Chapter 7
Database Applications and Privacy Implications

Objectives

- ✓ Explain what a database is and describe its basic structure.
- ✓ Identify what kinds of tasks can be carried on with database software.
- ✓ Describe different kinds of database management software, from simple file managers to complex relational databases.




Objectives (continued)

- ✓ Describe database operations for storing, organizing (e.g., sorting), updating, retrieving (via querying), integrating, and aggregating information.
- ✓ Give examples of ways in which large, easily accessible databases make lives safer or more convenient.
- ✓ Explain the ways databases may threaten our privacy.




The Electronic File Cabinet: Database Basics

What Good Is a Database?

- A database:
  - An organized collection of data stored on secondary memory (e.g., on disks)
- Database software:
  - Application software (somewhat similar to spreadsheet software)
  - Designed to manage data




The Electronic File Cabinet: Database Basics

- ✓ Advantages offered by computerized databases:
  - ➢ Make it easier to store large quantities of data
  - ➢ Make it easier to retrieve data quickly and flexibly
  - ➢ Make it easy to organize, integrate, and aggregate data
  - ➢ Make it easy to distribute data

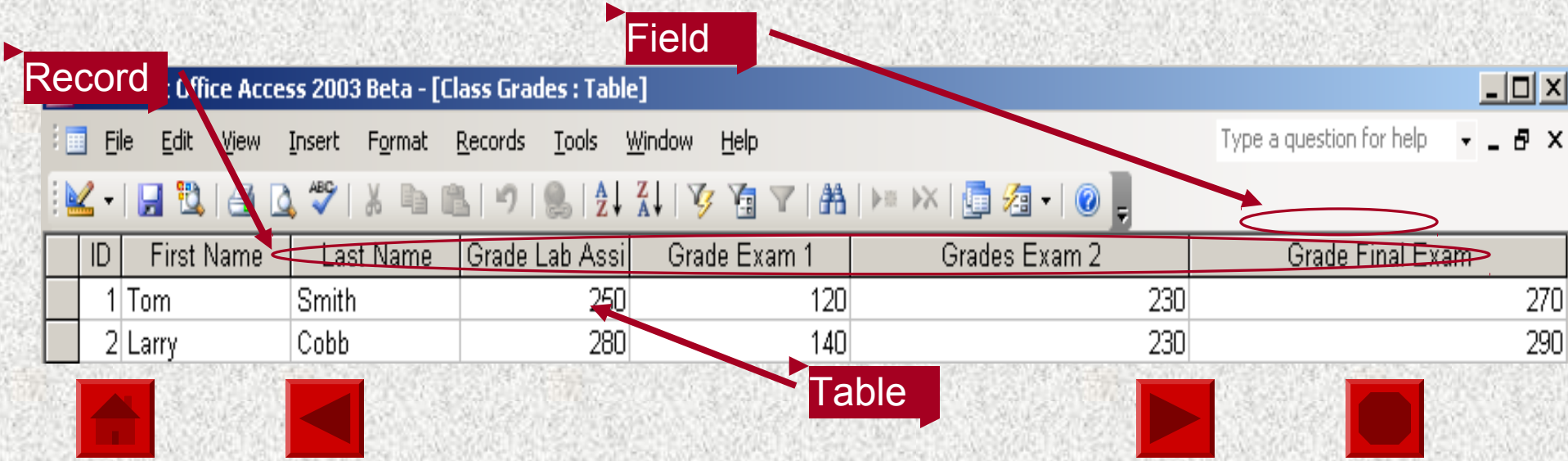



Database Anatomy

- **Database management system (DBMS):** a software tool for organizing the storage and retrieval of data
- **Database:** an organized collection of data stored in a secondary storage
  - Typically composed of one or more **tables**
    - A collection of related data
    - Basic units of organization: **records**



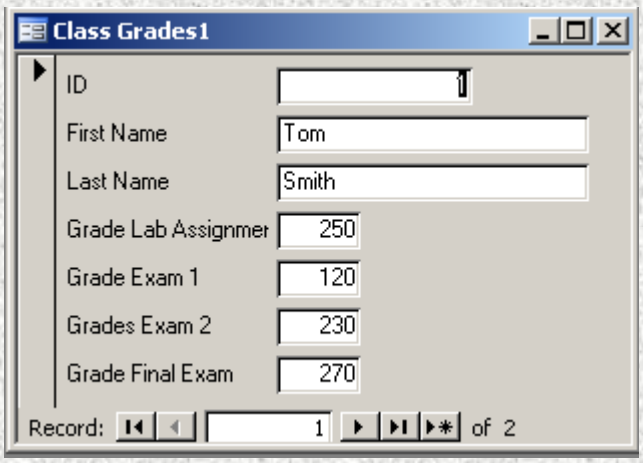

- A **record** is a structured set of data pertaining to one entity (person, product, event, etc.).
- The structure of a record: a list of one or more **fields** that contain individual data.
- The name of a field is called **attribute.**



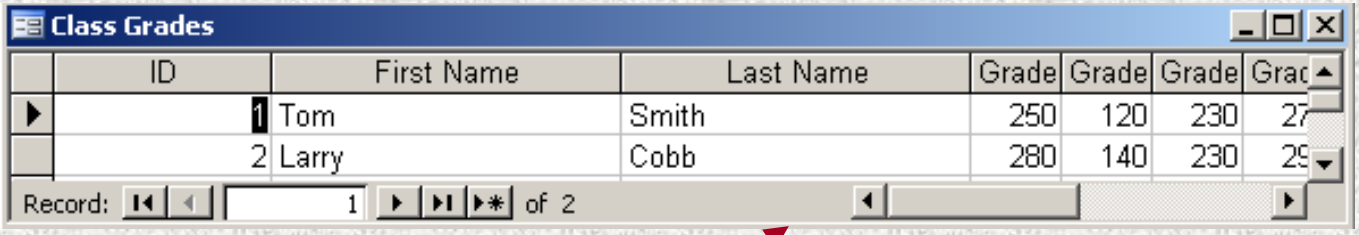
- The type of data a field can hold is determined by its:
  - **Field (attribute) type** or
  - **Data type**

- Database management systems provide the users with many ways to view data. Examples:
  - **Form views**
    - Show one record at a time
  - **List views**
    - Display several records in lists similar to the way a spreadsheet displays data

- In any view, fields can be rearranged without changing the underlying data.



