

Objectives

In this chapter, you will learn:

- The difference between data and information
- What a database is, the various types of databases, and why they are valuable assets for decision making
- The importance of database design
- How modern databases evolved from file systems

Database Systems, 10th Edition2

Objectives (cont'd.)

- About flaws in file system data management
- The main components of the database system
- The main functions of a database management system (DBMS)

Database Systems, 10th Edition3

Introduction

- Good decisions require good information derived from raw facts
- Data is managed most efficiently when stored in a database
- Databases evolved from computer file systems
- Understanding file system characteristics is important

Database Systems, 10th Edition

4

Why Databases?

- Databases solve many of the problems encountered in data management
 - Used in almost all modern settings involving data management:
 - Business
 - Research
 - Administration
- Important to understand how databases work and interact with other applications

Database Systems, 10th Edition

5

Data vs. Information

- Data are raw facts
- Information is the result of processing raw data to reveal meaning
- Information requires context to reveal meaning
- Raw data must be formatted for storage, processing, and presentation
- Data are the foundation of information, which is the bedrock of knowledge

Database Systems, 10th Edition

6

Data vs. Information (cont'd.)

- Data: building blocks of information
- Information produced by processing data
- Information used to reveal meaning in data
- Accurate, relevant, timely information is the key to good decision making
- Good decision making is the key to organizational survival
- Data management: focuses on proper generation, storage, and retrieval of data

Database Systems, 10th Edition

7

Introducing the Database

- Database: shared, integrated computer structure that stores a collection of:
 - End-user data: raw facts of interest to end user
 - Metadata: data about data
 - Provides description of data characteristics and relationships in data
 - Complements and expands value of data
- Database management system (DBMS): collection of programs
 - Manages structure and controls access to data

Database Systems, 10th Edition

8

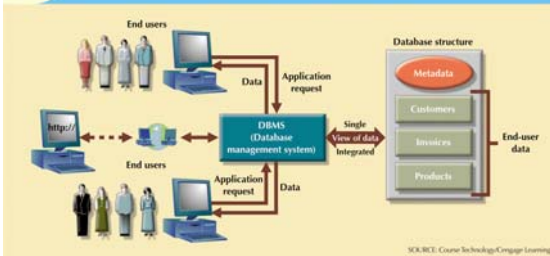
Role and Advantages of the DBMS

- DBMS is the intermediary between the user and the database
 - Database structure stored as file collection
 - Can only access files through the DBMS
- DBMS enables data to be shared
- DBMS integrates many users' views of the data

Database Systems, 10th Edition

9

FIGURE 1.2 The DBMS manages the interaction between the end user and the database



Database Systems, 10th Edition

10

Role and Advantages of the DBMS (cont'd.)

- Advantages of a DBMS:
 - Improved data sharing
 - Improved data security
 - Better data integration
 - Minimized data inconsistency
 - Improved data access
 - Improved decision making
 - Increased end-user productivity

Database Systems, 10th Edition

11

Types of Databases

- Databases can be classified according to:
 - Number of users
 - Database location(s)
 - Expected type and extent of use
- Single-user database supports only one user at a time
 - Desktop database: single-user; runs on PC
- Multuser database supports multiple users at the same time
 - Workgroup and enterprise databases

Database Systems, 10th Edition

12

Types of Databases (cont'd.)

- Centralized database: data located at a single site
- Distributed database: data distributed across several different sites
- Operational database: supports a company's day-to-day operations
 - Transactional or production database
- Data warehouse: stores data used for tactical or strategic decisions

Database Systems, 10th Edition

13

Types of Databases (cont'd.)

