Chapter 2: Database Management System

2.1 Introduction

In the field of Database Management Systems (DBMS), the concept of **data** is the starting point of all understanding. Every information system—whether it is MIS, DSS, or EIS—depends on data as its basic input. Without data, no processing can take place and no meaningful information can be generated.

In real life, organizations continuously generate large volumes of data through their daily activities. For example, when a customer makes a purchase, when a student registers for a course, or when an employee receives a salary, data is created at each step. However, this data in its original form is not directly useful for decision-making. It must first be organized, processed, and interpreted.

From a theoretical perspective, data is considered the raw material of information systems. Just like raw materials are processed to create finished products, data is processed to create useful information. Therefore, understanding data is essential to understand how databases and information systems function in an organization.

Meaning of Data

Data refers to **raw facts and figures** that are collected from various sources but do not have any meaning on their own until they are processed.

In simple words, data is: Unprocessed and unorganized facts

Data can exist in different forms such as numbers, text, symbols, or images. It becomes meaningful only when it is interpreted in a specific context.

Example

- "500, 700, 900" → This is **data** (just numbers)
- "Total sales are ₹2100" → This is **information**

Here, data is processed to produce meaningful information.

Definitions

1. Ralph M. Stair & George W. Reynolds

"Data consists of raw facts, such as employee numbers, hours worked, inventory levels, and sales figures."

2. Kenneth C. Laudon & Jane P. Laudon

"Data are streams of raw facts representing events occurring in organizations or the physical environment before they have been organized and arranged into a form that people can understand and use."

3. Gordon B. Davis

"Data is a collection of facts and figures which have not yet been organized or interpreted."

2.2 Key Characteristics of Data

1. Raw in Nature (Unprocessed Form)

Data exists in its original, unrefined state before undergoing any form of analysis or processing. It lacks interpretation and is often incomplete.

Example:

A list of numbers such as 1000, 2500, 1800 represents raw sales data without context (e.g., time period, product, region).

2. Unorganized (Lack of Structure Initially)

Data is typically collected from multiple sources and may not follow a predefined format or structure.

Example:

Customer feedback collected through emails, handwritten forms, and social media comments.

3. Context-Free (Lacks Meaning Without Processing)

Data does not inherently provide meaning unless it is interpreted within a specific context.

Example:

The number 5000 alone is meaningless unless specified as monthly sales, customer count, or profit margin.

4. Input for Processing (Foundation for Information)

Data acts as the primary input in the information processing cycle, where it is transformed into information through operations such as classification, calculation, and summarization.

Example:

Daily transaction records are processed to generate monthly financial reports.

5. Dynamic and Continuous in Nature

Data is continuously generated in modern digital environments, particularly with the growth of digital platforms and IoT systems.

Example:

Real-time data generated from online transactions or mobile applications.

6. Scalable and Voluminous

With advancements in technology, the volume of data generated has increased exponentially (often referred to as Big Data).

Example:

E-commerce platforms generating millions of transaction records daily.

7. Accuracy and Reliability Dependent

Data quality varies depending on how it is collected and maintained.

Example:

Incorrect survey responses or data entry errors.

2.3 Types of Data

1. Based on Nature

1. Quantitative Data (Numerical Data)

Quantitative data consists of measurable values expressed in numbers. It is used for statistical analysis and mathematical modeling.

Examples:

- Sales revenue (₹50,000)
- Number of students enrolled
- Market share percentage

2. Qualitative Data (Descriptive Data)

Qualitative data includes non-numerical information that describes attributes, opinions, or characteristics.

Examples:

- Customer satisfaction feedback
- Employee opinions
- Brand perception

2. Based on Source

1. Primary Data

Primary data is collected directly by the researcher or organization for a specific purpose.

Methods of Collection:

- Surveys
- Interviews
- Observations
- Experiments

Example:

A company conducting a survey to understand customer preferences.

2. Secondary Data

Secondary data is already collected and published by others for different purposes.

Sources:

- Government reports
- Research journals
- Company records
- Websites

Example:

Using census data for demographic analysis.

3. Based on Structure

1. Structured Data

Structured data is organized in a predefined format, typically in rows and columns.

Examples:

- Database tables
- Excel spreadsheets
- Transaction records

2. Unstructured Data

Unstructured data does not follow a specific format or structure.