Form View



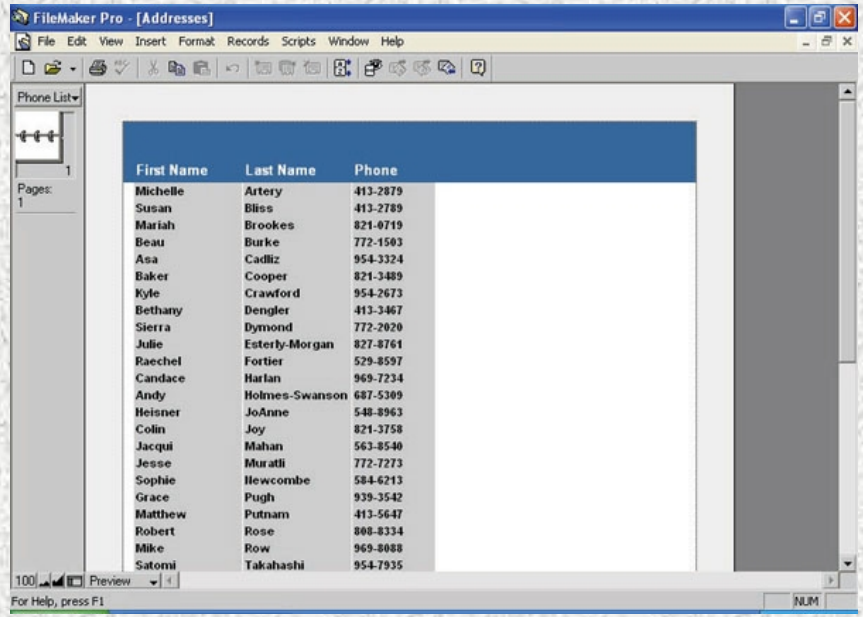
List View

Database Operations (examples)

- **Import:** receive data (that are, for instance, in the form of text files)
- **Browse:** navigate through data
- **Query:** find records that match specific criteria




- **Sort:** rearrange records with respect to values of selected attribute(s) (e.g., alphabetically, numerically, etc.)
- **Print reports, labels, and form letters:** A report is an ordered list of selected records and fields in an easy-to-read format.



- **Sort:** rearrange records with respect to values of selected attribute(s) (e.g., alphabetically, numerically, etc.)
- **Print reports, labels, and form letters:** A report is an ordered list of selected records and fields in an easy-to-read format.

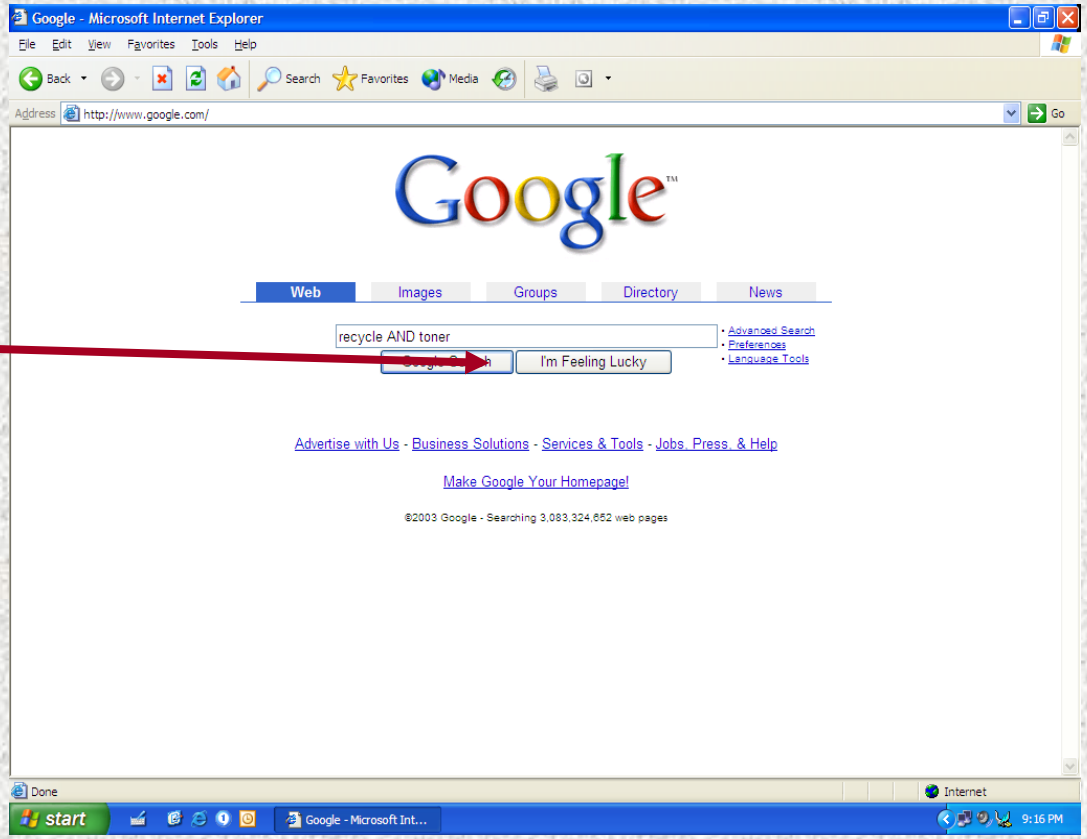



- Most modern relational database management systems support a standard language for programming complex queries called **SQL (*Structured Query Language*).**
  - SQL is available for many database management systems.
  - Programmers and sophisticated users don't need to learn new languages when they work with new systems.
  - The graphical user interfaces allow point-and-click queries that insulate users from the arcane details of the query language.