- Unstructured data exist in their original state
- Structured data result from formatting
 - Structure applied based on type of processing to be performed
- Semistructured data have been processed to some extent
- Extensible Markup Language (XML) represents data elements in textual format
 - XML database supports semistructured XML data

Database Systems, 10th Edition

14

TABLE 1.1 Types of Databases

PRODUCT	NUMBER OF USERS			DATA LOCATION		DATA USAGE		XML
	SINGLE USER	WORKGROUP	ENTERPRISE	CENTRALIZED	DISTRIBUTED	OPERATIONAL	ANALYTICAL	
MS Access	X	X		X		X		
MS SQL Server	X ¹	X	X	X	X	X	X	X
IBM DB2	X ²	X	X	X	X	X	X	X
MySQL	X	X	X	X	X	X	X	X
Oracle RDBMS	X ³	X	X	X	X	X	X	X

Database Systems, 10th Edition

15

Why Database Design Is Important

- Database design focuses on design of database structure used for end-user data
 - Designer must identify database's expected use
- Well-designed database:
 - Facilitates data management
 - Generates accurate and valuable information
- Poorly designed database:
 - Causes difficult-to-trace errors

Database Systems, 10th Edition

16

Evolution of File System Data Processing

- Reasons for studying file systems:
 - Complexity of database design is easier to understand
 - Understanding file system problems helps to avoid problems with DBMS systems
 - Knowledge of file system is useful for converting file system to database system
- File systems typically composed of collection of file folders, each tagged and kept in cabinet
 - Organized by expected use

Database Systems, 10th Edition

17

Evolution of File System Data Processing (cont'd.)

- Contents of each file folder are logically related
- Manual file systems
 - Served as a data repository for small data collections
 - Cumbersome for large collections
- Computerized file systems
 - Data processing (DP) specialist converted computer file structure from manual system
 - Wrote software that managed the data
 - Designed the application programs

Database Systems, 10th Edition

18

Evolution of File System Data Processing (cont'd.)

- Initially, computer file systems resembled manual systems
- As number of files increased, file systems evolved
 - Each file used its own application program to store, retrieve, and modify data
 - Each file was owned by individual or department that commissioned its creation

FIGURE 1.5 Contents of the CUSTOMER file

C_NAME	C_PHONE	C_ADDRESS	C_ZIP	A_NAME	A_PHONE	TP	AMT	REN
Albert A. Egan	815-844-2573	218 Fork Rd., Bala, TN	38122	Leslie F. Harte	815-852-1244	T1	100.00	05-Apr-2012
Leanne K. Dwyer	713-884-1238	Box 12A, Fox, KY	25246	Alan D. Abby	713-225-1249	T1	200.00	18-Jun-2012
Paul H. Gresh	815-884-2288	125 Oak Ln., Bala, TN	38122	Leslie F. Harte	815-852-1244	S2	150.00	28-Jun-2012
Paul F. Chowell	815-884-2180	217 Lee Ln., Bala, TN	38122	Leslie F. Harte	815-852-1244	S1	300.00	14-Oct-2012
Myra Cole	815-225-1672	Box 111, Fox, TN	80395	Alan D. Abby	713-225-1249	T1	100.00	25-Oct-2012
Angie B. O'Brien	713-440-3301	387 Todd Dr., Fox, KY	25246	John T. Olson	815-123-5689	T2	950.00	22-Sep-2012
James G. Brown	815-267-1238	21 Top Rd., Bala, TN	37119	Leslie F. Harte	815-852-1244	S1	120.00	26-Mar-2013
George Williams	815-267-2008	186 Maple, Bala, TN	37119	John T. Olson	815-123-5689	S1	200.00	17-Jul-2012
Anna G. Farris	713-362-7185	2109 Elm, Cree, KY	25432	Alan D. Abby	713-225-1249	T2	100.00	13-Dec-2012
Cheryl K. Smith	815-267-3889	2052 Main, Bala, TN	37119	John T. Olson	815-123-5689	S2	300.00	14-Mar-2013

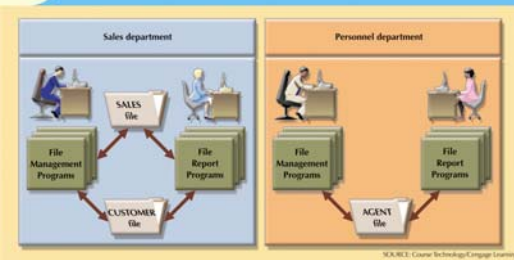
C_NAME = Customer name
C_PHONE = Customer phone
C_ADDRESS = Customer address
C_ZIP = Customer zip code
A_NAME = Agent name
A_PHONE = Agent phone
TP = Insurance type
AMT = Insurance policy amount, in thousands of \$
REN = Insurance renewal date

SOURCE: Course Technology/Cengage Learning.