Examples:

- Text documents
- Images
- Videos
- Social media posts

2.4 Difference between Data and Information

Basis	Data	Information
Meaning	Data refers to raw facts and figures that are collected from various sources. It does not carry any specific meaning on its own.	Information is processed and organized data that has meaning and can be understood clearly.
Nature	Data is unorganized and unstructured in its original form.	Information is organized, structured, and presented in a meaningful way.
Role in System	Data acts as an **input** in an information system.	Information acts as an **output** after processing the data.
Context	Data lacks context and interpretation, so it cannot be directly used.	Information has context and relevance, making it useful for understanding situations.
Usefulness	Data is not directly useful for decision-making.	Information is highly useful for decision-making and problem-solving.
Processing Requirement	Data requires processing such as sorting, classification, and calculation to become useful.	Information is already processed and ready to be used.
Dependency	Data is independent and exists in raw form.	Information depends on data for its creation.
Level of Detail	Data is detailed and may include large volumes of raw values.	Information is summarized and concise for easy understanding.
Example	Marks: 45, 60, 75	Average marks = 60 or Result: Pass/Fail
User Understanding	Difficult for users to interpret directly.	Easy for users to understand and apply in decisions.

Data and information are closely related but fundamentally different. Data is the raw input, while information is the meaningful output obtained after processing. Understanding this distinction is essential in DBMS, as it explains how systems convert raw data into valuable insights for decision-making.

2.5 Concept of Database and Database Management System (DBMS)

1. Concept of Database

1.1 Introduction

In the modern digital environment, organizations generate and handle vast amounts of data through their daily activities such as transactions, registrations, and communication. Managing this data in scattered files or manual records is inefficient and prone to errors. To overcome these limitations, data is organized systematically in the form of a database.

A database provides a structured way to store, manage, and retrieve data so that it can be used effectively for operations and decision-making. It ensures that data is not only stored safely but is also easily accessible when required. From banks and hospitals to colleges and e-commerce platforms, databases form the backbone of all information systems.

Thus, understanding the concept of a database is essential for understanding how modern organizations manage and utilize data efficiently.

1.2 Meaning of Database

A **database** is an organized collection of related data that is stored in a structured format and can be easily accessed, managed, and updated.

In simple words: A database is a **systematic and organized storage of data**.

A database ensures that data is:

- Stored in a logical manner
- Easily retrievable
- Consistent and accurate
- Shared among multiple users

Example

College Database

- Student details (Name, Roll No, Marks)

- Course information
- Faculty records

Bank Database

- Customer accounts
- Transaction history
- Loan details

1.3. Definitions

1. C. J. Date

"A database is a collection of persistent data that is used by the application systems of a given enterprise."

2. Elmasri and Navathe

"A database is a collection of related data."

3. Silberschatz, Korth & Sudarshan

"A database is a collection of interrelated data."

1.4 Characteristics of a Database

A well-designed database possesses the following characteristics:

- **Integrated Data**: Combines data from multiple sources into a unified system
- **Shared Access**: Allows multiple users to access data simultaneously
- **Data Persistence**: Data is stored permanently for future use
- **Reduced Redundancy**: Minimizes duplication of data
- **Data Independence**: Changes in structure do not affect applications

1.5 Components of a Database System

A database environment consists of:

1. **Data** – Raw facts stored in the system
2. **Hardware** – Physical devices such as servers and storage systems
3. **Software** – DBMS and related applications
4. **Users** – Administrators, developers, and end-users
5. **Procedures** – Rules governing database usage

Examples of Databases

- Banking systems storing customer account details
- University systems managing student records

- E-commerce platforms managing products and transactions

2. Concept of Database Management System (DBMS)

2.1 Introduction

In today's data-driven environment, organizations handle massive volumes of data generated from daily operations such as sales, banking transactions, student records, and online activities. Simply storing this data in a database is not enough; it must also be efficiently managed, controlled, and accessed. This is where a Database Management System (DBMS) becomes essential.

A DBMS acts as an interface between users and the database, allowing users to store, retrieve, update, and manage data in a systematic way. It eliminates problems associated with traditional file systems such as data redundancy, inconsistency, and lack of security. By providing structured control over data, DBMS ensures that information is accurate, secure, and readily available for decision-making.

Thus, DBMS is considered the backbone of modern information systems, enabling organizations to handle data efficiently and effectively.

2.2 Meaning of DBMS

A **Database Management System (DBMS)** is software that is used to create, store, manage, and manipulate data in a database.

In simple words:

DBMS is a **tool or software that manages databases**.