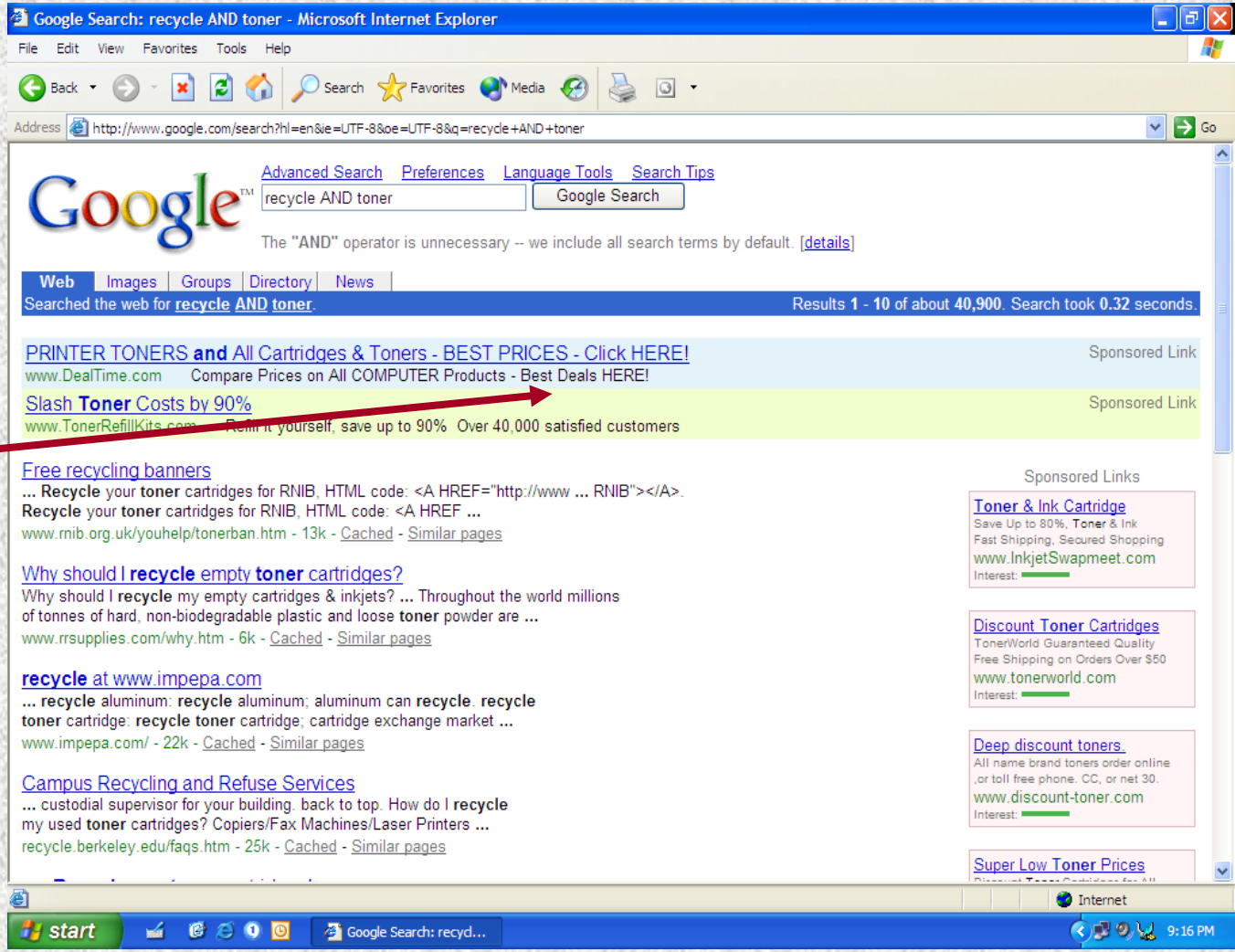



Screen Test: Querying a Web Search Database

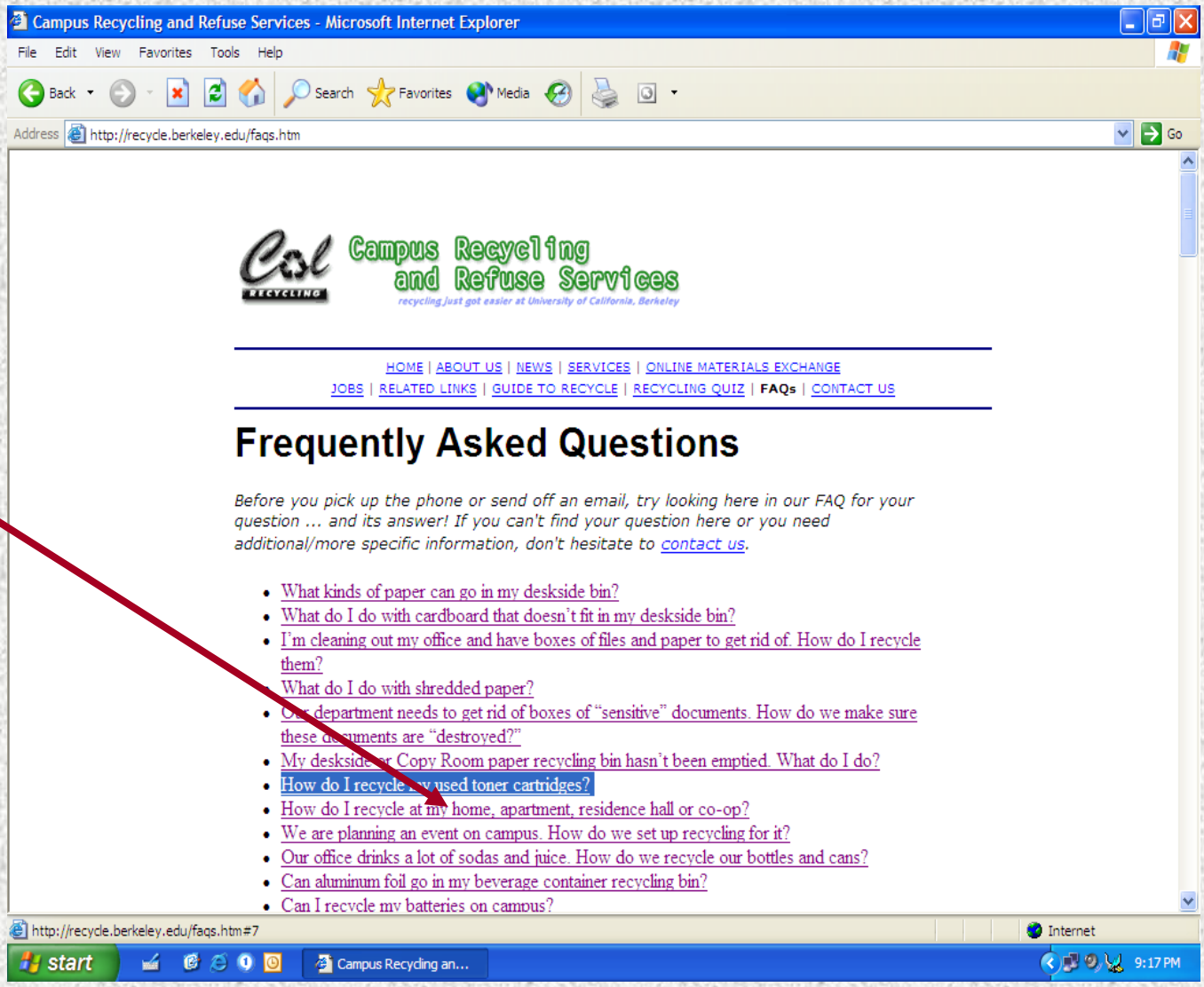
To search for articles online about a new method for recycling laser printer toner cartridges, you can visit a search engine, such as Google.



When you hit the "Search" button, the search engine will return a list of links, sorted by relevancy.



Selecting any of the links will cause the corresponding page to load.

