New Perspectives on Computer Concepts 2014

Chapter 11 Solutions: Databases

In this document, you will find the Solutions to the Chapter QuickChecks, Lab Assignments, Interactive Summary, Interactive Situation Questions, Projects, Learning Objectives Checkpoints, Concept Map, Issue Try It!, and Information Tools Try It!.

QuickChecks

QuickCheck A (page 621)
1. analytical
2. flat
3. ERD
4. relational
5. object^object-oriented^object oriented

QuickCheck B (page 632)
1. True^T
2. independence
3. client
4. server
5. XML

QuickCheck C (page 644)
1. key
2. True^T
3. validation
4. normalization
5. indexed

QuickCheck D (page 652)
1. Boolean
2. False^F
3. SELECT
4. UPDATE
5. False^F

QuickCheck E (page 659)
1. True^T
2. Encryption
3. access
4. audit
5. opt

Lab Assignments

Lab 11-A: Working with Database Software

Start the interactive part of the lab. Make sure you've enabled Tracking if you want to save your QuickCheck results. Perform each of the lab steps as directed, and answer all of the lab QuickCheck questions. When you exit the lab, your answers are automatically graded and your results are displayed. Student answers are saved in their Tracking files.

A friend wants to create a table to store information about a collection of old books. List the fields you might include in the table to store information about the books. For each field, specify the field name, data type (text, numeric, date, etc.), and field length. Indicate the primary key(s), and describe how you would sort and/or index the data.

Students should specify the field name, data type (text, numeric, date, etc.), field length, and primary

key(s). Then they should describe how they would sort and/or index the data in a database of old books. Answers may vary, but should be similar to the following:

Field Name	Data Type	Field Length	Primary Key
Title	Character	100	Yes
AuthorLastName	Character	100	Yes
Author First Name	Character	50	No
DatePublished	Date	8	No
Publisher	Character	50	No
Subject	Character	100	No
PlaceofPublication	Character	25	No
NumberofPages	Numeric (integer)	5	No
Remarks	Memo	200	No

[Note to instructor: Students should not consolidate first and last name into one field. If students wonder what to do about a book with multiple authors, the correct structure would be to create a related table containing author first names and last names.]

Use Microsoft Access or any available file or database management software to create the structure for the table you specified in Assignment 2. Enter at least 10 records. Print a list of all your data.
Assign this activity only if students have access to Microsoft Access or a similar database management package. Students should submit a printout showing at least ten records entered into a database structure from #2.

Make a list of five queries that might be useful if your database had hundreds of records. Try these queries on your table. For each query, list the records that were selected.
Students should submit five queries and results for the database created in #3. In the lab, students learned to use a variety of query methods: the Edit/Find menu option, the Filter button, and the Query button. They were briefly shown how to access the SQL facility in Access, too. The method that a student uses to perform the queries might affect the way in which they submit this assignment.

If a student uses the Edit/Find menu, a query is likely to be specified like this:

Find: Tolkein Look in: AuthorLastName Match: Whole field

If a student uses the Filter or Query button, a query is likely to specify like this:

Field: AuthorLastName Criteria: Tolkien

If a student uses SQL, a query is likely to be specified like this:

Select AuthorLastName, Title from Books where AuthorLastName = 'Tolkien'

The syntax for complex Access queries allows wild cards, such as *like T*,* and Boolean operators, such as *>1920 AND <1950,* and *NOT Rowling*.

Sketch a report on paper that uses some of the fields in your table. Make sure your report contains a file and headings for each field. Specify whether you would like to align your data at the right, center, or left of each column. Use your software to generate and print the report.
Students should sketch the plan for a report from the old books database, indicating titles, column headings, column contents, and alignment. If possible, students should create and print this report from Access.

Interactive Summary

Interactive Summary A

A **database** is a collection of information, typically stored as computer files. The information it contains can be stored, updated, organized, output, distributed, searched, and analyzed. A filing cabinet full of folders and papers would be classified as an **unstructured** file. A **structured** file uses a uniform format to store data for each person or thing in the file. The simplest model for storing data is a **flat** file that consists of a single, two-dimensional table of data elements. Each row in the table is a **record**, and each column of the table is a **field**. Each kind of record is referred to as a record **type**. A record that contains data is sometimes referred to as a record **occurrence**. In a database, records can be related by one-to-**one** relationships, one-to-many relationships, or many-to-many relationships. The number of associations that can exist between two record types is referred to as **cardinality**. Relationships can be depicted graphically by using **entity**-relationship diagrams. **Hierarchical** databases allow only one-to-many relationships. **Network** databases allow one-to-many and many-to-many relationships. **Relational** databases exist as a series of tables that can be related by common fields. A **dimensional** database organizes relationships over three or more dimensions. An **object** database stores data in objects that can be grouped into classes and defined by attributes and methods.

