
File Systems

Evolution of File System Data Processing

Manual File Systems

Accomplished through a system of file folders and filing cabinets



Computerized File Systems

Data processing (DP) specialist: Created a computer-based system that would track data and produce required reports



File System Redux: Modern End-User Productivity Tools

Includes spreadsheet programs such as Microsoft Excel

Basic File Terminology

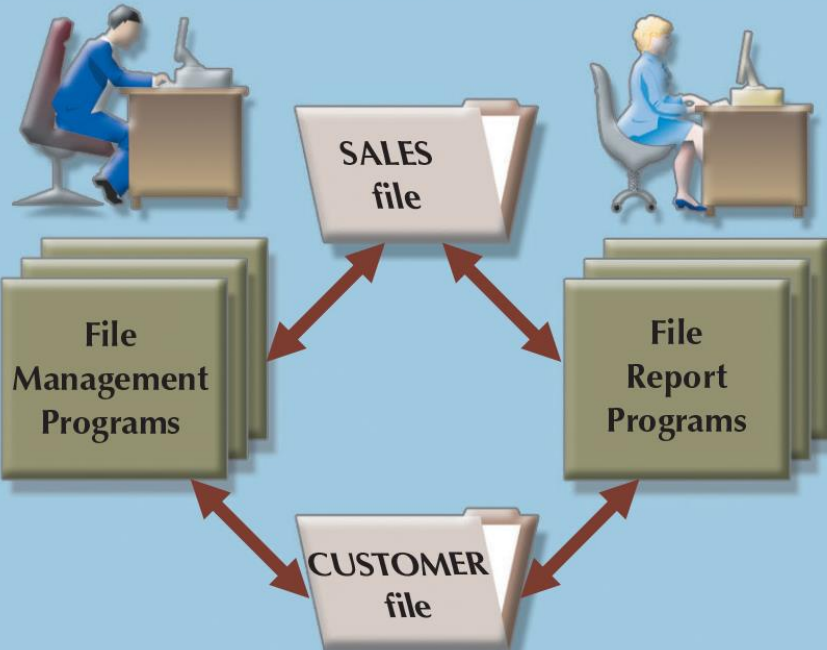
TABLE 1.2

BASIC FILE TERMINOLOGY

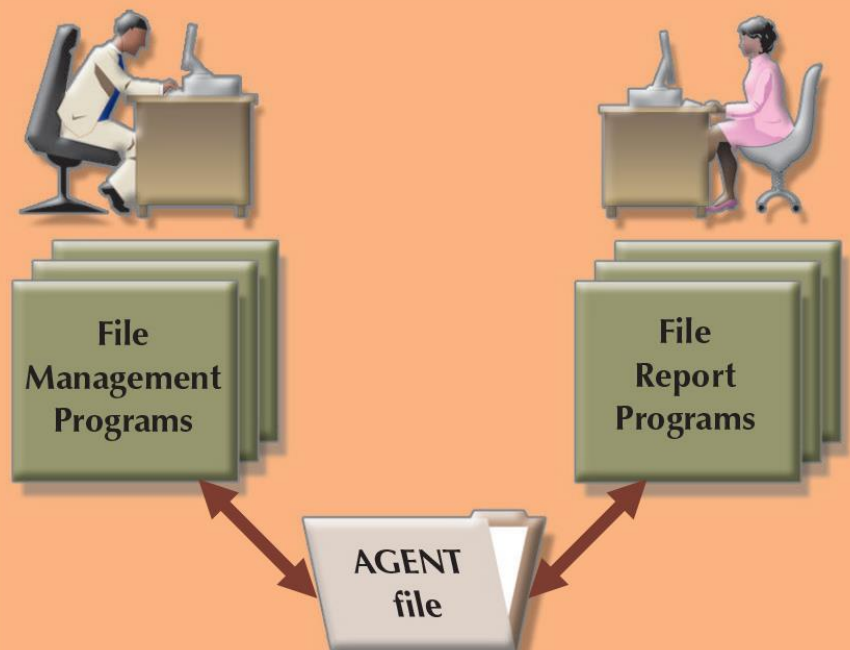
TERM	DEFINITION
Data	Raw facts, such as a telephone number, a birth date, a customer name, and a year-to-date (YTD) sales value. Data have little meaning unless they have been organized in some logical manner.
Field	A character or group of characters (alphabetic or numeric) that has a specific meaning. 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 record for a customer 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.