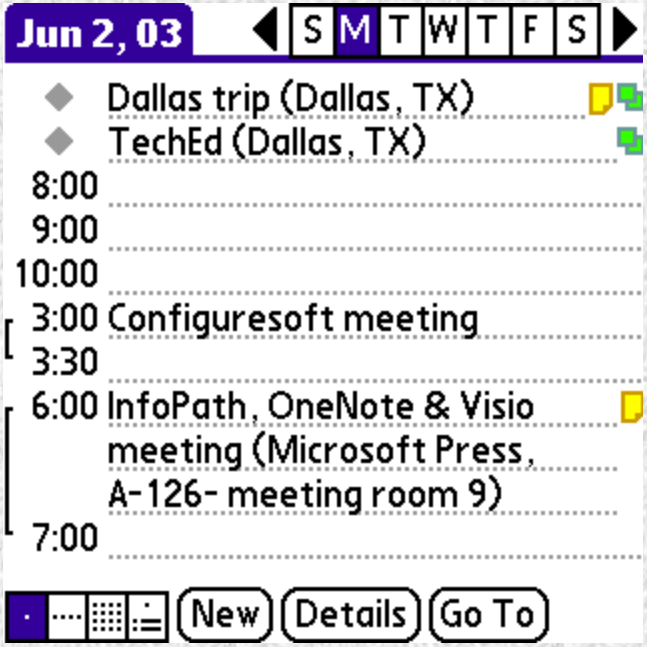


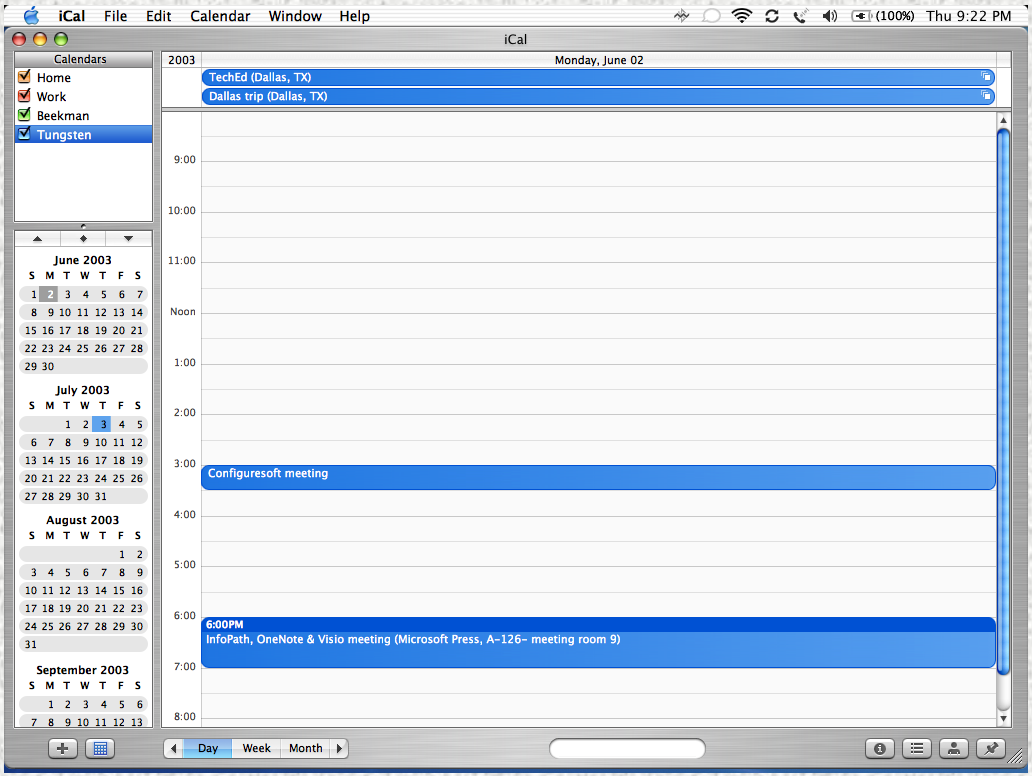
Special-Purpose Database Programs

- **Specialized database management systems**: preprogrammed for specific data storage and retrieval purposes
- **Geographical information systems (GIS):** include geographic and demographic data in map form

- **Personal information manager (PIM)**: an electronic organizer
- Automates some or all of the following functions:
  - Address/phone book
  - Appointment calendar
  - To-do list
  - Miscellaneous notes
- Handheld computers can share information with applications such as iCalendar running on PCs and Macintoshes.





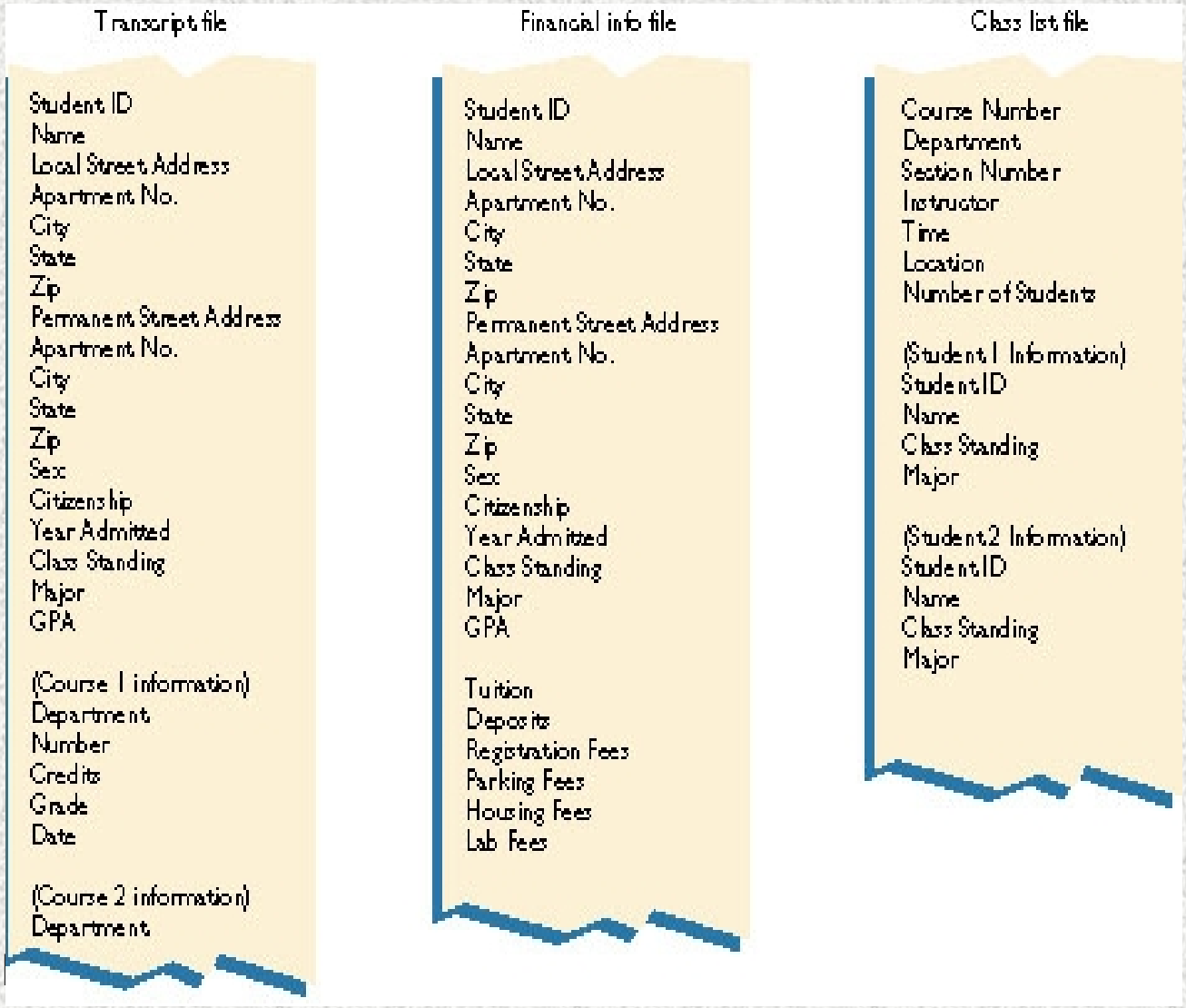
Beyond the Basics: Database Management Systems

From File Managers to Database Management Systems

- ✓ **File manager:** enables users to work with few files at a time
- ✓ **Database management system (DBMS):** manipulates data in a large collection of files, cross-referencing between files as needed




- ✓ A DBMS can be used interactively, or can be controlled directly by other programs.