TABLE 1.2 Basic File Terminology

TERM	DEFINITION
Data	Raw facts, such as a telephone number, a birth date, a customer name, and a year-to-date (YTD) sales value. Data have little meaning unless they have been organized in some logical manner.
Field	A character or group of characters (alphabetic or numeric) that has a specific meaning. A field is used to define and store data.
Record	A logically connected set of one or more fields that describes a person, place, or thing. For example, the fields that constitute a customer record might consist of the customer's name, address, phone number, date of birth, credit limit, and unpaid balance.
File	A collection of related records. For example, a file might contain data about the students currently enrolled at Gigantic University.

FIGURE 1.7 A simple file system



Database Systems, 10th Edition

22

File System Redux: Modern End-User Productivity Tools

- Ubiquitous use of personal productivity tools can introduce the same problems as the old file systems
- Microsoft Excel
 - Widely used by business users
 - Users have become so adept at working with spreadsheets, they tend to use them to complete tasks for which spreadsheets are not appropriate – database substitute

Database Systems, 10th Edition

23

Problems with File System Data Processing

- File systems were an improvement over manual system
 - File systems used for more than two decades
 - Understanding the shortcomings of file systems aids in development of modern databases
 - Many problems not unique to file systems
- Even simple file system retrieval task required extensive programming
 - Ad hoc queries impossible
 - Changing existing structure difficult

Database Systems, 10th Edition

24

Problems with File System Data Processing (cont'd.)

- Security features difficult to program
 - Often omitted in file system environments
- Summary of file system limitations:
 - Requires extensive programming
 - Cannot perform ad hoc queries
 - System administration is complex and difficult
 - Difficult to make changes to existing structures
 - Security features are likely to be inadequate

Database Systems, 10th Edition

25

Structural and Data Dependence

- Structural dependence: access to a file is dependent on its own structure
 - All file system programs must be modified to conform to a new file structure
- Structural independence: change file structure without affecting data access
- Data dependence: data access changes when data storage characteristics change
- Data independence: data storage characteristics do not affect data access

Database Systems, 10th Edition

26

Structural and Data Dependence (cont'd.)

- Practical significance of data dependence is difference between logical and physical format
- Logical data format: how human views the data
- Physical data format: how computer must work with data
- Each program must contain:
 - Lines specifying opening of specific file type
 - Record specification
 - Field definitions

Database Systems, 10th Edition

27

Data Redundancy

- File system structure makes it difficult to combine data from multiple sources
 - Vulnerable to security breaches
- Organizational structure promotes storage of same data in different locations
 - Islands of information
- Data stored in different locations is unlikely to be updated consistently
- Data redundancy: same data stored unnecessarily in different places

Database Systems, 10th Edition

28

Data Redundancy (cont'd.)

- Data inconsistency: different and conflicting versions of same data occur at different places
- Data anomalies: abnormalities when all changes in redundant data are not made correctly
 - Update anomalies
 - Insertion anomalies
 - Deletion anomalies

Database Systems, 10th Edition

29

Lack of Design and Data-Modeling Skills

- Most users lack the skill to properly design databases
 - Despite multiple personal productivity tools being available
- Data-modeling skills
 - Vital in the data design process
- Good data modeling facilitates communication between the designer, user, and the developer

Database Systems, 10th Edition

30

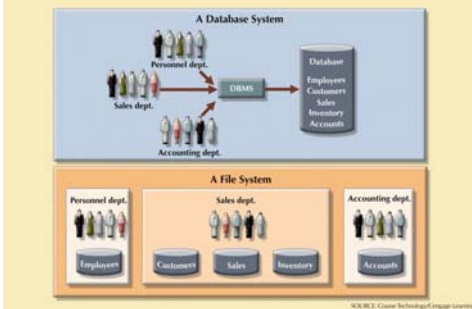
Database Systems

- Database system consists of logically related data stored in a single logical data repository
 - May be physically distributed among multiple storage facilities
 - DBMS eliminates most of file system's problems
 - Current generation stores data structures, relationships between structures, and access paths
 - Also defines, stores, and manages all access paths and components