1. database
2. unstructured
3. structured
4. flat
5. record
6. field
7. type
8. occurrence, instance
9. one
10. cardinality
11. entity
12. Hierarchical
13. Network
14. Relational
15. dimensional, multidimensional, multi-dimensional
16. object, object-oriented, object oriented, OO

Interactive Summary B

Flat files can be created and manipulated by using a variety of tools, including word processing and spreadsheet software. For databases composed of more than one record type, however, it is best to use a database management system, which is abbreviated as **DBMS**. An **entry**-level database management system typically handles many simultaneous searches, but has limited capability to deal with multiple simultaneous updates. Handling billions of records and performing hundreds of transactions every second requires database **server** software. The data in a database can be accessed over the Web. A simple process called **static** Web publishing converts a database report into an HTML document, which can be displayed by a browser. More sophisticated **dynamic** Web publishing produces data from a database on demand. HTML forms not only provide search capabilities, but can also be used to add or modify data in a database with a Web browser. **XML** provides a Web-based data management tool that uses special tags as field names within a document.

1. DBMS
2. entry
3. server
4. static
5. dynamic
6. XML, Extensible Markup Language

Interactive Summary C

The first step in designing a relational database is to define its fields by specifying a field name and data type. Integer, date, and **real** data types are used for fields containing data that might be mathematically manipulated. The **text** data type is used for fixed-length fields containing text that is not intended to be mathematically manipulated. The **memo** data type is a variable-length field for entering text. The **logical** data type is used to store true/false or yes/no data. The **hyperlink** data type can be used to store URLs. The **BLOB** data type is used to store binary data, such as MP3 files or graphics. When designing fields, a database designer can also include field formats, field **validation** rules, and lookup routines to reduce data entry errors. The number of tables in a database can be determined by a process called **normalization**, which helps a database designer group fields into record types and avoid data redundancy. A database designer must also consider how to sort or index records. The **sort** key for a table specifies the order in which records are stored and indicates where new records are inserted in a

table. A database provides an alternative way to organize records, using a series of keys and pointers to temporarily arrange data without affecting the physical sequence of records specified by the sort order.

1. real
2. text, character, string
3. memo
4. logical, Boolean
5. hyperlink
6. BLOB, binary large object
7. validation
8. normalization
9. sort

Interactive Summary D

SQL is a database query language that typically works behind the scenes as an intermediary between the database **client** software provided to users and the database itself. Although the specifications for searches and other database tasks are collected by easy-to-use graphical user interfaces, those specifications are converted into SQL **queries**, which can communicate directly with the database. An SQL query contains SQL **keywords**, such as SELECT, FROM, INSERT, JOIN, and WHERE, plus **parameters** that specify the details of the command. Records can be removed from a database using the SQL **DELETE** command. Records can be added to a table using the SQL **INSERT** command. To search for data, you can use the SQL **SELECT** command. To change or replace the data in a field requires the SQL **UPDATE** command. SQL also provides a **JOIN** command that can be used to temporarily consolidate two tables so that data can be accessed simultaneously from both of them.

1. client
2. queries, commands, statements
3. keywords, key words, commands
4. parameters
5. DELETE
6. INSERT
7. SELECT
8. UPDATE
9. JOIN

Interactive Summary E

Databases are vulnerable to theft, hacking, and unauthorized access. Although **encryption** cannot prevent a database from being lost or stolen, it can make the data it contains unusable. This security technique, however, is typically used for archived databases, rather than operational databases. Access controls can be used to restrict physical access to a database, limit user privileges, and regulate data **views** that establish which fields a particular user is allowed to access. Tracking database use is also an important security tool. If an intruder breaches database security, a database **audit** can help to identify the damage and correct it. A patchwork of database security law in the U.S. leaves loopholes that allow some types of personal data to be gathered, aggregated, and shared with third parties. Individuals can take steps to protect their personal data by vigilantly monitoring what information is released, using strong passwords, running security software, and becoming familiar with the **privacy** policy for sites that store personal data.

1. encryption
2. views
3. audit, database audit
4. privacy

Interactive Situation Questions

1. relational
2. one, 1
3. text, character, string
4. validation
5. template
6. global
7. record types, tables
8. XML, Extensible Markup Language

Learning Objectives Checkpoints

Study Tip: Make sure you can use your own words to correctly answer each of the red focus questions that appear throughout the chapter.

The Study Tip is designed for the student's own review and is quite lengthy. It is not recommended that instructors assign or grade it.

1. List eight ways the information in a database can be used and applied.

See page 610-612.
Collect Data
Store data
Update data
Organize data
Output data
Distribute data
Find data
Analyze data

2. Create a descriptive example that would help explain the concepts of data mining, data warehouses, predictive analytics, and OLAP to an average adult who has no technical expertise.

There is no right or wrong answer to this question. Students should use the terminology correctly, but should also provide definitions and examples of what they mean. A good explanation uses analogies to situations in an adult's world. Explanations should be short - no more than a few paragraphs.

