Introduction to databases part 1

- Databases: The Unseen Services
- Databases: The Ubiquitous Information Provider
- Database Skills for Knowledge-based Careers
- The Database Approach
- Information Structuring in Databases (in part 2)

Learning Objectives

After reading this chapter, you will be able to:

- Understand the importance of databases in information access
- Describe the characteristics of information
- Explain the major components of a DBMS
- Describe the need for database knowledge in business careers

Databases: The Unseen Services

In today's modern societies, we depend on many services from organizations and governments that we take for granted. Often we are not even aware of the underlying and invisible components that make these services available to us. For example, we live in houses, work and go to school in buildings, exercise in gymnasiums, and go to sports activities in stadiums. We observe and interact with the visible components of these edifices, but we seldom think about the unseen components that provide us comfort and stability.

Just as edifices contain both visible structure and an unseen infrastructure, the services that we use extensively every day also contain two separate components. We use our smartphones for all types of activities. We also received services from our computers with the Internet or in our education activities. We receive services from people and companies as we shop, go to the doctor, or get a driver's license. The unseen component of all these activities is the databases that provide the support in our information rich, connected, and mobile world. Of course, almost everything we do with an electronic device depends on some underlying databases. In most cases, we do not even think about the source and format of the information that we are retrieving.

Let us examine some examples.

- **Login and use Facebook, Twitter or Instagram**: Your login and personal information is kept in a database, as is the information for all of your friends, posts, and messages.

- **Make a phone call or send a text message:** The lists of phone numbers and locations are maintained by databases. All the information about your phone calls and messages is maintained at a very detailed level in databases.
- **Read your email:** Not only is your login information kept in a database, but also all the history of incoming and outgoing emails is kept in very large databases.
- **Go shopping (Groceries, clothes, gasoline):** First, the product, inventory levels, and price information is kept in a database so that the checkout register can identify it correctly. Then if you use a credit or debit card, that information must be retrieved from another database.
- **Check a book out of the library:** The database is checked to make sure you are a valid patron of the library. Then information about the book, its availability, and check out status is all verified and updated for your library loan. The database is used to obtain both patron information and book information.
- **Withdraw money from an ATM**: First, your authorization information must be verified, and then the availability of sufficient funds is verified. Finally, your account balances must be updated. All this information is stored in the bank's databases.
- **Have a physical exam with your doctor:** All of your patient information is maintained in your medical database. Your records will be updated from the results of your exam. The accounting, billing, and insurance records will also all be updated.

We seldom think about these databases, unless they become unavailable or there has been a security breach where private data has been accessed by unauthorized individuals or been made public. When they do not work correctly, we get frustrated and upset that the level of service that we expect is not being provided. The availability and integrity of databases is an important issue in the daily activities of each of our lives.

Databases: The Ubiquitous Information Provider

We are all aware that we live in an information age. In fact, we could say that there are three major characteristics of today's society: pervasive connectivity, universal mobility, and abundant information. Pervasive connectivity and universal mobility enable the abundance of information that is available. However, the primary contributor in our access to abundant information is database technology. Let's see how databases support our need to obtain in-depth information.

- **Searching to buy a specific item (e.g. an appliance, a service, clothing, a book...):** We have all used an Internet search provider. The information provided is usually a list of the particular item for which we are searching as identified by a key word. It is mind boggling to think of the size and complexity of the databases that provide indexes to find what we want. The amount of data that must be stored, searched, and retrieved to answer a simple query for a product or service in a particular location is staggering.
- **Researching more in-depth before purchasing a major item such as a car: The information you need for this includes items like repair history,**

performance data, and safety record. You may also be interested in other people's opinions. This information can be obtained in such places as Consumer Reports, which is maintained in databases. You might find people's experiences in ratings or comments or even some blogs; these are provided by databases.

