBMSs have a built-in backup utility while others require a separate utility.

Backup and Recovery

- More complex DBMSs maintain a **log**, which is a listing of activities that modify the contents of the database.
    - Ex. A registration clerk modifies a student's address, the change will appear in the log.
- The log contains:
    - A copy of the record prior to the change called the *before image*
    - The change being made
    - And a copy of the record after the change called the *after image*
    - The log may also store who made the change and when.

Backup and Recovery

- DBMSs that create logs usually provide a **recovery unit**, which uses the logs and/or backups to restore a database when it becomes damaged or destroyed using rollforward and rollback techniques.

- In a *rollforward*, or *forward recovery*, the DBMS uses the log to reenter changes made to the database since the last save or backup.

- In a *rollback*, or *backward recovery*, the DBMS uses the log to undo any changes made to the database during a certain period.

Backup and Recovery

- *Continuous backup* is a backup plan in which all data is backed up whenever a change is made.

- Continuous backup provides recovery in a matter of seconds.

- This technique can cost more but is growing in popularity due to its benefits.

Relational, Object-Oriented, and Multidimensional Databases

- Every database and DBMS is based on a specific data model.

- A **data model** consists of rules and standards that define how the database organizes data.
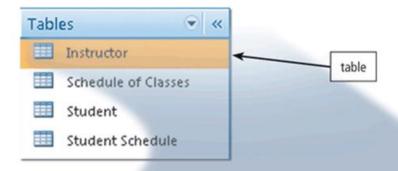
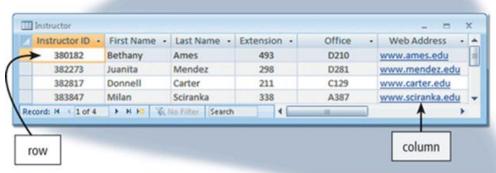# Relational, Object-Oriented, and Multidimensional Databases

- Three popular data models in use today are relational, object-oriented, and multidimensional.

- Some are called *object-relational databases* because they combine features of the relational and object-oriented data models.
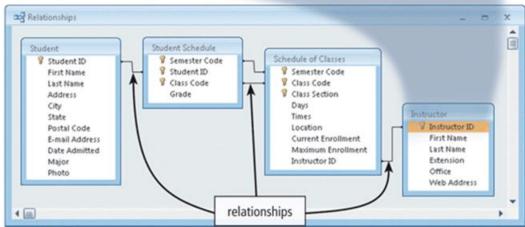
Data Models for Popular DBMSs					
Data Model	**Popular DBMSs**		**Data Model**	**Popular DBMSs**	
Relational	Access Adabas FileMaker Informix Ingres InterBaseSMP MySQL	SQL Server Sybase Teradata	Object-relational	DB2 Oracle Polyhedra PostgreSQL Visual FoxPro Teradata	
Object-oriented	FastObjects GemFire KE Texpress	ObjectStore Versant	Multi-dimensional	D^3 Essbase	Oracle Express Edition

Relational Databases

- A **relational database** is a database that stores data in tables that consist of rows and columns.

- Each row has a primary key and each column has a unique name.

- A developer of a relational database refers to a file as a *relation*, a record as a *tuple*, and a field as an *attribute*.

Relational Databases

- A user of a relational database refers to a file as a **table**, a record as a **row**, and a field as a **column**.

- In addition to storing data, a relational database also stores data **relationships**, which are links within the data.

- With a relational database, you can set up a relationship between tables with common fields.

Relational Databases

Relational Databases

- A developer of relational databases uses normalization to organize the data in the database.

- *Normalization* is a process designed to ensure the data within the relations (tables) contains the least amount of duplication.