3. Define basic database terminology, such as fields, records, record types, record occurrences, and cardinality.

Field The smallest meaningful unit of information contained in a data file; column in a database table.
Record The fields of data that pertain to a single entity in a database; a row in a database table.
Record occurrence A record that has been filled with data for a particular entity.
Record type The structure of a record, including the names, length, and data types for each field.
Cardinality A description of the numeric relationship (one-to-one, one-to-many, or many-to-many) that exists between two record types.

Students might also define the following:
Data type The characteristics of data that can be entered into a field in a data file; data types include character, numeric, date, logical, and memo.
Database A collection of information that might be stored in more than one file or in more than one record type.
Database model The underlying structure or category of a database, such as relational, hierarchical, network, or object.
Database structure The arrangement of the fields, tables, and relationships in a database.
Primary key A field in a database that contains data, such as a Social Security number, that is unique to a record.
Structured file A file that consists of a collection of data organized as a set of similarly structured records.
Unstructured file A file that contains data, but that is not in a structured format of fields and records.
Flat file A single file that is the electronic version of a box of index cards, in which all records use the same record format

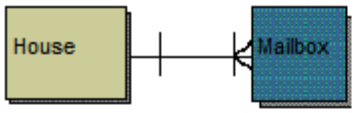
4. For each of the following pairs of record types, draw an ERD showing whether the relationship is one-to-one, one-to-many, or many-to-many. Author—Book Person—Social Security number House—Mailbox Musician—CD

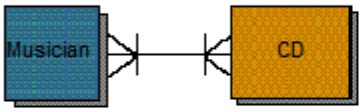
Author—Book (many-to-many)

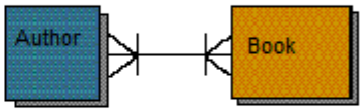
Musician—CD (many-to-many)

Person—Social Security number (one-to-one)

House—Mailbox (one-to-many for apartment buildings, or one-to-one if single family homes are considered)







5. Describe flat files and six other database models. Give examples that illustrate each model.

See pages 617-621. Student examples will vary.

Flat file A single file that is the electronic version of a box of index cards, in which all records use the same record format. Example: Shopping list.

Hierarchical database: Nodes arranged as an upside-down tree in one-to-many relationships. This is an easy-to-use, simple structure. Application limited to data that fits a one-to-many structure. Example: Windows Registry.

Network database: Allows many-to-many relationships. This is a more complex structure than a hierarchical database. Example: Domain Name System (DNS).

Relational database: Stores records in tables that can be related to each other. This is the most commonly used type of database. Example: Music shop database.

Dimensional database: Organizes relationships over three or more dimensions. This is an extension of the relational database model. Example: Publisher database.

Object database: Stores data as objects grouped within classes. Objects can have methods that define their behavior. Example: Database for a business that has store and online orders.

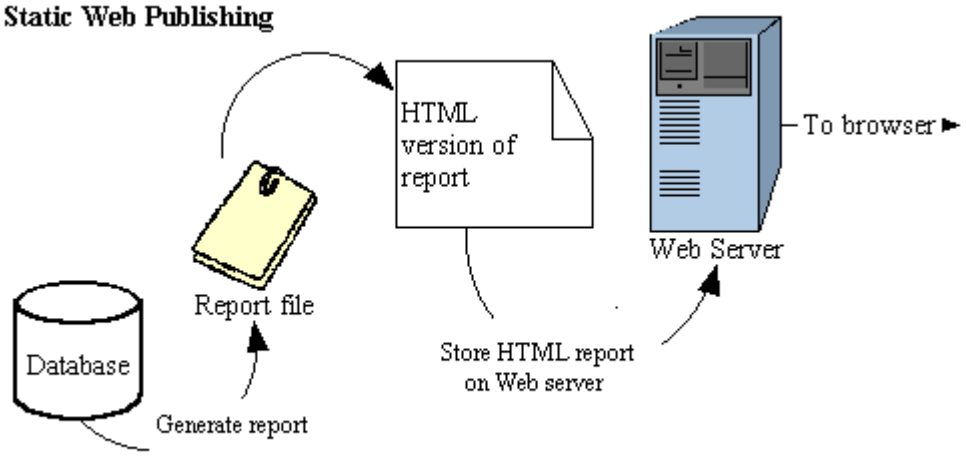
Object-relational database: Combine object-oriented and relational concepts. This is organized as one or more tables, with flexibility to store unique types of data and program code necessary to access that data. Example: iTunes Playlist.

6. Explain the capabilities of various data management tools, such as commercial applications, word processing software, spreadsheet software, custom data management software, and database management software.

