Database Systems

Data - Raw facts, or facts that have not yet been processed to reveal their meaning to the end user.

Information – The result of processing raw data to reveal its meaning. Information consists of transformed data and facilities decision making. (Coronel and Morris, 2017, p. 4)

Database is a shared, integrated computer structure that houses a collection of the following:

- End-user data – that is, raw facts of interest to the end user.
- Metadata, or data about data, through which the end-user data is integrated and managed. (Coronel and Morris, 2017, p. 6).

Roles and Advantages of DBMS (Coronel and Morris, 2017)

A **Database Management System** (**DBMS**) is a collection of programs that manages the database structure and controls access to the data stored in the database. The **DBMS** serves as the intermediary between the user and the database. The database structure itself is stored as a collection of files, and the only way to access the data in those files is through the DBMS.

Advantages of DBMS:

- **Improved data sharing** - The DBMS serves as the intermediary between the user and the database. The database structure itself is stored as a collection of files, and the only way to access the data in those files is through the DBMS.
- **Improved data security** - The more users access the data, the greater the risks of data security breaches. Corporations invest considerable amounts of time, effort, and money to ensure that corporate data is used properly. A DBMS provides a framework for better enforcement of data privacy and security policies.
- **Better data integration** - Wider access to well-managed data promotes an integrated view of the organization's operations and a clearer view of the big picture. It becomes much easier to see how actions in one segment of the company affect other segments.
- **Minimized data inconsistency** - Data inconsistency exists when different versions of the same data appear in different places.
- **Improved data access** - The DBMS makes it possible to produce quick answers to ad hoc queries. From a database perspective, a query is a specific request issued to the DBMS for data manipulation—for example, to read or update the data.
- **Improved decision making** - Better-managed data and improved data access make it possible to generate better-quality information, on which better decisions are based. The quality of the information generated depends on the quality of the underlying data.
- **Increased end-user productivity** - The availability of data, combined with the tools that transform data into usable information, empowers end users to make quick, informed decisions that can make the difference between success and failure in the global economy.

Types of Databases

Database Management System (DBMS) cab be used to build many different types of databases. The number of users determines whether the database is classified as single user or multiuser.

Types of Databases:

- **Single-user database** – A type of database that supports only one user at a time.
- **Desktop database** – A single user database that runs on a personal computer.
- **Multiuser database** – A type of database that supports multiple users at the same time.
- **Workgroup database** – A type of database that supports a relatively small number of users or a specific department within an organization.
- **Enterprise database** – A type of database that is used by the entire organization and supports many users across many departments.
- **Centralized database** – A type of database that supports data located at a single site.
- **Distributed database** – A type of database that supports data distributed across several different sites.
- **Cloud database** – A database that is created and maintained using cloud services, such as Microsoft Azure or Amazon AWS.
- **General-purpose database** – A database that contains a wide variety of data used in multiple disciplines.

- **Discipline-specific database** – A type of database that contains data focused on specific subject areas.
- **Operational database** – A type of database designed primarily to support a company's day-to-day operations.
- **Analytical database** – A type of database focused primarily on storing historical data and business metrics used for tactical or strategic decision making.

Importance of Database Design

Database Design refers to the activities that focus on the design of the database structure that will be used to store and mange end-user data. A database that meets all user requirements does not just happen; its structure must be designed carefully. In fact, database design is such as crucial aspect of working with databases that most of this book is dedicated to the development of good database design techniques. (Coronel and Morris, 2017, p. 11)

Oftentimes the database design does not get the attention it deserves. This can occur for numerous reasons such as:

- Insufficient specifications and/or poor logical data modeling
- Not enough time in the development schedule
- Too many changes occurring throughout the development cycle
- Database design assigned to, or performed by novices

The first step in constructing a physical database should be transforming the logical design using best practices. The transformation consists of the following:

- Transforming entities into tables
- Transforming attributes into columns
- Transforming domains into data types and constraints
- Transforming relationships into primary and foreign keys

File System Data Processing Issue (Coronel and Morris, 2017)

The file system method of organizing and managing data was a definite improvement over the manual system, and the file system served as useful purpose in data management for over two (2) decades. Nonetheless, many problems and limitations became evident in this approach.

A critique of the file system method serves two (2) major purposes:

- Understanding the shortcomings of the file system enable you to understand the development of modern databases.
- Many of the problems are not unique to file systems. Failure to understand such problems is likely to lead their duplication in a database environment, even though database technology makes it easy to avoid them.

The following problems associated with file systems, whether created by Data Processing (DP) specialist or through a series of spreadsheets, severely challenge the types of information that can be created from the data as well as the accuracy of the information:

- **Lengthy development times** – The first and most glaring problem with the file system approach is that even the simplest data-retrieval task requires extensive programming. With the older file systems, programmers had to specify what must be done and how to do it. As you will learn in upcoming chapters, modern databases use a nonprocedural data manipulation language that allows the user to specify what must be done without specifying how.
- **Difficulty of getting quick answers** – The need to write programs to produce even the simplest reports makes ad hoc queries impossible.
- **Complex system administration** – System administration becomes more difficult as the number of files in the system expands. Even a simple file system with a few files requires creating and maintaining several file management programs.
- **Lack of security and limited data sharing** – Another fault of a file system data repository is a lack of security and limited data sharing. Data sharing and security are closely related. Sharing data among multiple geographically dispersed users introduces a log of security risks.
- **Extensive programming** – Making changes to an existing file structure can be difficult in a file system environment.

Structural dependence – A data characteristic in which a change in the database schema affects data access, thus requiring changes in all access programs.

Structural independence – A data characteristic in which changes in the database schema do not affect data access.

Data dependence – A data condition in which data representation and manipulation are dependent on the physical data storage characteristics.

Data independence – A condition in which data access is unaffected by changes in the physical data storage characteristics.

Data redundancy – It exists when the same data is stored unnecessarily at different places.

Uncontrolled data redundancy sets the stage for the following:

- **Poor data security** – Having multiple copies of data increases the chances for a copy of the data to be susceptible to unauthorized access.
- **Data inconsistency** – Data inconsistency exists when different and conflicting versions for the same data appear in different places.
- **Data-entry errors** – Data-entry errors are more likely to occur when complex entries are made in several different files or recur frequently in one or more files.
- **Data integrity problems** – It is possible to enter a nonexistent sales agent's name and phone number into the Customer file, but customers are not likely to be impressed if the insurance agency supplies the name and phone number of an agent who does not exist.

Data Anomalies

- A data abnormality in which inconsistent changes have been made to a database.
- A data anomaly develops when not all of the required changes in the redundant data are made successfully.

REFERENCES:

Coronel, C. and Morris, S. (2017). *Database systems: design, implementation, and management, 12th edition.* USA: Cengage Learning.

Elmasri, R. and Navathe, S. (2016). *Fundamentals of database systems, 7th edition.* USA: Pearson Higher Education.

Kroenke, D. and Auer, D. (2016). *Database processing: fundamentals, design, and implementation.* England: Pearson Education Limited.