- **Planning a vacation or trip: In this case, we want geographical and resort information – information about places and things to do**. We also seek for information about travel arrangements including schedules, costs, and availability. The information systems may have to access multiple databases – airline databases, detailed resort or accommodations databases, sites databases, activities databases, and even geographical databases for road maps.
- **Learning about a particular hobby or sport that you enjoy such as skiing, boating, or geocaching**: Almost all hobbies and sports have organizations and clubs that generate articles and conferences on that sport. The articles and information are largely stored in databases maintained by these organizations.
- **Researching a particular company before investing: Many institutions maintain databases of financial records**. There are databases that provide detailed histories of every publicly traded company. Some of these databases are publicly available, but even more details are available through private databases maintained by individual financial services companies.
- **Researching a particular medical condition or medical problem:** There are also extensive databases with information about various illnesses, medications and treatments. Many individuals also comment and write blogs about personal experiences with particular illnesses. Of course, not all of this information is valuable and correct. However, databases in almost every case provide the information repository.
- **Solving a technical problem or learning a new technical skill such as writing SQL queries:** One of the interesting phenomena in today's technical world is that for almost every problem that a developer or technical person encounters, someone else has already encountered it and there is a solution or additional training available. Most of these answers are on forums, blogs, or on-line documentation. All of these information sources are maintained by databases.
- **Writing a research paper or researching for a school project**: Although many research articles and books are not stored in databases per se, keyword and topic search capability is always supported by search engine databases. In addition, depending on the research project, databases are available from newspapers, research organizations, and on-line libraries.
- **Wikipedia**: Who of us has not used Wikipedia from time to time to get a quick understanding of a particular subject? Wikipedia, as well as the keyword indexes, are maintained in a database.

Database Skills for Knowledge-based Careers

It should be apparent from this section and the previous one that database impact our lives in a multitude of ways in almost everything we know and do. The ability to understand and retrieve data from databases is an important and powerful skill to have.

It should be evident from the above examples that an understanding and knowledge of how to use, design, and build databases is essential for anyone desiring a career in information systems. The core of every type of information system is a database.

Every type of technical position requires extensive or substantial database skills. The database skills required range from being able to design and build a database to being able to monitor, tune, and optimize the database performance. However, database skills are not only required by technical people. Every type of knowledge-based career also requires substantial database knowledge and skill.

Table 1-1 identifies a few other types of knowledge-based careers. A knowledge-based career is one that requires an understanding of how to solve problems using knowledge and analytical skills. A very large percentage of knowledge-based skills now require computer proficiency, along with the ability to understand and use databases to find and extract information. As can be seen from the table, knowledge workers must be able to understand the structure of the data in the database so that they can extract information from the raw data.

Table 1-1
Database skills required for knowledge-based careers.

Job Title	Job Description	DB Knowledge Required	Skill Level Required
Financial planner	Analyze security market and recommend financial portfolios	Understand financial information available in databases. Know how to extract information in various combinations	Moderate
Market research analyst	Analyze sales and economic data to predict future trends	Understand economic databases. Be able to use databases in novel ways to discover and extract information.	Substantial
Advertising manager	Manage advertising campaigns and budgets	Understand sales and performance data. Be able to extract information in various ways.	Moderate
Human resource manager	Manage all the hiring, evaluation, and monitoring of the work force.	Understand information about employees as well as all financial information.	Moderate

Job Title	Job Description	DB Knowledge Required	Skill Level Required
Accountant	Support all the accounting requirements of an organization	Understand all financial databases. Identify problems or potential fraudulent activity through database analysis	Extensive
Economist	Analyze and predict economic trends based on historical data	Understand economic databases, their structure and data. Be able to extract information in various forms using novel approaches.	Extensive
Sales manager	Manage the sales persons and sales activities of an organization	Understand sales data and be able to extract information in various forms. Evaluate performance.	Moderate
Sociologist	Analyze social trends and social problems in communities and nations	Understand demographic and social data from diverse and disparate databases. Be able to extract information using multiple techniques	Extensive
Management consultant	Analyze organizational issues and recommend solutions	Understand financial and operational data of an organization as found in various databases	Substantial
Public Administrator	Manage a public government unit	Be able to understand data and information provided in various and diverse databases from the governmental database.	Extensive
Political Scientist	Understand and research demographic data and trends	Be able to analyze data and draw conclusions from demographic databases and questionnaire data.	Substantial

The Database Approach

An information system that uses a Database Management System (DBMS) to manage its information has a particular structure, comprising three components: Data, DBMS, and Application software. This structure as described below is referred to as the database approach to information system development.