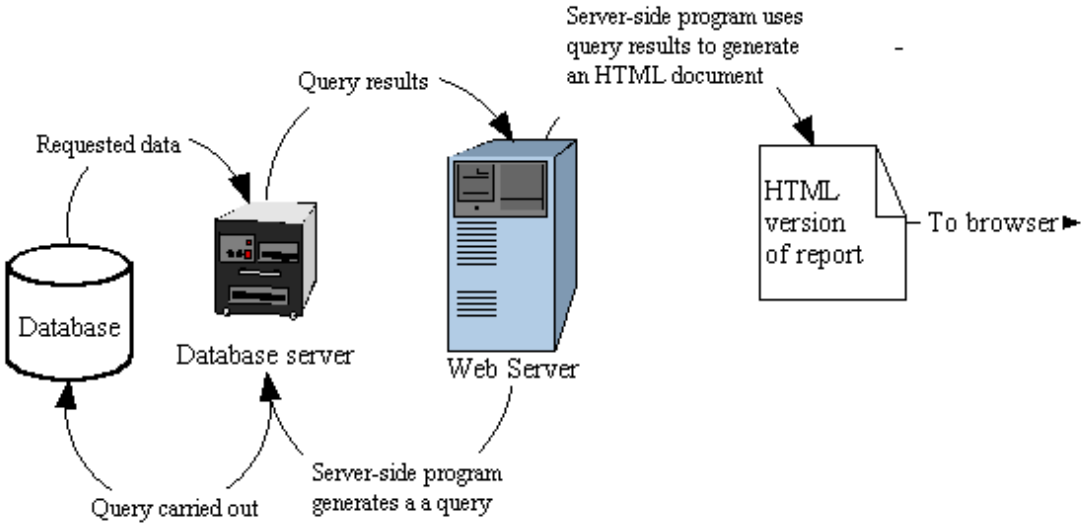
See pages 622-625.

Commercial applications: Appointment book and financial management software. Easy to use, but they don't generally allow you to create new record types because the record types are predefined.
Word processing software: Includes a tool that allows you to specify fields, enter data, and manipulate it. Flat files only. Used most often for mailing lists.
Spreadsheet software: Includes data functions that allow you to manipulate flat files. Rows and columns similar to a database table. Data functions work similarly to other spreadsheet features.
Custom data management software: Can be tailored to the exact needs of a business or individual, but requires time and expertise to create. Must be careful to avoid data dependency.
Database management software: Might specialize in one data model (i.e., relational model). Can be used and reused for different databases. Includes tools needed to define database structure, enter data, organize, find, view and report data.

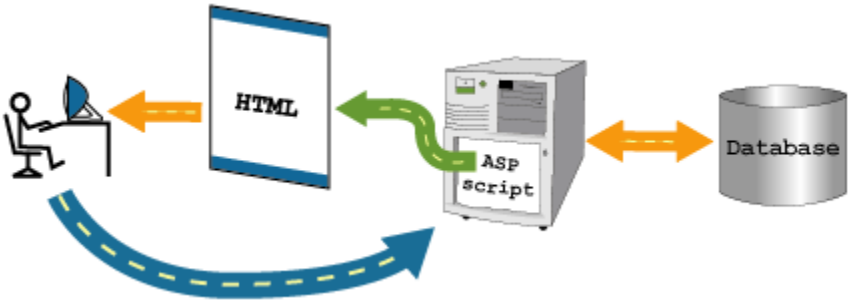
7. Use diagrams to explain different ways of providing Web access to the data in a database.
See pages 627-629



Dynamic Web Publishing



Interactive Web Database Access



8. Provide five examples of data you would store in real, integer, text, logical, date, memo, and BLOB data fields.

See page 635. Responses will vary, but here are some samples:

Real: price, discount, number of tracks, length of song, length of a ship

Integer: number of payments, number of pages, rank, number of children, quantity ordered

Text: name, address, city, state, ZIP code, hair color, item name, manufacturer, song title, book title

Logical: high-school graduate, CPR certified, U.S. citizen, approved, licensed

Date: expiration date, birth date, publication date, renewal date, registration date

Memo: comments, evaluation, description

BLOB: music, video clip, virtual reality tour, program code, photo of employee, photo of CD cover, book cover photo

Computed field: discount (calculated from price and discount percent), current age (calculated based on birthdate and today's date), tax (calculated based on purchase amount and tax rate), number of items purchased, reorder flag (calculated based on quantity in stock and minimum needed).

9. List the techniques that a database designer can use to reduce data entry errors.

See pages 636-637. These techniques include:

- Turn case sensitivity on or off, depending on application
- Provide field formats to make data entry more consistent
- Include field validation rules to filter out invalid data
- Limit data entry to the items on a list
- Use lookup routines to validate and cross-check data

10. Using a real-world entity, like a comic book collection or a recipe file that can be stored in a database, divide the information into fields and record types. Explain the steps you need to take to normalize the data.

Students' answers will vary, depending on the example that they choose. A sample answer for a comic book database that uses three record types might look like this:

COMICS
Title
Publication Date
ISBN
Publisher

AUTHORS
AuthorLastName
AuthorFirstName
ISBN

CHARACTERS
CharacterName
ISBN

11. Using your own examples, explain the differences between sorting and indexing.

Pages 639-640 provide the following information:
A table's **sort order** is the order in which records are stored on disk. Sorted tables typically produce faster queries and updates because they take advantage of clever algorithms that quickly pinpoint records. In a sorted table, new records are inserted to maintain the order. If no sort order is specified, new records are appended to the end of the file, resulting in a file that is not in any particular order. Queries and updates within an unsorted database are slow because the only algorithm for searching an unsorted table requires a sequential look at each record.