It allows users to:

- Create databases
- Insert, update, and delete data
- Retrieve required information
- Maintain data security and integrity

2.3 Definitions

1. C. J. Date

"A DBMS is basically a computerized record-keeping system."

2. Elmasri and Navathe

"A DBMS is a collection of programs that enables users to create and maintain a database."

3. Silberschatz, Korth & Sudarshan

"A DBMS is a general-purpose software system that facilitates the processes of defining, constructing, and manipulating databases."

2.4 Characteristics of Database Management System (DBMS)

List of Characteristics

1. **Data Independence**
2. **Reduced Data Redundancy**
3. **Data Consistency**
4. **Data Integrity**
5. **Data Security**
6. **Data Sharing (Multi-user Access)**
7. **Concurrency Control**
8. **Backup and Recovery**
9. **Data Abstraction**
10. **Efficient Query Processing**

1. Data Independence

Data independence refers to the ability to change the structure of the database without affecting the application programs. This means that modifications in data storage or schema do not require changes in user applications. It improves flexibility and reduces maintenance effort.

2. Reduced Data Redundancy

A DBMS minimizes duplication of data by storing it in a centralized and organized manner. Instead of keeping multiple copies of the same data in different files, DBMS ensures that data is stored once and shared wherever needed. This reduces storage cost and avoids unnecessary repetition.

3. Data Consistency

Since data redundancy is reduced, DBMS ensures that data remains consistent across the system. Any update made in one place is automatically reflected throughout the database, preventing conflicting or mismatched data.

4. Data Integrity

Data integrity ensures that the data stored in the database is accurate, valid, and reliable. DBMS enforces rules and constraints (such as primary keys, data types, and validation rules) to maintain correctness of data.

5. Data Security

DBMS provides mechanisms to protect data from unauthorized access. It allows administrators to define access rights and permissions, ensuring that only authorized users can view or modify specific data.

6. Data Sharing (Multi-user Access)

DBMS allows multiple users to access and use the database simultaneously. It supports a shared environment where different users can retrieve and update data without interfering with each other.

7. Concurrency Control

Concurrency control ensures that when multiple users access the database at the same time, transactions are executed in a controlled manner. It prevents conflicts and maintains data consistency during simultaneous operations.

8. Backup and Recovery

DBMS provides facilities to back up data and recover it in case of system failure, errors, or crashes. This ensures that data is not permanently lost and can be restored to its previous state.

9. Data Abstraction

Data abstraction hides the complexity of the database from users by providing different levels of view. Users interact with the system without needing to understand how data is physically stored. This makes the system easier to use.