The central component of the database approach is the DBMS. This software is also referred to as the “database engine” or the “back end.” With regard to the data it manages, it has several responsibilities including the following:

- Data Definition: providing a way to define and build the database
- Data Manipulation: providing a way to insert and update data in the database
- Query Execution: retrieving information from the data in the database
- Data Integrity: ensuring that data stored is well formed
- Data Security: enforcing restrictions about who is able to access what data
- Provenance: logging capabilities to provide an audit trail for data changes
- Multiuser Concurrency: supporting the activities of many users at the same time

As can be seen from the above list, a DBMS is a complex software application. While all database management systems may not provide all of these features, these are the general characteristics of today’s DBMSs. Using a database requires considerable expertise and knowledge about the specific DBMS being used. Some of the more popular DBMS's today are MySQL, Microsoft SQL Server, Oracle, PostgreSQL, Microsoft Access, and IBM's DB2.

The second component in the database approach is the data. Although the physical location or manner in which the data are stored may be important for performance reasons, the location of the data does not determine whether a system is developed using the database approach. As long as the DBMS has access and is able to perform its responsibilities with respect to the data, the details of the data storage are not relevant.

The final component of the database approach is the application, also called "front-end" software. Application software interacts with the DBMS to provide information to a user. It may also provide a way for a user to invoke other functionality of the DBMS. In fact, the DBMS software itself is non-visual, meaning that the user does not interact directly with the DBMS. Any software that provides an interface for user to invoke procedures in the DBMS we will defined as application software.

Once the application has determined what the user is trying to accomplish, it sends a request to the DBMS. The request may be an instruction to change data or a request for information such as the list of employees who were hired on a particular date. All relational databases use a standard language to receive and process requests. The standard language is called Structured Query Language (SQL).

The DBMS receives the request and determines if the operation requested is allowed for the authenticated user. If the operation is allowed, the DBMS completes the operation and sends a response to the application. The application then communicates the information to the user. If the operation is not authorized for the user or if there is an error in fulfilling the operation, the DBMS responds with an appropriate message. Again, it is up to the application to display that to the user. It is a critical feature of the database approach that the application never bypasses the DBMS to access stored data directly.

Figure 1-1 illustrates some of the primary components of a typical DBMS and how they are used in an information system. The User interacts with the DBMS generally by writing SQL statements through the front end. (Although a sophisticated front end could format the SQL statements itself based on other types of user input.) These SQL statements are interpreted and executed by the DBMS by either updating the data or by returning results from the data. In this class, we will focus on query statements, whose purpose is to retrieve data from the database and present it in a form that is understandable by the User.

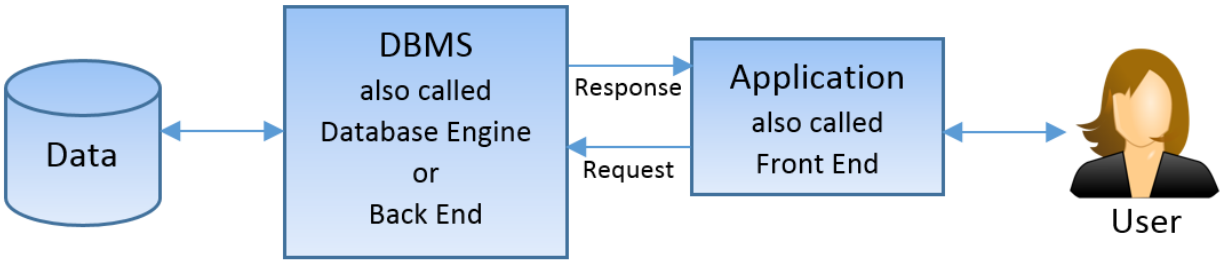

Figure 1-1: The database approach