A **database index** is very similar to an index in a book that contains a list of keywords and a pointer to the page where they can be found. A database index contains a list of keys, and each key provides a pointer to the record that contains the rest of the fields related to that key.

Key points:

- A sort order is the physical arrangement of records.
- An index is a separate file with pointers to the original table.
- Only one sort order is possible for a table.

- A table can have multiple indexes.

12. Describe how a database report template works. List five principles for creating effective report templates.

See page 642.

A **report template** contains the outline or general specifications for a report, including such elements as the report title, fields to include, fields to subtotal or total, and report format specifications. The template does not, however, contain data from the database. Data is merged into the template when you actually run a report.

• Supply only the information required. Too much information can make it difficult to identify what is essential.

• Present information in a usable format. For example, if subtotals are necessary for making a decision, include them. The people who use reports should not have to make additional manual calculations.

• Information should be timely. Reports must arrive in time to be used for effective decision making. Some decisions require periodic information— for example, monthly sales reports. Other decisions require ongoing information, such as current stock prices, that will be best satisfied by a continuous display.

• Information should be presented in a clear, unambiguous format and include necessary titles, page numbers, dates, labels, and column headings.

• Present information in the format most appropriate for the audience. In many cases, a traditional report organized in rows and columns is most appropriate. In other cases, graphs might be more effective.

13. Imagine that you must access a library card catalog using SQL. Write an SQL query that you would use to search for any books by J. K. Rowling in a table called Books, where authors' names are stored in a field called AuthorName and book titles are stored in a field called Title.

SELECT Title FROM Books WHERE AuthorName = 'J. K. Rowling'

14. Explain the extent to which encryption, user privileges, and audits can secure a database.

Encryption:

- If stolen, the data in an encrypted data is unlikely to be compromised by hackers
- Makes the process of accessing data more resource intensive
- Should not be the only means of securing a database

User privileges:

- Are a type of access control
- Allows access to certain fields or records based on who is accessing the database
- Lowest level use privilege is read only
- Other privileges include adding records, changing database structure, changing the data in fields, indexing, and copying
- Should not be the only security measure protecting a database

Database audits:

- A procedure that monitors and records user activity within a database
- Can identify potential intruders before they can compromise a system
- Tracking who makes changes in a database helps an organization maintain accountability, trace the source of errors, and make comprehensive corrections
- Might only reveal a database breach after the fact, so should not be only technology used to secure a database

15. Describe the status of privacy legislation in your country, providing examples of relevant laws and regulations.

Student responses will vary by country. See pages 656-658. The text mentions the following:

- The European Union's Privacy Directive 95/46/EC mandates that database breaches from accidental loss or electronic attack are disclosed.
- Canada's Personal Information Protection and Electronics Documents Act (PIPEDA) mandates that database breaches from accidental loss or electronic attack are disclosed.
- Despite progress, governments around the world still have much work to do in designing and enforcing effective privacy laws.
- Canada, the European Union, and Japan have more stringent privacy regulations than the U.S.
- The U.S. Privacy Act of 1974 requires government agencies to disclose to an individual the contents of his or her records. It places restrictions on how agencies can share an individual's data with other people and agencies. The Act also requires agencies to follow "fair information practices" when gathering and handling personal data, and allows individuals to sue the government for violating its provisions.
- The USA Patriot Act further delineates the rules governing the collection and use of data about U.S. citizens and visitors.
- Laws dating back to the 1950s cover disclosure of telephone and cable records.
- The Video Privacy Protection Act, passed in 1988, requires video stores to provide consumers with the opportunity to opt out from mailing lists that might be sold to other businesses.
- The Health Insurance Portability and Accountability Act (HIPAA) of 1996 addresses the security and privacy of medical records.
- The Gramm-Leach-Bliley Act of 1999 requires financial institutions to establish security standards that protect customer data from internal and external threats, including unauthorized access that occurs through networks and online systems.

16. List the steps you can take to protect your privacy and identity when working with databases.
See pages 658-659.

- Be aware of any activities that can possibly collect information about you for a database, such as registering to use a Web site, participating in a survey, submitting your resume to an online data bank, sending for a product rebate, participating in online discussion groups, clicking pop-up ads, and so on.
- When asked for information, supply only what's required.
- Never divulge your Social Security number and be cautious about other personal information, such as your telephone number or address.
- When in doubt about how your data might be used, you might consider using phony data.
- Make sure you look for the opt-out button if you don't want your data distributed to third parties.
- Consider using a USB flash drive containing a portable password manager that stores strongly encrypted passwords and enters them when you want to access a password-protected database or site, but can be removed from your computer when it is not in use to foil intrusion attempts
- Be wary of offers for free services in exchange for personal information or permission to track your online footprints.
- Make sure your antivirus software is running and up-to-date.
- Use a personal firewall and pop-up blocker to protect your computer from unauthorized
- intrusions.
- Never send personal information in response to an e-mail request that could be a phishing attack from an illegitimate source.
- When a privacy policy exists, check it out—especially the part that explains whether your data might be shared with a third party