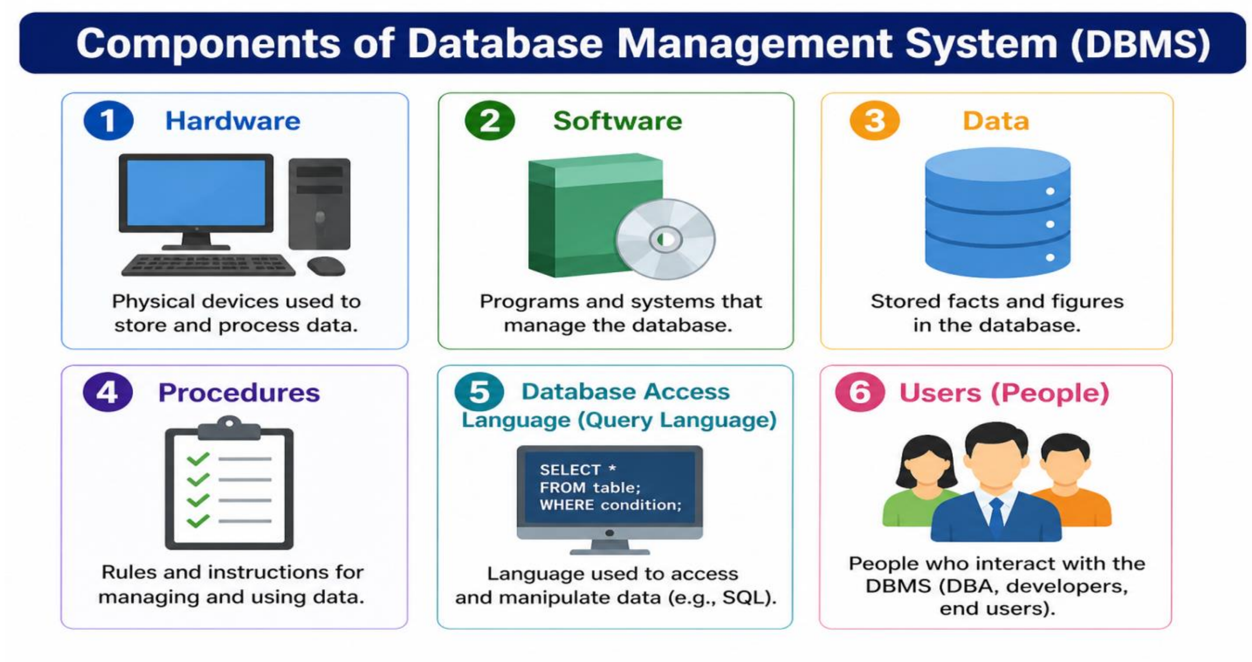
10. Efficient Query Processing

DBMS allows users to retrieve and manipulate data efficiently using query languages such as SQL. It optimizes queries to provide fast and accurate results, even when dealing with large volumes of data.

2.5 Components of Database Management System (DBMS)

1. **Hardware**
2. **Software**
3. **Data**
4. **Procedures**

5. **Database Access Language (Query Language)**
6. **Users (People)**



1. Hardware

Hardware refers to the **physical devices** on which the database system operates. It includes computers, servers, storage devices, and networking equipment. The performance and capacity of hardware directly affect the efficiency and speed of the DBMS.

2. Software

Software is the core component that manages the database. It includes:

- The DBMS software itself (such as MySQL, Oracle)
- Operating system
- Application programs

The DBMS software acts as an interface between users and the database, controlling access, processing queries, and managing data.

3. Data

Data is the most important component of DBMS. It consists of:

- Actual stored data (records, tables)

- Metadata (data about data, such as structure and definitions)

DBMS organizes and maintains data in a structured manner to ensure easy access and consistency.

4. Procedures

Procedures are the **rules and instructions** that define how the database system is used and managed. They include guidelines for:

- Data entry
- Data processing
- Backup and recovery
- Security measures

Procedures ensure proper and consistent use of the DBMS.

5. Database Access Language (Query Language)

This component refers to the language used to interact with the database. The most common is **SQL (Structured Query Language)**.

It allows users to:

- Retrieve data (SELECT)
- Insert data (INSERT)
- Update data (UPDATE)
- Delete data (DELETE)

This language makes communication between users and DBMS possible.

6. Users (People)

Users are individuals who interact with the DBMS. They include:

- **Database Administrator (DBA):** Manages and controls the database
- **Application Programmers:** Develop applications using DBMS
- **End Users:** Use the system for daily operations

Each user has a specific role in maintaining and using the database system.

