COMPUTER ORGANIZATION AND ARCHITECTURE

Lecture 10

Databases

Dr Victoria Mukami

INTRODUCTION

This lecture is an introduction to the world of databases. Specifically, we will review some database definitions, the hierarchy of data and various characteristics of data. We will then look at the types or classifications of databases and we will focus on a term known as DBMS. We will finally focus on the people who work with databases

Learning objectives

By the end of this topic, you should be able to:

1. Give various definitions of terms associated with databases
2. Understand the characteristics of various databases
3. Identify the types of databases
4. Discuss the role of various database users

OVERVIEW

We have so far learnt about the differences between hardware and software. We have also reviewed the different categories of software like operating system and application software. We are now going to focus on databases which is a type of application software. As computer users, programmers, or developers, we generate loads of data. Whether it's a bank or a hospital, this data needs to be stored somewhere. This is where databases come in.

A database is a collection of different types of data that is organized in a manner that allows for easy access. Remember we defined data as unprocessed information. There when a database holds a lot of data that is organized in essence it becomes information storage. Recall that we also defined information as meaningful or processed data.

DATABASES

Databases have come a long way to where they are today. Before there were databases there existed the file system. The file system used tabled data (database tables) that captured records and were stored individually. The file system was based on the physical cabinet system. Where each user had a file opened and all their records went to the file. However, you would find that the payment details might be stored in another file in another cabinet. As a result, the file system sometimes had repetitive information, incorrect information (imagine a user changes their phone

number, it may be changed on only one file record) and took time requesting for the files (tables) from other departments. This then led to the evolution of the database management system (DBMS).

Database Management System

This approach utilizes a central repository of data that can be accessed by a variety of authorized users. The DBMS approach has many advantages over the previous file system. The following are some of the advantages of the DBMS:

1. Low redundancy within the records stored [4].
2. Improved data integrity which points to more accurate data as users work on a single file [1]
3. Users can share data and get the most up to date each time they access the database
4. The DBMS approach allows for non-technical users to easily access and modify data.

Of course, as with everything the DBMS system has its share of disadvantages. The greatest limitation of the DBMS approach is how vulnerable data can be. Records could easily be modified or deleted by users whether intentionally or non-intentionally. To mitigate this limitation, user roles must be separated, and the security of the data is taken seriously.

DATA CHARACTERISTICS

Data is considered an asset for many organizations. Data is used to determine whether an organization can thrive and keep up with market trends. Therefore, several characteristics are used to define data.

1. Qualities of valuable information

Information is processed data. For information to be considered useful, several items need to be met. First, information needs to be **accurate**. This means that the information does not contain any errors [1]. Information needs to be **verifiable**. This means that you can check whether the information is correct or incorrect. Information also needs to be **accessible** by the relevant party. Information that cannot the accessed by the person who needs it is considered useless. Information also needs

to be **timely** and **organized**. Timely in the sense that receiving late information means that it does not help anyone. Organized as it means someone can make decisions with the information. Finally, information needs to have **meaning** and should be **cost-effective** [1]. The cost of gathering the information should be lower than what the person can do with the information.

2. Data Hierarchy

Data has various levels hence the hierarchy. The first level is considered as the **characters**. Remember when we defined the smallest unit as a bit and a group of 8 bits making a byte? Now a character is defined as a group of bytes and could represent a letter, number, or symbol. Next, we have a **field**. A field is made up of a group of characters and each field is uniquely identified by a **field name** [1]. An example of a field would be Jane with the field name being Student Name. Fields normally accept various types of data as shown in the figure such as number, short text, long text.