Concept Map

a. Structured
b. Unstructured
c. Flat
d. Database
e. relational
f. Relational

g. Object, Object-oriented, Object oriented
h. Table
i. Record, Tuple
j. Field, Attribute
k. Methods, Method

Projects

CRITICAL THINKING

C11 Type your name into the Google search engine. How much information can you currently find about yourself? How much privacy are you willing to release? Create a table with rows of the following: Name, address, e-mail address, SS#, credit card #, phone number, height, weight, gender, hobbies, medical history, employer, spouse name, children's names. Put the following columns across the top: Applying for a job, visiting the doctor, subscribing to a Web site, making an online purchase, opening a bank account, responding to an e-mail message, registering for classes. Add two more situations to the list. Use a check mark or "OK" to indicate if you would be willing to provide each type of information. Use a question mark when you're not sure if the information is required.

Deliverables:
A table like the following:

	Apply for a job	Visit the doctor	Subscribe to a Web site	Make an online purchase	Join a chat room	Take an online survey	Register for classes	**
Name	OK	OK	OK			OK	OK	
Address	OK	OK					OK	
E-mail address	OK	OK	OK			OK	OK	
Social Security number							?	
Credit card number		OK					OK	
Phone number	OK	OK					OK	
Height		OK						
Weight		OK						
Gender		OK				OK		
Hobbies								
Medical history		OK						
Employer		OK						
Spouse's name		OK						
Children's names								

** Student should add two situations, such as:
Making a flight reservation
Applying for a passport

Grading suggestion: Pass/no pass

GROUP

C11 Form a group of four to five students. Suppose you must design the record structure for a new information system that holds drivers' license data. Using your own driver licenses, design an effective record structure, indicating record types, field names, field lengths, and data types. Where appropriate, indicate ways that you could minimize data entry errors. Submit your record structure to your instructor.

Deliverables:
A record structure for driver's licenses

Grading suggestion:
For each field, award 1 point if the field name is descriptive and within the 10-character limit; if the field length seems appropriate; and if the data type is accurately identified. Subtract 5 points from the total if any required field is missing.

CYBER CLASSROOM

C11 Look at the Internet Movie Database (*www.imdb.com*) and search for your favorite movie. Examine the field names and content (i.e., title, genre, tagline) imdb uses to organize data about movies. How would you distribute these fields among tables in a relational database? Create the structure for a database that you think duplicates the fields for the Internet Movie Database. You can use Microsoft Access or similar relational database management software. Send your structure to one other member of your team who will try to enter the data for at least three movies (you can limit the number of actors to five per movie). Your team partner should also send you comments for changes or improvements. When you are satisfied with the structure of your database, send it to your instructor.

Instructional tip: You might want to spend a few minutes familiarizing students with available DBMS software. A movie database should be relational. Depending on the sophistication of your students, you might want to limit it to a flat file. Alternatively, you could ask students to design the relational tables, but not implement them. You might also want to limit the assignment to the Overview, Cast, Details, and Technical Specs sections of the IMDB database.

Deliverables:
Plan for a movie database similar to imdb.com
Database structure created in DBMS software

Grading suggestion: Based on the parameters of the assignment (see instructional tip above)
Check the database before grading assignments. Periodically, IMDB changes its format.

MULTIMEDIA

C11 Create a multimedia database. At least two of the field types should be defined to hold media, such as photos, sounds, narration, or video. Set up a prototype of your database using software, such as Microsoft Access. You can create the media or download it from the Web, giving appropriate citations for its source. Populate your database with at least five records. Follow your instructor's guidelines for submitting your database.

Instructional suggestion: Provide your students with access to database software, such as Microsoft Access, that allows the use of BLOBs. A demonstration of how to use the software might be in order. This assignment can be given in conjunction with a database application lab.

Deliverables:
A database in which at least two of the fields hold media, such as photos, sounds, narration, or video. The database must contain at least five records.

Grading suggestion: Pass if the assignment is completed as specified.

RESUME BUILDER

C11 When you look for a job, you might have query letters and resumes out to several employers. To track the status of each, you can use a database. Use database or spreadsheet software to create a job-hunting database that includes company contact information plus fields to indicate the status of your inquiry. Load the database with at least five sample contacts. Make sure you can use the database to generate mailing labels. You should also be able to make queries to determine which job prospects are still "active" possibilities. Follow your instructor's guidelines for submitting your database and mailing labels.