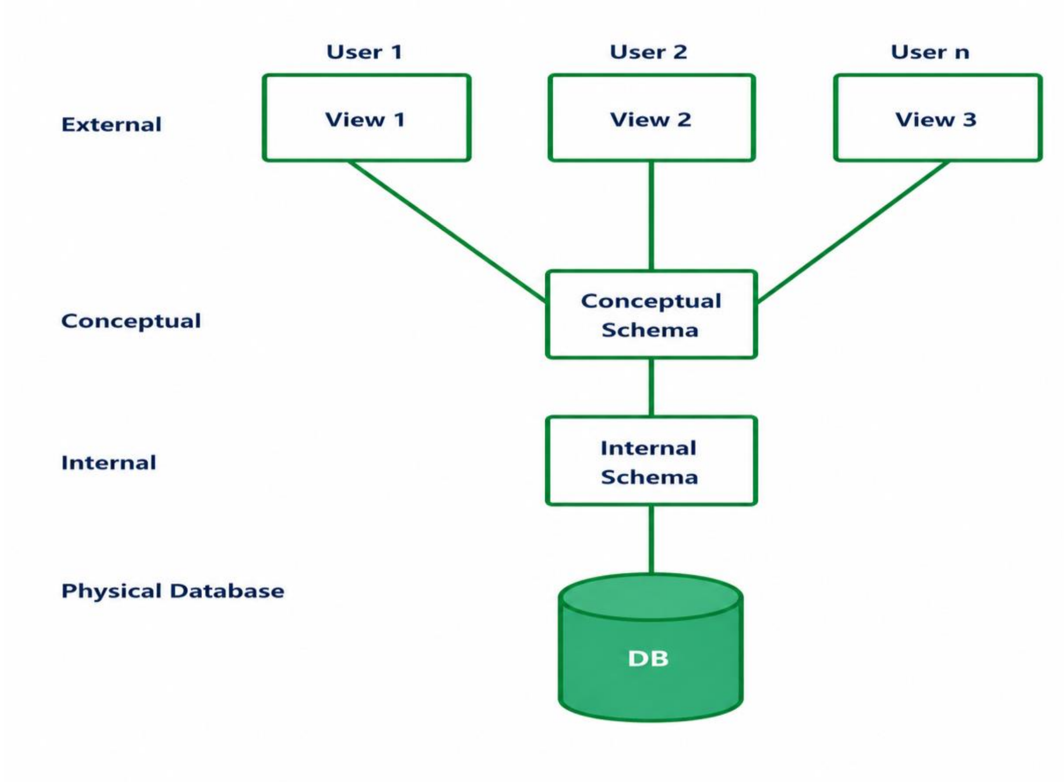
2.6 DBMS architecture

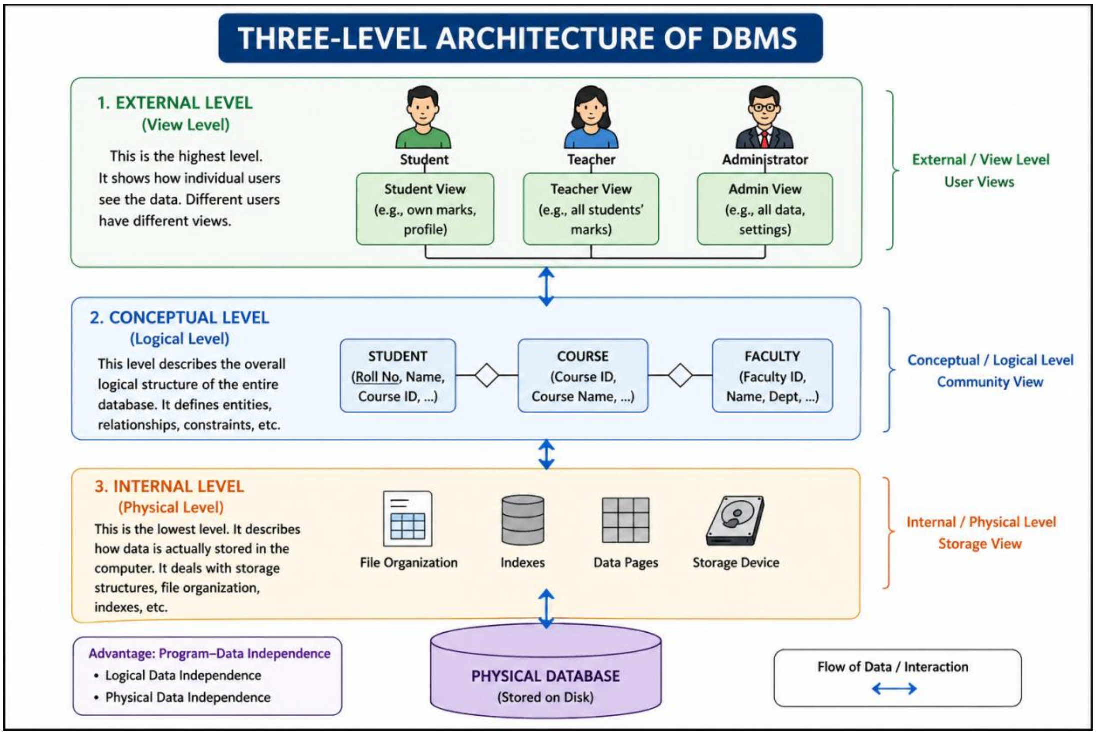
The **architecture of a Database Management System (DBMS)** refers to the overall design and structure that defines how data is stored, accessed, and managed within the system. It explains how different components of DBMS interact with each other and with users.

Understanding DBMS architecture is important because it helps in:

- Organizing data efficiently
- Ensuring data security and consistency
- Providing different views of data to different users

The concept of DBMS architecture is primarily explained through the **three-level architecture model**,





1. External Level (View Level)

The **external level** is the highest level of abstraction. It defines how individual users view the data.

At this level:

- Different users see different parts of the database
- Data is presented in a simplified and customized form
- Sensitive data can be hidden

Example:

A student can see only their marks, while a teacher can see marks of all students.