SQL

- **Structured Query Language (SQL)** is a popular query language that allows users to manage, update, and retrieve data.
- SQL has special keywords and rules that users include in SQL statements.

```
SELECT CLASS_TITLE, CLASS_CODE, MAXIMUM_ENROLLMENT -
  CURRENT_ENROLLMENT AS SEATS_REMAINING
FROM SCHEDULE_OF_CLASSES, CLASS_CATALOG
WHERE SCHEDULE_OF_CLASSES.CLASS_CODE =
  CLASS_CATALOG.CLASS_CODE
ORDER BY CLASS_TITLE
```

Class Title	Class Section	Seats Remaining
Algebra 1	51	14
Art Appreciation	52	19
English Composition 1	02	5
Introduction to Sociology	01	14

Object-Oriented Databases

- An **object-oriented database (OODB)** stores data in objects.
- An **object** is an item that contains data, as well as the actions that read or process the data.
- OODBs can store more types of data than relational databases, access data faster, and allow programmers to reuse objects.

Object-Oriented Databases

- A *multimedia database* stores images, audio clips, and/or video clips.
    - Ex. Voice mail system
- A *groupware database* stores documents such as schedules, calendars, manuals, memos, and reports.
    - Ex. Schedules for meeting times.

Object-Oriented Databases

- A *computer-aided design (CAD) database* stores data about engineering, architectural, and scientific designs.
  - Contains a list of components of the item being designed, the relationship amongst components, and design drafts.
- A *hypertext database* contains text links to other types of documents.
- A *hypermedia database* contains text, graphics, video, and sound.

Object Query Language

- Object-oriented and object-relational databases often use a query language called *object query language (OQL)* to manipulate and retrieve data.
- OQL is similar to SQL and uses many of the same rules, grammar, and keywords.

Multidimensional Databases

- A **multidimensional database** stores data in dimensions.

- Whereas a relational database is a two-dimensional table, a multidimensional database can store more than two dimensions of data.

- These multiple dimensions, known as *hypercube*, allow users to access and analyze any view of the database data.

Multidimensional Databases

- The number of dimensions varies.
  - Ex. A retail business might have 4: products, customers, regions, and time.
- Nearly every multidimensional database has a dimension of time.
- Multidimensional databases can consolidate data much faster than a relational database.

Data Warehouses

- One application that uses multidimensional databases is a **data warehouse**, which is a huge database that stores and manages the data required to analyze historical and current transactions.

- A data warehouse typically has a user-friendly interface so users can easily interact with its data.

Data Warehouses

- Data in a *distributed database* exists in many separate locations through a network or the Internet and is accessible through a single server.
- Data warehouses often use a process called *data mining* to find patterns and relationships amongst data.
  - Ex. E-commerce for customer preferences
- A smaller version of a data warehouse, a *data mart*, contains a database that helps a specific group or department make decisions.

Web Databases

- Much of the information on the Internet exists in databases stored on the Web.
- Some Web databases are *collaborative databases*, where users store and share photos, videos, recordings, and other personal media with registered users.
- The Web page is used as the front end to the database.

Web Database

- A Web database usually resides on a *database server*, which is a computer that stores and provides access to a database.
- One type of program that manages the sending and receiving of data between the front end and the database server is a *CGI (Common Gateway Interface) script.*

Database Administration

- Managing a company's database requires a great deal of coordination.

- The role of coordinating the use of the database belongs to the database analysts and administrators.

Database Design Guidelines

- A carefully designed database makes it easier for a user to query the database, modify its data, and create reports.
  - Determine the purpose of the database
  - Design the tables or files
  - Design the records and fields for each table
  - Determine the relationship among the tables

Role of the Database Analysts and Administrators

- The **database analyst (DA)**, or *data modeler*, focuses on the meaning and usage of data.
- The DA decides on the proper placement of fields, defines the relationships among data, and identifies users' access privileges.
- The **database administrator (DBA)** requires a more technical inside view of the data.
- The DBA creates and maintains the data dictionary, manages security of the database, monitors the performance of the database, and checks backup and recovery procedures.

Role of the Employee as a User

- Today, employees access databases from their office desktop computers, notebook computers, or even smart phones and other mobile devices.

- Employees interact with databases related to inventory and identify new data in the database.