Deliverables: A database with at least five job contacts; a set of mailing labels (electronic or paper—you can specify)

Grading suggestion:
60% for submitting a database
10% if the database contains at least 5 records
20% for a set of mailing labels
5% if the database has a field that indicates if the prospect is "active." Typically, you would expect a status field that could contain "active," "rejection," and so on.
5% for good data structure: correct data types, descriptive field names, appropriate field lengths

GLOBALIZATION

C11 Computer databases often store personal information about the citizens of more than one country, yet privacy expectations and laws differ. For this project, explore the differences in privacy laws around the globe. Which countries have the strongest privacy laws and which have the weakest? Which laws would govern the data stored by multinational companies? As a global consumer, which databases would concern you most for potential privacy violations? Consolidate your thoughts into a one- to two-page paper and submit it to your instructor.

Deliverables: A document about privacy law differences worldwide

The Information Shield site has a list of privacy laws by country here:
http://www.informationshield.com/intprivacylaws.html

EUguidelines for computerized data files can be found here:
http://ec.europa.eu/justice_home/fsj/privacy/instruments/un_en.htm

Google's high-profile call for global privacy standards contains some international comparisons and is detailed here:
http://peterfleischer.blogspot.com/2007/09/need-for-global-privacy-standards.html

A map showing the leading surveillance societies can be found here, along with ranking on various categories:
http://www.privacyinternational.org/article.shtml?cmd[347]=x-347-545269

Grading suggestion:
A: Student obviously examined many components and made some coherent observations about privacy laws in at least 5 countries
B: Student examined information about a few countries and made some generalizations about privacy laws
C: Student examined a few countries, had no insight about privacy laws
D: Student made a lackluster effort to look at components, did not include coherent ideas about global privacy laws

TECHNOLOGY IN CONTEXT

C11 The Technology in Context section of this chapter focused on several ways computers are used to enhance health care. It pointed out that many Web-based medical resources are available. For this project, explore the Web to find five sites: 1) a site that attempts to make a diagnosis based on a set of symptoms, 2) a site that provides information on various drugs used to treat diseases, 3) a site dedicated to providing information and patient support for a particular disease, 4) a site that evaluates medical care given at hospitals or nursing homes, and 5) a government site dedicated to health or medicine. Record the URL and name for each site, along with the organization that maintains it. Write a paragraph that describes each site's content and who would find the site most useful.

Making the assignment: Students should read the Computers in Context section before beginning this project. The project is best assigned as homework.

For this project, students will explore the Web to find five sites as noted in the assignment.

Make sure students understand the types of sites they are required to find. You might want to provide some context for their search. For example, #4 asks for a site that evaluates medical care given at hospitals or nursing homes. You might tell your students, "Suppose one of your grandparents no longer wants to live at home. Where could you find information that would help you find the best facility?"

Tell students how to submit the project. This assignment can be easily e-mailed to you.

Deliverables: A document that includes the URLs for five health care Web sites, plus a one-paragraph description of each site.

Grading:

Allow 10 points for each URL and its description. Criteria for each include:

- Site provides the information specified in the assignment
- URL is accurate
- Site description is clear and well written

Issue Try It!

1. As yet, there is not a fail-proof method for automatic terrorist identification. Suppose that a researcher devises an algorithm that is correct 99% of the time. For every 100 names it produces, one will be a false positive—someone who is not a terrorist. It will also miss one person who is a terrorist. If this algorithm is used to analyze the 308,745,538 people in the U.S., how many innocent people would be falsely identified as terrorists and how many real terrorists would go undiscovered?

3,087,455 innocent people will be identified as terrorists. Assuming that there are 1,000 real terrorists, the algorithm will miss 100 of them.

2. People leave personal digital tracks whenever they make a purchase, take a trip, access their bank account, make a phone call, file an income tax return, stroll past a security camera, obtain a prescription, mail a package, apply for a loan, e-mail a friend, rent a video, or download music. Create a fictitious person and provide five examples of online activities he or she might perform that would look suspicious to a data mining algorithm, but are actually perfectly innocent activities.

Your answers will vary. Here are some:

Sending and receiving e-mail from a relative in a country that is a known sponsor of terrorists

Sending an e-mail message like "Have you seen the new class President; he's da bomb!"

Researching Iran by going directly to its government Web sites

Researching terrorists by going to their Web sites

Having lots of Facebook friends in Yemen

Doing research for a fiction book about terrorists

3 After a flurry of post-9/11 controversies over data mining, privacy advocates were relieved to see articles in *Wired* magazine and several mainstream press sources citing a government-funded study on data mining that concluded it is not an effective tool against terrorism. Look through the executive summary for the 2008 study "Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment." Find and copy the passage stating that data mining is not an effective tool for combating terrorism.

The report contains the following statement at the bottom of page 3:

"Automated identification of terrorists through data mining (or any other known methodology) is neither feasible as an objective nor desirable as a goal of technology development efforts."

However, the recommendations of the report do not include any wording that the government desist from data mining. On the contrary, it assumes that information-based counterterrorism programs will be deployed. The report recommends only that these programs be periodically reviewed.

