

## Introduction

07 August 2023 15:01

Data	Information
<ul style="list-style-type: none"> <li>Raw material that serves as foundation for knowledge</li> <li>Raw and unorganized, collection of facts</li> <li>Needs to be processed to be meaningful</li> </ul>	<ul style="list-style-type: none"> <li>When data is processed, organized and structured to be made meaningful</li> <li>Puts facts into context</li> <li>Utilized to make decisions and draw conclusions</li> </ul>

### Database

- Collection of related data representing an aspect of the real world
- Also known as mini world or UoD (universe of discourse)
- Logically coherent collection
- Designed, built and populated for a purpose (caters to application that interacts with it)
- Can vary by size and complexity

### Database Management System

- Contains information about an enterprise
- Convenient and efficient environment : user friendly interface and query language for data interaction
- Collection of interrelated data
- Set of programs to access data (how will different users access data? Web portal/ app?)
- Managing large, complex data
- Accessed by multiple users and applications: concurrent access
- Highly valuable and relatively large data

A general purpose software system that facilitates the process of defining, constructing, manipulating and sharing databases. DCMS

- Defining
  - Specifying the data types, structures, and constraints of the data to be stored in the database
  - Ensures the consistency and integrity of the stored data.
- Constructing
  - Organizes such that there is efficient retrieval & manipulation
  - Storing securely, only authorized parties can access
- Manipulating
  - Querying the database to retrieve specific data
  - Updating the database to reflect changes
  - Generating reports from the data
  - Any changes made follow predefined constraints and maintain data consistency
- Sharing
  - Allows concurrent access if user has necessary permissions

### Databases help us

- Store data(file structures, disk management)
- Understand data (data models)
- Keep data secure (security, recovery)
- Find data and use it (query languages, concurrency control and data analysis tools)
- Get accurate info (built in constraints and checks help with this)
- Maintain data integrity

### Applications:

Traditional Applications:

- Numeric and Textual Databases

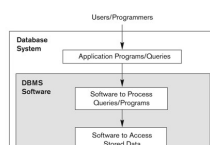
More Recent Applications:

- Multimedia Databases
- Geographic Information Systems (GIS)
- Data Warehouses
- Mobile databases
- Real-time and Active databases
- In-memory databases

### How Are Databases Used?

- Online Transaction Processing (OLTP):**
  - Used by a large number of users for small data retrieval and updates
  - common in most database applications like banking, universities, and airlines.
- Data Analytics:**
  - Involves processing data to draw conclusions and create predictive models for business decisions.
  - Examples include loan approval, targeted advertisements, and manufacturing decisions.
  - Data mining combines AI and statistical techniques for efficient analysis of large databases.

### Database system environment





a

### Example of a simple database

COURSE			
Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

SECTION				
Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	04	King
92	CS1310	Fall	05	Anderson
102	CS3320	Spring	05	Knuth
112	MATH2410	Fall	05	Chang
119	CS1310	Fall	05	Anderson
135	CS3380	Fall	05	Stone

GRADE REPORT		
Student_number	Section_identifier	Grade
17	112	B
17	119	C
9	85	A
9	92	A
9	102	B
9	135	A

PREREQUISITE	
Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

**Figure 1.2**  
A database that stores student and course information.

### Drawbacks of File Processing System

- Data Redundancy and Inconsistency**
  - Takes up more storage space
  - Can lead to inconsistencies if changes are made

**Solution:**

  - Store in centralized location
- Difficulty in accessing the data**
  - Not designed for a quick and efficient data retrieval

**Solution:**

  - Central information store
  - Powerful query language
  - Efficient ways to search and filter data
- Data isolation**
  - Data scattered across files in different formats

**Solution:**

  - Centralized system
  - Queries: standardized way to access data
- Integrity problems**
  - Careful consideration of constraints, cannot easily add

**Solution:**

  - Constraints specified directly into schema
  - Enforced automatically
- Atomicity problems**
  - Tasks may not get fully completed, resulting in integrity issues

**Solution:**

  - Atomicity must be ensured
  - Automatic rollback on failure
- Concurrent access anomalies**
  - Concurrent access not well supported, could result in integrity issues

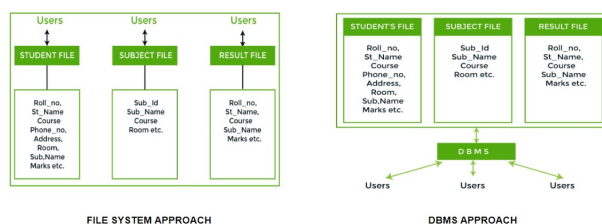
**Solution:**

  - Transaction and locking to be implemented
    - Transaction: Assures atomicity
    - Locking: prevents conflicting changes
- Security problems**
  - Access controls and permissions not well defined centrally

**Solution:**

  - Define access controls and permissions at database level
  - DBMS supports
    - User Authentication
    - Data Encryption
    - Auditing and Logging
    - Centralized security management

### DIFFERENCE BETWEEN FILE SYSTEM AND DBMS



### Characteristics of Database Approach

- Self-describing nature of a database system
  - The database system contains
    - Database
    - Metadata ( complete definition of db structure and constraints)
- Insulation between programs and data, and data abstraction
  - The structure of data files is stored in the DBMS catalog separately from the access programs (program-data independence)
- Support multiple views of the data
  - Support different types of users, each of whom may require a different perspective or view of the database.
  - A view may be a subset of the database or it may contain virtual data that is derived from the database files but is not explicitly stored.
- Sharing of data and multiuser transaction processing
  - Allows concurrent access
  - It ensures that data changes are effected in a controlled manner to maintain integrity

**Advantages of Data Base System**

- Controlling redundancy in data storage.
- Sharing of data among multiple users.
- Restricting unauthorized access to data
- Providing persistent storage for program Objects
  - E.g., Object-oriented DBMSs make program objects persistent
- Providing Storage Structures and search techniques for efficient Query Processing
- Providing backup and recovery
- Providing multiple interfaces to different classes of users.
- Representing complex relationships among data.
- Enforcing integrity constraints on the database.
- Drawing inferences and actions from the stored data using deductive and active rules and triggers
- Potential for enforcing standards
- Reduced application development time
- Flexibility to change data structures
- Availability of current information
- Economies of scale