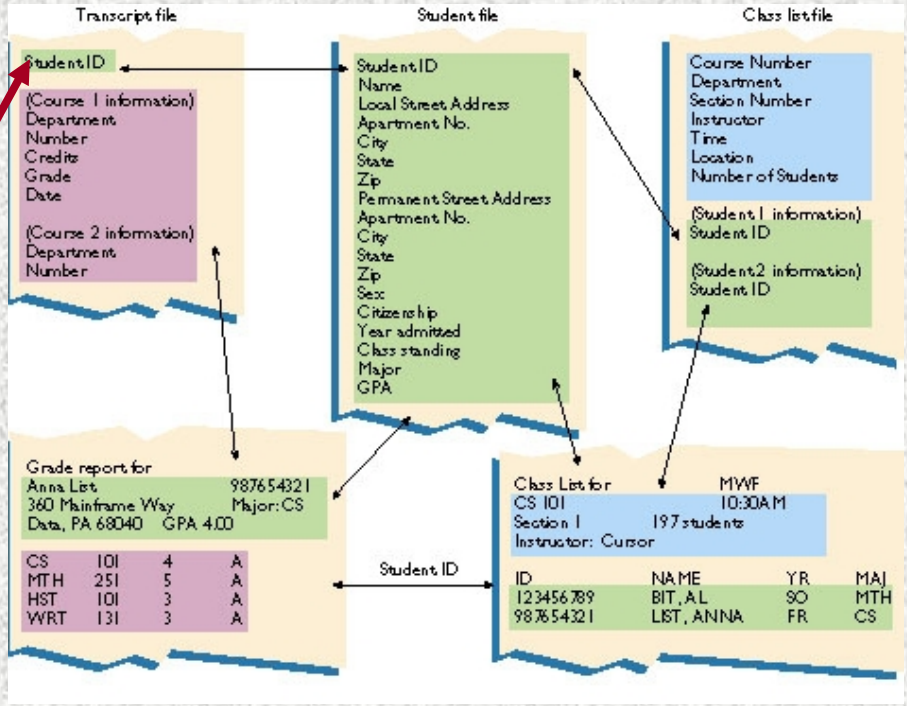
What Makes a Database Relational?

- ✓ A **relational database** is organized in tables in a way similar to how mathematical relations are organized.
- Each table is a collection of tuples, usually implemented as fixed-structure records.
- ✓ The adjective *relational* that qualifies noun *database* indicates:
  - ➢ The structure of the data
  - ➢ The ways of manipulating data




- The structure of a relational database is based on the relational model—a mathematical model (of relation) that combines data in tables.

A database is **relational** when files are **related** to each other, such as this **Student ID** field in the Student file.



- ✓ In relational database queries are based on predicate calculus (a system of logical expressions, often with quantifiers).
- ✓ For instance, a logical expression (a query):

$(\exists s)(s > 1 \;\&\&\; s < p \;\&\; \text{divisor}(s, p) \;\&\; p > 1 \;\&\; p < 100)$

will return the set of all composite numbers between 1 and 100.




Examples of SQL queries

UPDATE customers SET salesperson = "Mike"
WHERE state = "NH"

```
SELECT customer_id, customer_name,
COUNT(order_id) as total
FROM customers INNER JOIN orders ON
customers.customer_id = orders.customer_id
GROUP BY customer_id, customer_name
HAVING COUNT(order_id) > 5
ORDER BY COUNT(order_id) DESC
```

The Many Faces of Databases

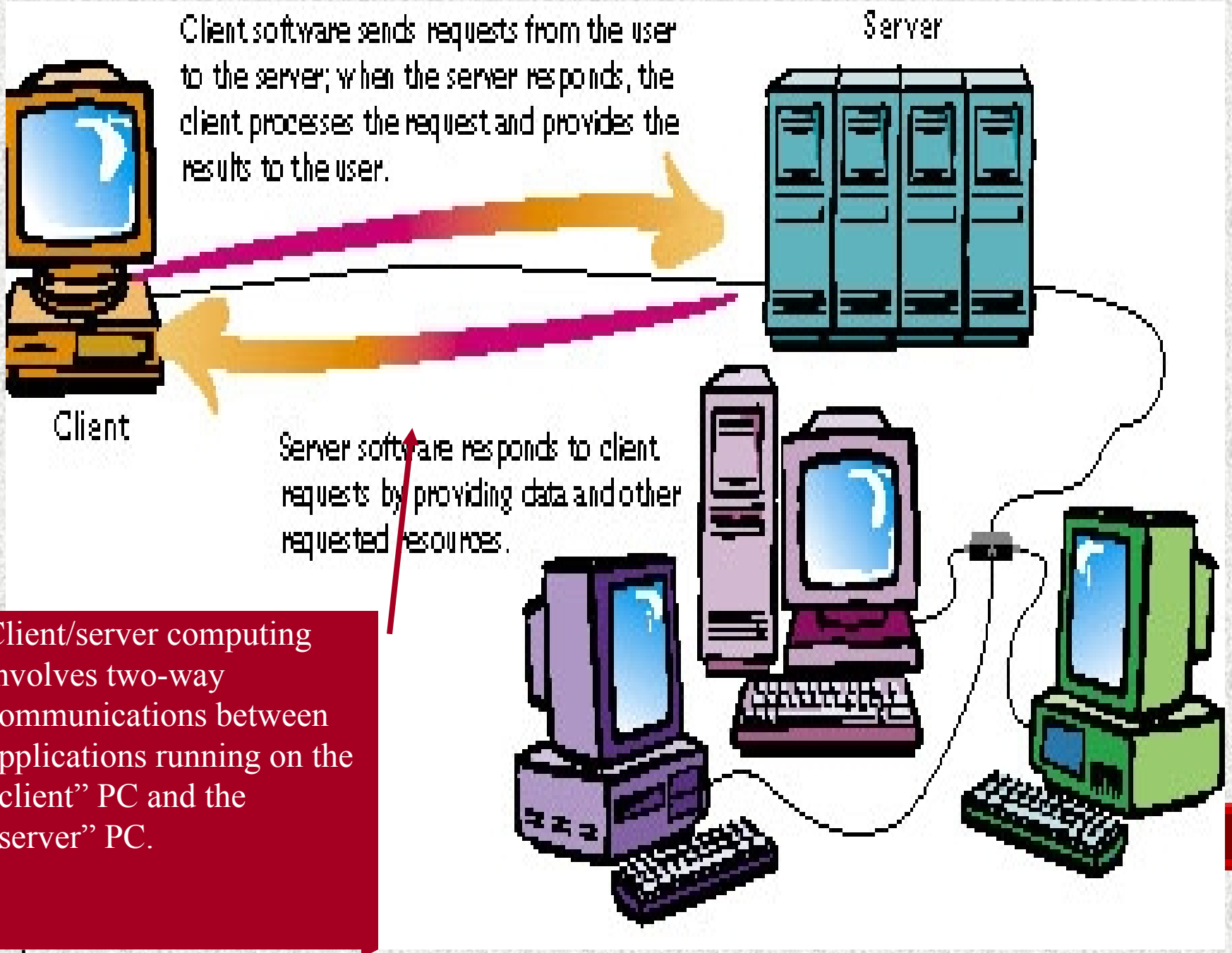
- ✓ Large databases often contain hundreds of interrelated tables.
- ✓ A database management system shields the end users from the complex inner workings of the system, providing them with only the information and commands they need to get their jobs done.