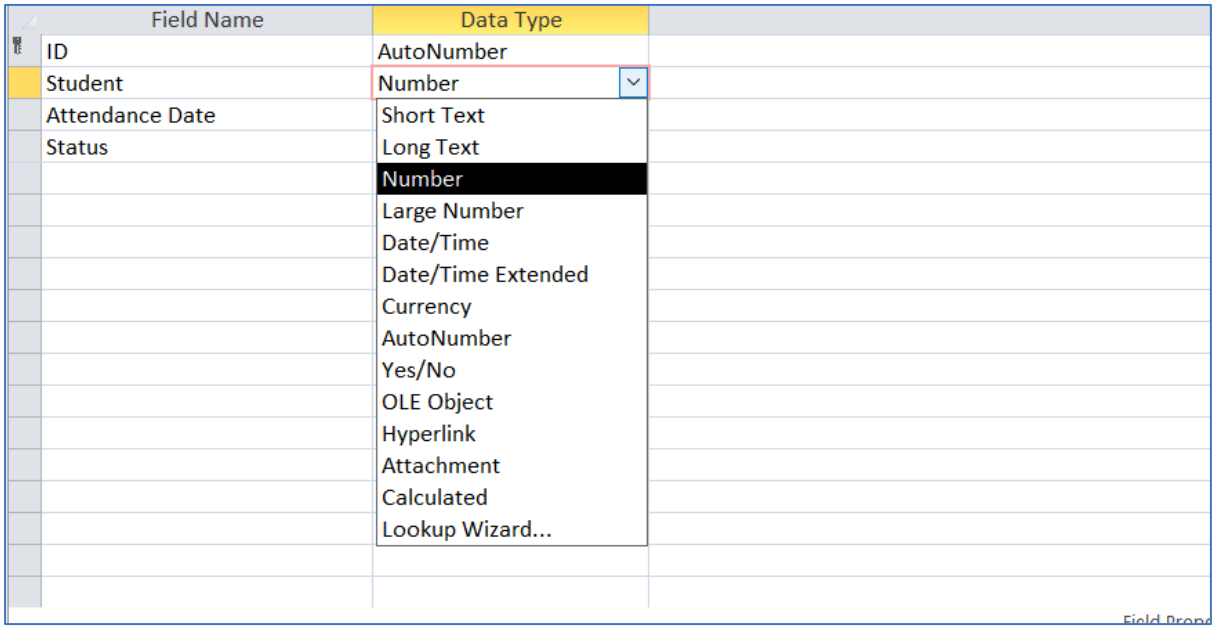

Figure 1: Various data types

A **record** is a group of related fields [1]. For instance, a student-related database would have one student's record containing the various field names such as the student's name, date of birth, gender, country etc. Each record has what is called a **primary key** which is used to uniquely used to identify a record. A primary key cannot be empty neither can it be duplicated therefore a field like date of birth, name or country would

all be bad primary keys. On the other hand, a field like Student ID, national ID, bank account number would be good examples of primary keys.

A **file** is considered a collection of all related records stored within one table. From the figure, these are student attendance records of various students.

ID	Student	Attendance Date	Status
1	Jane John	20/10/2021	Present
2	Jane John	21/10/2021	Absent - Excused
3	Jane John	23/10/2021	Absent - Unexcused
4	Jane John	25/10/2021	Present
5	John Jones	22/10/2021	Present
6	John Jones	25/10/2021	Present
7	John Jones	26/10/2021	Present

Figure 2: Data file example

3. Relationships

A relationship is defined as an association between one or more entities [4]. An **entity** has similarities to a noun and is considered as anything such as a person, object or event that is important to any organization. The relationship will show data that is related to other data. There are three types of relationships:

One-to-One: This is where one entity is related to exactly one other entity. It is shown by 1:1 and a good example is a school where a specific teacher is assigned a specific class. Therefore, one teacher is assigned to one class only.

One-to-Many: The is where one entity is related to many other entities and is represented by a 1:M. These are the most common type of relationships. For example, a headteacher oversees many teachers and many teachers report to one headteacher.

Many-to-Many: this is where many entities are related to many other entities and is represented by M: M. For instance, a student can take many subjects and one subject can be taken by many students.

4. Data Dictionary

A data dictionary contains the various data definitions of a database. Data definitions describe the properties of data for each data file. In addition, the data dictionary contains security information and relationships within the database.

5. Data Security

As with all things computers, security is key when managing databases. DBMS allow for various access levels to ensure accountability. These access levels are used to limit what each user can view and edit. In addition, some levels of access allow users to only view and not change data.

DATABASE MODELS

A data model consists of rules and standards that define how the database organizes data [1]. The data model oversees how organizational data is viewed by the various users. Specific models exist and are discussed below

Relational Database

This is a database that stores data in tables. Each table is made up of related records represented by field names. Tables are considered independent, but data can be retrieved through the relationships within the tables. A good example of a relational database model is Microsoft Access. Microsoft access is characterized by tables, forms, queries and reports. Tables are related to other tables using relationships as shown.