Database Systems, 10th Edition

31

FIGURE 1.8 Contrasting database and file systems



Database Systems, 10th Edition

32

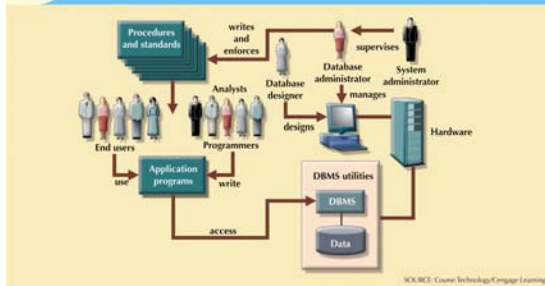
The Database System Environment

- Database system: defines and regulates the collection, storage, management, use of data
- Five major parts of a database system:
 - Hardware
 - Software
 - People
 - Procedures
 - Data

Database Systems, 10th Edition

33

FIGURE 1.9 The database system environment



Database Systems, 10th Edition

34

The Database System Environment (cont'd.)

- Hardware: all the system's physical devices
- Software: three types of software required
 - Operating system software
 - DBMS software
 - Application programs and utility software

Database Systems, 10th Edition

35

The Database System Environment (cont'd.)

- People: all users of the database system
 - System and database administrators
 - Database designers
 - Systems analysts and programmers
 - End users
- Procedures: instructions and rules that govern the design and use of the database system
- Data: the collection of facts stored in the database

Database Systems, 10th Edition

36

The Database System Environment (cont'd.)

- Database systems are created and managed at different levels of complexity
- Database solutions must be cost-effective as well as tactically and strategically effective
- Database technology already in use affects selection of a database system

Database Systems, 10th Edition

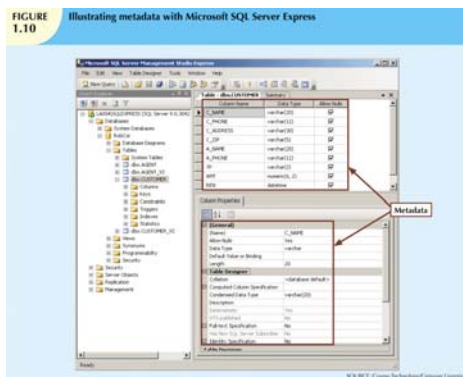
37

DBMS Functions

- Most functions are transparent to end users
 - Can only be achieved through the DBMS
- Data dictionary management
 - DBMS stores definitions of data elements and relationships (metadata) in a data dictionary
 - DBMS looks up required data component structures and relationships
 - Changes automatically recorded in the dictionary
 - DBMS provides data abstraction and removes structural and data dependency

Database Systems, 10th Edition

38



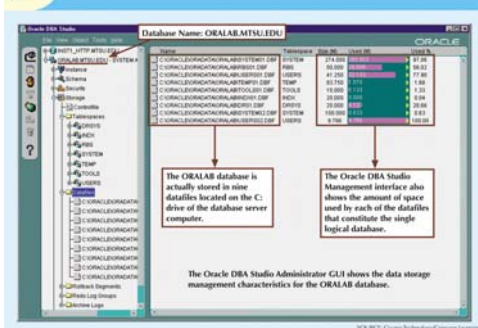
Database Systems, 10th Edition

39

DBMS Functions (cont'd.)

- Data storage management
 - DBMS creates and manages complex structures required for data storage
 - Also stores related data entry forms, screen definitions, report definitions, etc.
 - Performance tuning: activities that make the database perform more efficiently
 - DBMS stores the database in multiple physical data files

FIGURE 1.11 Illustrating data storage management with Oracle



DBMS Functions (cont'd.)

- Data transformation and presentation
 - DBMS transforms data entered to conform to required data structures
 - DBMS transforms physically retrieved data to conform to user's logical expectations
- Security management
 - DBMS creates a security system that enforces user security and data privacy
 - Security rules determine which users can access the database, which items can be accessed, etc.