Database Trends

- **Batch processing:** users accumulate transactions and input them into the computer in large batches
- **Real-Time computing**: allows instantaneous access to information
- **Interactive processing:**
  - Users can interact with data through terminals, viewing and changing values online.



Client/server computing involves two-way communications between applications running on the "client" PC and the "server" PC.

Downsizing and Decentralizing

- Using a client/server approach
  - Today many businesses use a **client/server** approach, using database servers.
  - Users can take advantage of the PC's simple user interface and convenience, while still having access to data stored on large server systems.




- Data Mining (no more GIGO):
  - The discovery and extraction of hidden or implicit trends and patterns between data in huge repositories (even in "garbage")
  - Uses statistics and artificial intelligence
    - Identifies trends and patterns in data that, for instance, could have been overlooked while working with specific queries




- Object-Oriented Databases
  - Make database design and use more flexible but also less systemic and, sometimes, theoretically unsound.
  - Store **objects** that contain data together with programs that manipulate these data.
  - May be used in conjunction with **object-oriented programming languages**




- Object-Oriented Databases: A simplistic example
  - A database of images, containing a ***class*** for photos
  - One instance of this class – one ***object*** for every photograph in the database
  - ***Data encapsulated*** by this object: photographer's name, date when created, textual description of photo, copyright status, and the image itself
  - One ***operation*** for the class – producing a thumbnail




Object-Oriented

- Easy manipulation of various types of data
- Saves time by reusing objects
- Associates actions with the data
- A lack of clear logical structure (may be chaotic)
- **Difficult to formally design, analyze, and prove correctness**