A Simple File System

Sales department



Personnel department



Problems with File Systems

Problems with File System Data Processing

Lengthy development times

Difficulty of getting quick answers

Complex system administration

Lack of security and limited data sharing

Extensive programming

Structural and Data Dependence

A file system may exhibit:

- **Structural dependence:** Access to a file is dependent on its own structure
 - All file system programs are modified to conform to a new file structure
 - **Structural independence:** File structure is changed without affecting the application's ability to access the data
-

Structural and Data Dependence

- Data dependence
 - Data access changes when data storage characteristics change
 - Data independence
 - Data storage characteristics is changed without affecting the program's ability to access the data
 - Practical significance of data dependence is difference between logical and physical format
-

Data Redundancy

- Unnecessarily storing same data at different places
 - **Islands of information:** Scattered data locations
 - Increases the probability of having different versions of the same data
-

Data Redundancy Implications

- Poor data security
 - Data inconsistency
 - Increased likelihood of data-entry errors when complex entries are made in different files
 - **Data anomaly:** Develops when not all of the required changes in the redundant data are made successfully
-

Types of Data Anomaly

Update Anomalies

Insertion Anomalies

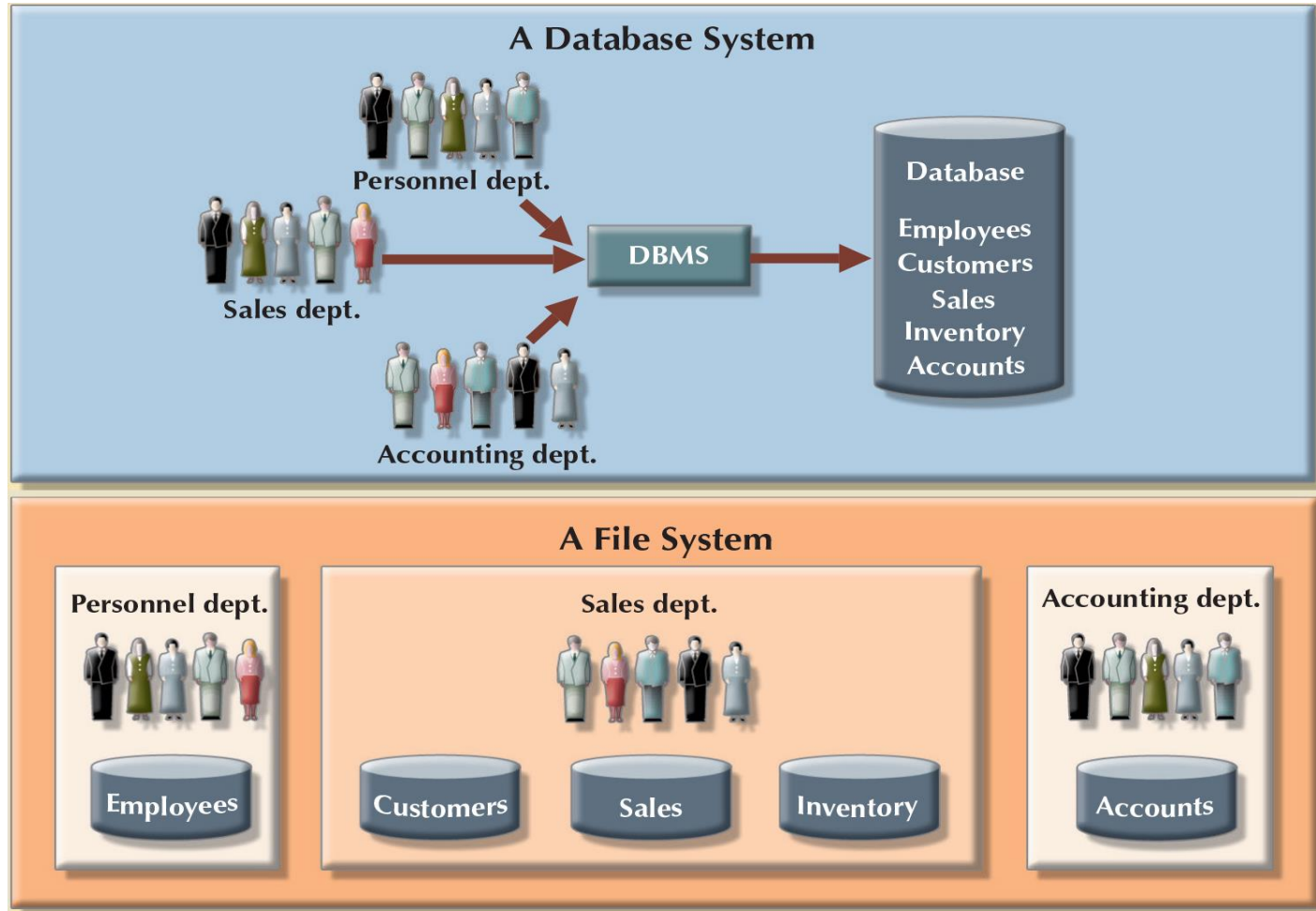
Deletion Anomalies

Database Systems

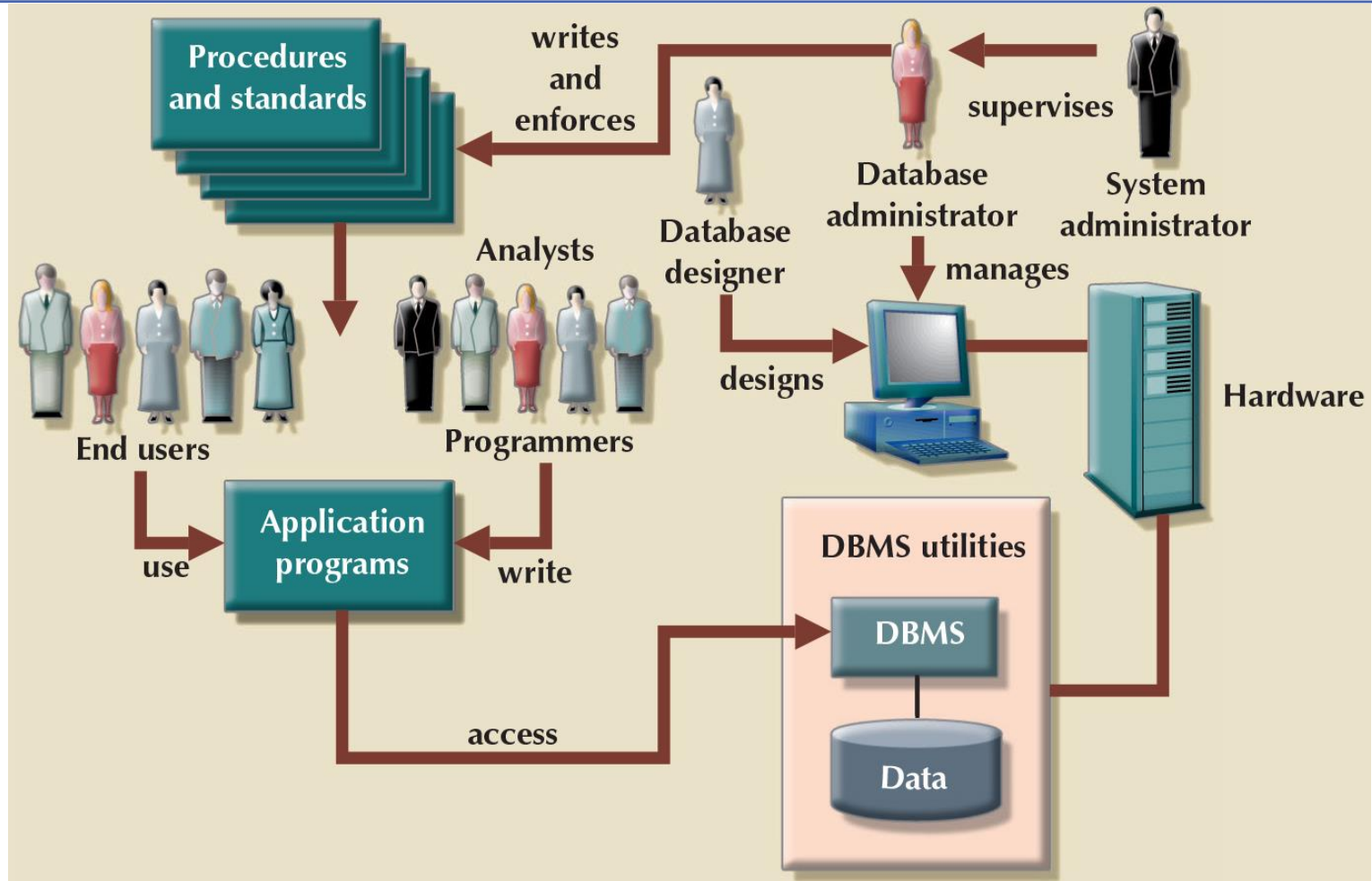
Database Systems

- Logically related data stored in a single logical data repository
 - Physically distributed among multiple storage facilities
 - DBMS eliminates most of file system's problems
 - Current generation DBMS software:
 - Stores data structures, relationships between structures, and access paths
 - Defines, stores, and manages all access paths and components
-