DBMS Functions (cont'd.)

- Multiuser access control
 - DBMS uses sophisticated algorithms to ensure concurrent access does not affect integrity
- Backup and recovery management
 - DBMS provides backup and data recovery to ensure data safety and integrity
 - Recovery management deals with recovery of database after a failure
 - Critical to preserving database's integrity

Database Systems, 10th Edition

43

DBMS Functions (cont'd.)

- Data integrity management
 - DBMS promotes and enforces integrity rules
 - Minimizes redundancy
 - Maximizes consistency
 - Data relationships stored in data dictionary used to enforce data integrity
 - Integrity is especially important in transaction-oriented database systems

Database Systems, 10th Edition

44

DBMS Functions (cont'd.)

- Database access languages and application programming interfaces
 - DBMS provides access through a query language
 - Query language is a nonprocedural language
 - Structured Query Language (SQL) is the de facto query language
 - Standard supported by majority of DBMS vendors

Database Systems, 10th Edition

45

DBMS Functions (cont'd.)

- Database communication interfaces
 - Current DBMSs accept end-user requests via multiple different network environments
 - Communications accomplished in several ways:
 - End users generate answers to queries by filling in screen forms through Web browser
 - DBMS automatically publishes predefined reports on a Web site
 - DBMS connects to third-party systems to distribute information via e-mail

Database Systems, 10th Edition

46

Managing the Database System: A Shift in Focus

- Database system provides a framework in which strict procedures and standards enforced
 - Role of human changes from programming to managing organization's resources
- Database system enables more sophisticated use of the data
- Data structures created within the database and their relationships determine effectiveness

Database Systems, 10th Edition

47

Managing the Database System: A Shift in Focus (cont'd.)

- Disadvantages of database systems:
 - Increased costs
 - Management complexity
 - Maintaining currency
 - Vendor dependence
 - Frequent upgrade/replacement cycles

Database Systems, 10th Edition

48

Preparing for Your Database Professional Career

TABLE 1.3

Database Career Opportunities

JOB TITLE	DESCRIPTION	SAMPLE SKILLS REQUIRED
Database developer	Creates and maintains database-based applications	Programming, database fundamentals, SQL
Database designer	Designs and maintains databases	Systems design, database design, SQL
Database administrator	Manages and maintains DBMS and databases	Database fundamentals, SQL, vendor courses
Database analyst	Develops databases for decision support reporting	SQL, query optimization, data warehouses
Database architect	Designs and implements database environments (conceptual, logical, and physical)	DBMS fundamentals, data modeling, SQL, hardware knowledge
Database consultant	Helps companies leverage database technologies to improve business processes and achieve specific goals	Database fundamentals, data modeling, database design, SQL, DBMS, hardware, vendor-specific technologies
Database security officer	Implements security policies for data administration	DBMS fundamentals, database administration, SQL, data security technologies

Database Systems, 10th Edition

49

Summary

- Data are raw facts
- Information is the result of processing data to reveal its meaning
- Accurate, relevant, and timely information is the key to good decision making
- Data are usually stored in a database
- DBMS implements a database and manages its contents

Database Systems, 10th Edition

50

Summary (cont'd.)

- Metadata is data about data
- Database design defines the database structure
 - Well-designed database facilitates data management and generates valuable information
 - Poorly designed database leads to bad decision making and organizational failure
- Databases evolved from manual and computerized file systems

Database Systems, 10th Edition

51

Summary (cont'd.)

- In a file system, data stored in independent files
 - Each requires its own management program
- Some limitations of file system data management:
 - Requires extensive programming
 - System administration is complex and difficult
 - Changing existing structures is difficult
 - Security features are likely inadequate
 - Independent files tend to contain redundant data
 - Structural and data dependency problems

Database Systems, 10th Edition

52

Summary (cont'd.)

- Database management systems were developed to address file system's inherent weaknesses
- DBMS present database to end user as single repository
 - Promotes data sharing
 - Eliminates islands of information
- DBMS enforces data integrity, eliminates redundancy, and promotes security

Database Systems, 10th Edition

53