4 To find out how extensively your government is using data mining, you can go directly to the source. For example, in the U.S., the Department of Homeland Security Privacy Office issues an annual Data Mining Report to Congress. Read the most recent report. List and briefly describe the data mining projects that are operational.

The 2010 report covers the following projects:

(1) the Automated Targeting System (ATS) Inbound, Outbound, and Passenger modules administered by U.S. Customs and Border Protection (CBP); (2) the Data Analysis and Research for Trade Transparency System (DARTTS) administered by U.S. Immigration and Customs Enforcement (ICE); and (3) the Freight Assessment System (FAS) administered by the Transportation Security Administration (TSA).

The report also states:

"While each of the programs described below engages to some extent in data mining, none uses data mining to make unevaluated automated decisions about individuals. These programs do not make decisions about individuals solely on the basis of data mining results. In all cases, DHS employees conduct investigations to verify (or disprove) the results of data mining, and then bring their own judgment and experience to bear in making determinations about individuals initially identified through data mining activities.

Which in real terms means that when data mining identifies an individual, that person is investigated (without thier knowledge).

The report is also notable for its extensive use of acronyms and its extensive verbiage about compliance regulations that require the DHS to publish the report. The actual amount of information about surveillance programs is quite sparse.

5 Technology is increasing the amount and type of data that is available to data mining operations. Although data mining originally was designed to analyze text data, advances in facial recognition and translation provide multimedia input for data mining algorithms. You can test facial recognition systems at online sites that find celebrities with facial features similar to yours. Upload a photo (it doesn't have to be yours) at a celebrity look-alike site such as Picadilo.com, MyHeritage.com, or FaceDouble. Based on the accuracy of finding similar faces, how efficient would this system be for identifying a terrorist based on his or her photo?

You probably found that some of the images weren't even close to the original photo. However, a little additional research reveals that facial recognition is a viable law enforcement tool. See the Wikipedia page on Facial Recognition and http://www.nist.gov/itl/iad/ig/frgc.cfm

"In 2006, the performance of the latest face recognition algorithms were evaluated in the Face Recognition Grand Challenge (FRGC). High-resolution face images, 3-D face scans, and iris images were used in the tests. The results indicated that the new algorithms are 10 times more accurate than the face recognition algorithms of 2002 and 100 times more accurate than those of 1995. Some of the algorithms were able to outperform human participants in recognizing faces and could uniquely identify identical twins."

Information Tools Try It!

1 Although forums are sometimes billed as "advice and recommendations from a community of experts," participants are not always experts and the advice offered is not always accurate. The first step to identifying the source of information is to click any links available to the user's profile. Select an article from a technology news source, such as Engadget or HuffPost Tech, that includes lots of comments. Click the picture and name links for three commenters, then fill in the following table, entering "NA" if the information is not available.

Answers will vary, depending on the commenters selected.

2 Verifying tweets requires a twofold approach: gathering information on the tweeter, and triangulating the location and timing of the content. To see how it works, check out a technology Twitter feed. You can access tweets even without a Twitter account by googling *twitter search* and then using the Search box at the top of the screen. Enter *Cisco* to find information about the technology company that specializes in network equipment. Select a tweet that appears to come from an individual, then answer the following questions:

a. Does the tweeter provide a name, picture, and biography?

b. If you search for the tweeter's name in Google, do you find further clues to the person's identity?

c. What is the person's track record of blog posts, forum responses, or tweets?

d. Does the source have a large following on Twitter or an extensive network of friends on Facebook, LinkedIn, or Google+?

e. Can you determine where the source is located and if the location might be relevant to verifying the information (as in an eyewitness account)?

f. Are other sources reporting similar information?

g. If the information is a link or retweet, can you identify the original source?

h. Does the language of the message sound appropriate for the source?

Answers will vary, depending on the Tweeter selected.

3 When you're looking for a solution to a technical problem and you come across forum postings such as "I don't know" or "I've never tried this but...," you probably wonder why anyone would waste his or her time (and yours) with information that is totally useless. In addition to lame comments, online information sources are plagued by trolls and sockpuppets. Look for information about "trolls" and "sockpuppets" in Wikipedia. Make a list of characteristics that can help you stay alert for trolls and sockpuppets on Twitter feeds, forums, and other online sources.

According to Wikipedia, "a sockpuppet is an online identity used for purposes of deception." They are frequently used to make comments on blogs and other interactive sites. For example, a blogger might use several sock puppets to make comments about his or her blog to make it appear that many people agree with his or her viewpoint. Sockpuppets are also used to make snide comments when the commenter does not want to be identified.

Although administrators can access IP data to track down sockpuppets, participants in chat sessions and blogs have only contextual clues to identify sockpuppets. Characteristics that might make one suspicious that a sockpuppet is in use:

Writing style similar to the blogger or to other commenters.

Phrasing similar to the blogger.

Grammar or punctuation errors similar to the blogger.

Unable to trace identity to a Facebook page or personal Website

Strong viewpoints repeatedly expressed