Contrasting Database and File Systems



The Database System Environment



DBMS Functions

Data dictionary management

- **Data dictionary:** Stores definitions of the data elements and their relationships

Data storage management

- **Performance tuning:** Ensures efficient performance of the database in terms of storage and access speed

Data transformation and presentation

- Transforms entered data to conform to required data structures

Security management

- Enforces user security and data privacy
-

DBMS Functions

Multiuser access control

- Sophisticated algorithms ensure that multiple users can access the database concurrently without compromising its integrity

Backup and recovery management

- Enables recovery of the database after a failure

Data integrity management

- Minimizes redundancy and maximizes consistency
-

DBMS Functions

Database access languages and application programming interfaces

- **Query language:** Lets the user specify what must be done without having to specify how
- **Structured Query Language (SQL):** De facto query language and data access standard supported by the majority of DBMS vendors

Database communication interfaces

- Accept end-user requests via multiple, different network environments
-

Disadvantages of Database Systems

Increased costs

Management complexity

Maintaining currency

Vendor dependence

Frequent upgrade/replacement cycles

Database Career Opportunities

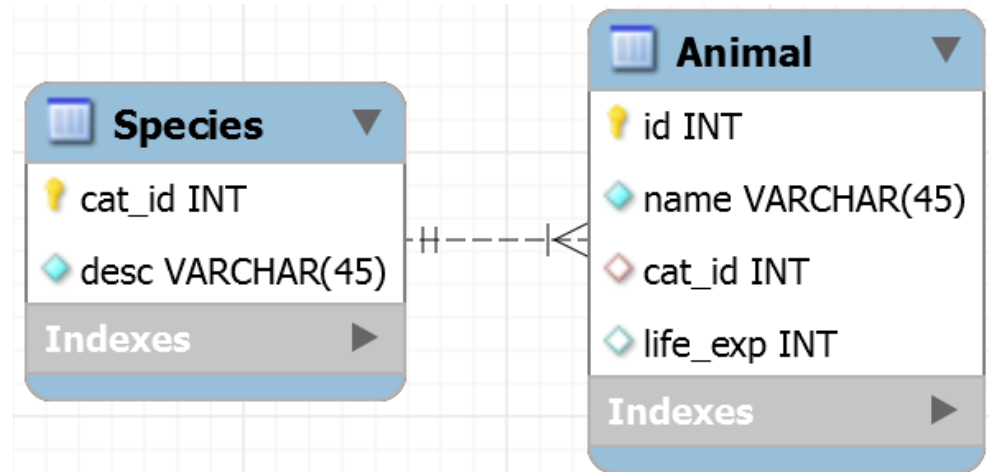
JOB TITLE	DESCRIPTION	SAMPLE SKILLS REQUIRED
Database Developer	Create and maintain database-based applications	Programming, database fundamentals, SQL
Database Designer	Design and maintain databases	Systems design, database design, SQL
Database Administrator	Manage and maintain DBMS and databases	Database fundamentals, SQL, vendor courses
Database Analyst	Develop databases for decision support reporting	SQL, query optimization, data warehouses
Database Architect	Design and implementation of database environments (conceptual, logical, and physical)	DBMS fundamentals, data modeling, SQL, hardware knowledge, etc.
Database Consultant	Help companies leverage database technologies to improve business processes and achieve specific goals	Database fundamentals, data modeling, database design, SQL, DBMS, hardware, vendor-specific technologies, etc.
Database Security Officer	Implement security policies for data administration	DBMS fundamentals, database administration, SQL, data security technologies, etc.
Cloud Computing Data Architect	Design and implement the infrastructure for next-generation cloud database systems	Internet technologies, cloud storage technologies, data security, performance tuning, large databases, etc.

What Databases Look Like

CAT_ID	DESC
1	INSECT
2	BIRD
3	FISH
4	MAMMAL

ID	NAME	CAT_ID	LIFE_EXP
1	Cat	4	20
2	Elephant	4	70
3	Trout	3	5
4	Shark	3	25
5	Canary	2	20
6	Albatross	2	40
7	Swift	2	5

Table Data



Entity-Relationship Diagram

Group Exercise

- In groups of 4 or 5, discuss and sketch a database that stores salient information about countries of the world
-

Group Exercise

- (Optional!)
 - Take a picture of your group's sketch and share it on the course's Facebook page