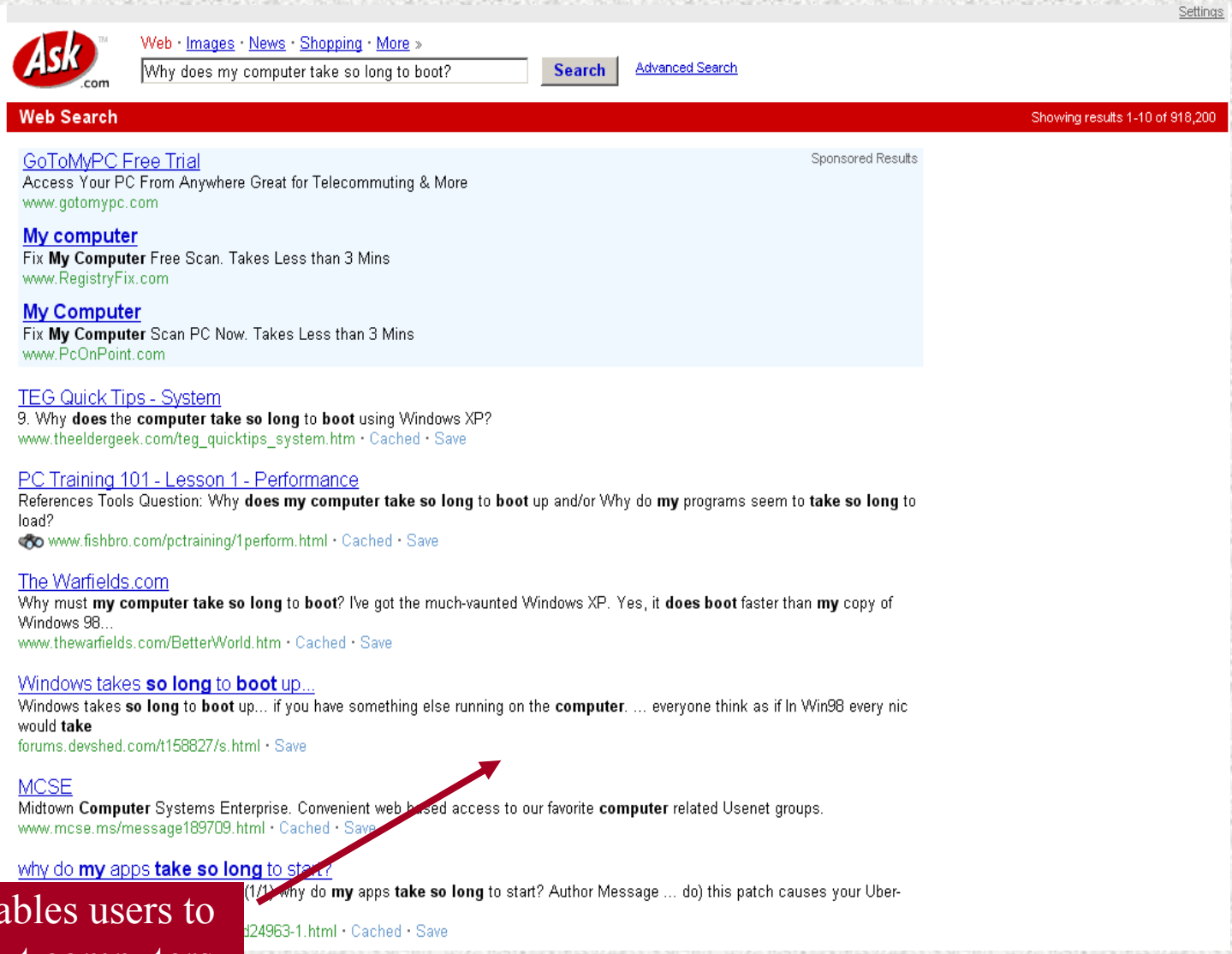
Relational

- A very powerful and reliable tool for storage and retrieval of logically structured data.
- Gives advantage to those with good math skills
- Not intended to use with logically unstructured data (audio, video clips)
- Is being expanded to incorporate objects.




- Multimedia Databases
  - Serve as indexes for art, photographs, maps, video clips, and other media files
- Natural Language Databases
  - Allow users to ask for data using the same language used to address humans (***natural language*** queries)
  - Future databases will undoubtedly incorporate more methods of artificial intelligence.



www.ask.com enables users to ask questions about computers by stating their queries in plain English and other natural languages.

Inc.

No Secrets: Computers and Privacy

Personal Data: All About You

- More than 15,000 specialized marketing databases contain 2,000,000,000 names.
- These databases contain personal characteristics like age, income, religion, etc..




What Is *Privacy*?

- A common theme in privacy is the notion of ***access***:
  - Physical proximity to the person
  - Knowledge about that person
- People need a certain amount of privacy to maintain their security, dignity, and freedom.
  - How much dignity or freedom would you have if everyone could read your mind?
- Information about people can be of great value to society. But also to a tyrant.
  - Many parents would like to know the identities of convicted sex offenders.




The Privacy Invasion Problem

- Protection against some forms of unlawful invasion of privacy is implied by 4th and 9th Amendment to the U.S. Constitution.
  - The right to privacy is also implied by other constitutional guarantees.
  - Debates rage about what this means.
- Federal and state laws provide forms of privacy protection.
  - Most of those laws were written years ago.
- Most European countries have had strong privacy protection laws for years.




Big Brother and Big Business

- Other information technologies amplify the threat to personal privacy:
  - **Computer networks** make it possible for electronic gathering and transmition of personal data.
  - Examples of information gathered electronically:
    - Passwords
    - Credit card numbers
    - Other consumer information




Big Brother and Big Business

- ✓ Big Brother means big, powerful, and intrusive government.
- ✓ The phrase comes from the George Orwell's book “1984” (highly recommended reading).