This level ensures **data security and user-specific views**.

2. Conceptual Level (Logical Level)

The **conceptual level** describes the overall structure of the entire database.

At this level:

- All data is represented logically
- Relationships between data are defined
- Constraints and rules are specified

Example:

A database defines tables like Student, Course, and Faculty and shows how they are related.

This level acts as a **bridge between the user view and physical storage**.

3. Internal Level (Physical Level)

The **internal level** is the lowest level of abstraction. It describes how data is actually stored in the system.

At this level:

- Data is stored in files and memory
- Storage structures and indexing are defined
- Performance optimization is handled

Example:

How student records are stored in hard disks or servers.

This level focuses on **efficient storage and fast data retrieval**.

Data Independence

One of the key advantages of this architecture is **data independence**, which means:

- Changes at one level do not affect other levels
- Users are not concerned with how data is stored physically

There are two types:

- **Logical Data Independence:** Changes in conceptual level do not affect external level
- **Physical Data Independence:** Changes in internal level do not affect conceptual level

2.7 Types of DBMS Architecture (Based on Deployment)

Apart from the three-level model, DBMS can also be categorized based on system structure:

1. 1-Tier Architecture

- Database and user interface are on the same system
- Used for development purposes

2. 2-Tier Architecture

- Client interacts directly with database server
- Common in small applications

3. 3-Tier Architecture

- Client → Application Server → Database
- Most widely used in web applications

2.8 Functions of Database Management System (DBMS)

1. **Data Definition (DDL Processing)**
2. **Data Storage, Retrieval and Update**
3. **User Access Control (Security Management)**
4. **Data Integrity Enforcement**
5. **Transaction Management**
6. **Concurrency Control**
7. **Backup and Recovery Management**
8. **Data Dictionary (Metadata) Management**

1. Data Definition (DDL Processing)

DBMS allows users to define the structure of the database using Data Definition Language (DDL). This includes creating tables, defining attributes, specifying data types, and setting constraints. It establishes the framework in which data is stored.

2. Data Storage, Retrieval and Update

One of the core functions of DBMS is to store data efficiently and allow users to retrieve and update it when required. It uses query languages like SQL to perform operations such as inserting, modifying, and deleting data.

3. User Access Control (Security Management)

DBMS provides security by controlling access to data. It ensures that only authorized users can access or modify certain data. This is achieved through authentication and authorization mechanisms.

4. Data Integrity Enforcement

DBMS ensures that data remains accurate and consistent by enforcing integrity constraints. These include rules such as primary keys, foreign keys, and validation checks.

5. Transaction Management

A transaction is a sequence of operations performed as a single unit. DBMS ensures that transactions are executed reliably using properties such as **Atomicity, Consistency, Isolation, and Durability (ACID)**. This guarantees correct and complete execution of transactions.

6. Concurrency Control

DBMS manages multiple users accessing the database at the same time. It ensures that concurrent transactions do not interfere with each other and maintains data consistency.

7. Backup and Recovery Management

DBMS provides mechanisms to back up data and recover it in case of system failure, errors, or crashes. This ensures data safety and continuity of operations.

8. Data Dictionary (Metadata) Management

DBMS maintains a data dictionary that stores metadata, which is information about the data. This includes definitions of tables, fields, data types, and relationships. It helps in managing and organizing the database effectively.

References

Davis, G. B. (1985). Management information systems: Conceptual foundations, structure, and development (2nd ed.). McGraw-Hill.

Date, C. J. (2004). An introduction to database systems (8th ed.). Pearson Education.

Elmasri, R., & Navathe, S. B. (2016). Fundamentals of database systems (7th ed.). Pearson.

Laudon, K. C., & Laudon, J. P. (2022). Management information systems: Managing the digital *firm* (17th ed.). Pearson.

Silberschatz, A., Korth, H. F., & Sudarshan, S. (2019). Database system concepts (7th ed.). McGraw-Hill Education.

Stair, R. M., & Reynolds, G. W. (2021). Principles of information systems *(14th ed.)*.

Connolly, T., & Begg, C. (2015). Database systems: A practical approach to design, implementation, and management (6th ed.). Pearson.

Coronel, C., & Morris, S. (2019). Database systems: Design, implementation, & management (13th ed.). Cengage Learning.

Rob, P., & Coronel, C. (2007). Database systems: Design, implementation, and management (7th ed.). Thomson Course Technology.