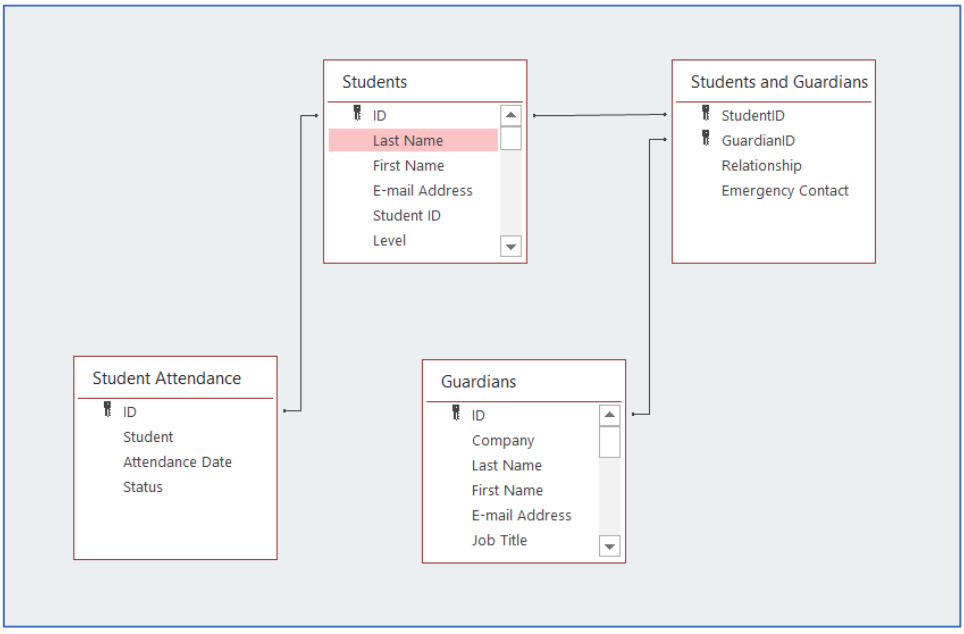

Figure 3: A relationships view

A form is used to insert data into a table while a query is used to ask questions of the data.

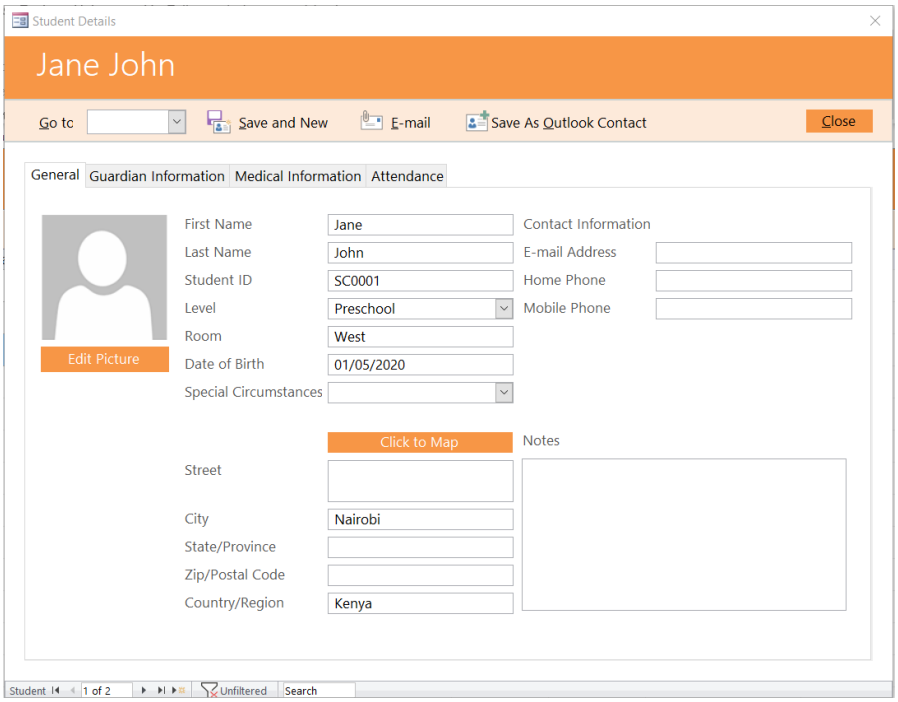

Figure 4: An example of a form

For instance, a user might want to ask questions of data like how many students attended a class on a certain date. Once this data is queried then it is best presented using reports.