- ✓ Here are links (make sure you don't violate copyright of these materials; while use for educational purposes only is usually exempt from copyright, unauthorized downloading may constitute a crime; if in doubt, please, consult a lawyer)




Big Brother and Big Business

Here are links (make sure you don't violate copyright of these materials; while use for educational purposes only is usually exempt from copyright, unauthorized downloading may constitute a crime; if in doubt, please, consult a lawyer)

http://www.mondopolitico.com/library/1984/1984.htm

http://video.google.com/videoplay?docid=-5464625623984168940#



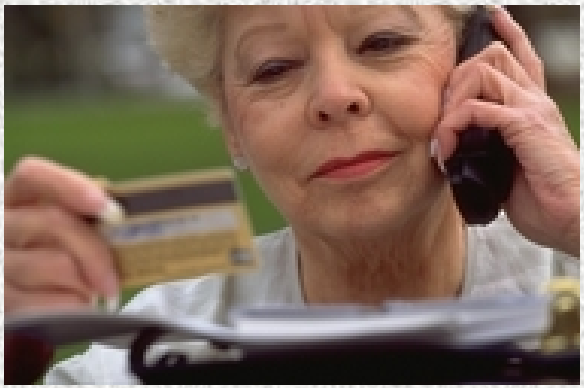

No Secrets: Computers and Privacy

- **Workplace monitoring technology** enables managers to learn more than ever before about the work habits and patterns of workers.
- **Surveillance cameras** are used increasingly for nabbing routine traffic violators and detecting security violators. Their data can be combined with picture databases to locate criminals—and others.
- **Surveillance satellites** can provide permanent peepholes into our lives for anyone willing to pay the price.
- **Cell phones** are now required, by law, to include technology to determine and transmit their locations to emergency personnel responding to 911 calls.

No Secrets: Computers and Privacy

Rules of Thumb: Your Private Rights

- ✓ Your Social Security number is yours —don't give it away.
- ✓ Say "no" to direct mail and phone solicitations, sharing of personal information, and pollsters.
- ✓ Know your electronic rights.
- ✓ Support organizations that fight for privacy rights.



Lesson Summary

- ✓ Database programs enable users to quickly and efficiently store, organize, retrieve, communicate, and manage large amounts of information.
- ✓ Database programs enable users to view data in a variety of ways, sort records in any order, and print reports, mailing labels, and other custom printouts.
- ✓ Database management systems (DBMSs) can work with several data sources at once, cross-referencing information among files when appropriate.




Lesson Summary (continued)

- ✓ The trend today is clearly away from large, centralized databases accessible only to data-processing staff.
- ✓ Organizations are moving toward a client/server approach that enables users to have access to data stored in servers throughout the organization's network.
- ✓ The accumulation of data by government agencies and businesses is a growing threat to our right to privacy.
- ✓ While there are many legitimate uses for these procedures, there is also a great potential for abuse.