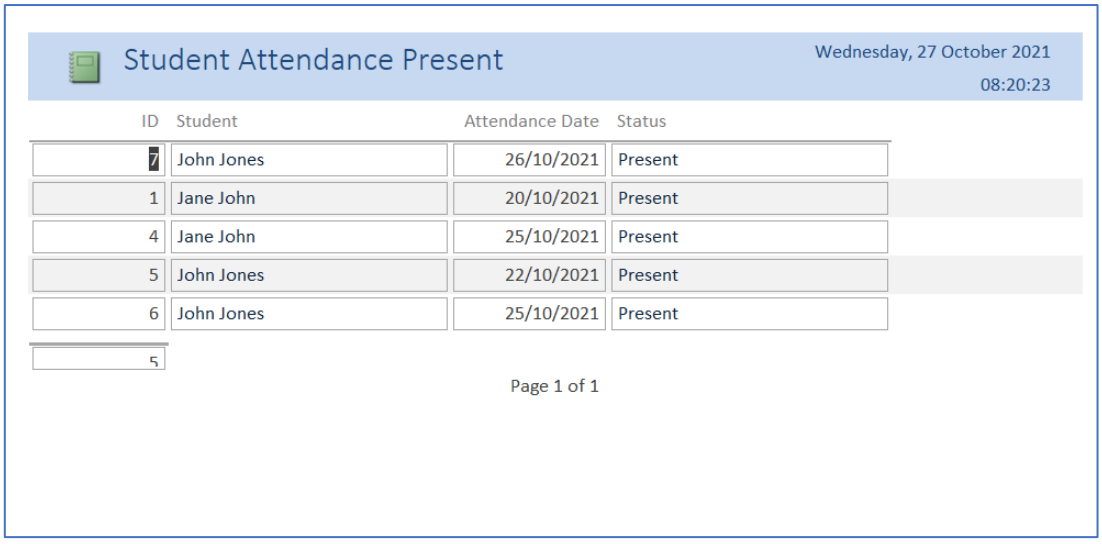

Figure 5: A report view

The relational database model best works when it is designed correctly to ensure consistent and quality data.

Object-Oriented Database

The object-oriented database stores data in objects whereby objects are items that contain data. These objects will contain attributes of a person or place or event that is of importance to the organization. One of the greatest advantages of an object-oriented database is the ability to retrieve complex unstructured data than the relational database. This data includes images, videos, audio and even documents.

Multi-dimensional Databases

A multi-dimensional database stores data in various dimensions. Dimensions are where you think of 2-D, 3-D. A multi-dimensional database stores more than two dimensions of data. This type of database allows users to view data from different dimensions.

DATABASE USERS AND ROLES

Various users will access a database and use the database for a variety of their needs. These users are made up of end-users as well as the developers and designers of the database.

Database designers

As with any developed system, design is a critical area of any system. The database designer is normally responsible for creating the design of a database. The designers work with other developers to craft the data views including tables, the relationships expected, and other areas related to the design of the database.

Database developers

These are the people who create the database based on the design given by the database designers. The development includes creating the actual structure including the relationships. Sometimes data needs to be coded and this is done by the developers.

Database administrators

These are individuals who oversee managing the database. If any errors should arise or any updates need to be done, the database administrator would oversee those tasks. The administrators oversee security and user access and will also create backup and perform restores when needed.

Users

This is the end-user who works with the database. This is the user who will add, delete, and update data directly to the database. These users do not understand how the database works on the backend with their only intention being the frontend.

SUMMARY

Throughout this lecture, we have reviewed the terms database and have gone further to look at the various characteristics of databases. This then led to the various types of database models with an exclusive view of the relational database management system model. Finally, a review of the various database users was conducted.

DISCUSSION TOPIC

While databases are mainly associated with technological fields, they are used throughout various sectors and industries. One of the sectors that use databases is the health sector. Patient data, results, diagnosis, and treatment plans are all stored within databases. The patient data needs to be secure so that only the authorized doctor can view the data. The Lab technician can only see the tests requested but not the patient history or previous diagnosis. What kind of database is ideal for such scenarios and what sort of security would be implemented to ensure ultimate patient privacy?

REFERENCES

[1] G. Shelly and M. Vermaat, Discovering Computers — Fundamentals: Your Interactive Guide to the Digital World. Boston, MA: Course Technology, 2012.

[2] W. Stallings, Computer Organization and Architecture Designing for Performance. Hoboken, NJ: Pearson Education, Inc, 2016.

[3] A. Evans, K. Martin and A. Poatsy, Technology in Action. New York, NY: Pearson, 2020

[4] D. Morley and C. Parker, Understanding Computers: Today and Tomorrow. Boston, MA: Course Technology, 2017.