

## ***CPS106 (DATA PROCESSING) COURSE GUIDE AND STUDY UNITS***

### ***INTRODUCTION***

CPS106 Data Processing is a two [2] credit unit course of Fifteen units. It deals the fundamentals concepts of data processing, basic components of computer systems, data processing personnel, file access and organization techniques. In addition it deals with the tools for developing data processing application and provides an overview of Information Systems and Structured Query Language. The lecture material provides adequate background information that is relevant for students pursuit of a career in data processing or related disciplines. The course is divided into four modules:

Module 1 introduces the basic concepts of data processing, explains the meaning of data, information and data processing. A detailed description of the computer hardware and software components required for data processing is covered. The various input and output devices such as pointing devices, pen-based devices, optical character recognition, printers, plotters etc and their suitability for data processing are explained.

Module 2 provides information on data storages devices and the parameters for disk performance. File access and organization techniques - serial, sequential, indexed sequential and random file organization are covered in this module. In addition, you will learn about data processing techniques such as Batch processing, Online Processing, Transaction Processing, Real-time Systems and Distributed processing.

Module 3 begins with database processing. The meaning of database, components of database and database models i.e. relational database, hierarchical database, network database and object-oriented database are explained. Data processing personnel are covered. In addition, tools for developing data processing applications such as programming languages, report generators, computer aided software engineering, query languages etc are covered in this module.

Module 3 provides an overview of Information Systems and Structure Query Language (SQL). The fundamental concepts of information systems are covered. Also, you will learn how to write simple p to write simple SQL queries by using the CREATE, SELECT, INSERT, UPDATE and DELETE commands

The aim of this course is to equip you with basic information and understanding of data processing concepts as well as acquiring the relevant skills to manage and implement data processing systems. At the end of the course, you would have gained the relevant knowledge and acquired the skills to effectively manage the data processing department of any organization.

### ***What you will learn in this course***

The main purpose of this course is to provide the required knowledge and business programming skills for you to effectively data processing systems and manage the data processing department of an organization. The basic concepts of data processing, computer systems required for data processing, data processing personnel, tools for developing data processing application, and basic programming skills are covered in this study material. This, we intend to achieve through the following:

## ***Course Aims***

- i. Explain the concepts, systems, personnel, tools for data processing
- ii. Describe data processing techniques
- iii Acquire skills for developing, implementing and managing data processing systems or department
- iv. Introduce the concepts associated with Information systems development;

## ***Course Objectives***

Certain objectives have been set out to ensure that the course achieves its aims. Apart from the course objectives, every unit of this course has set objectives. In the course of the study, you will need to confirm, at the end of each unit, if you have met the objectives set at the beginning of each unit. By the end of this course you should be able to:

- i. Explain the meaning of data, information, and data processing;
- ii. Identify and describe the basic hardware components of a computer system;
- iii Describe some software use for data processing;
- iv Explain file access and organization techniques;
- v Describe batch processing, real-time processing, transaction processing etc;
- vi Explain the meaning of Database Management Systems and Database Model;
- vii Describe various tools for developing data processing applications;
- viii Write SQL statements

## ***Working through This Course***

In order to have a thorough understanding of the course units, you will need to read and understand the contents, practice the what you have learn, and be committed to learning and implementing your knowledge.

This course is designed to cover approximately fifteen weeks, and it will require your devoted attention.

## ***Course Materials***

These include the Course Guide , Study Units, and Recommended Texts.

## **STUDY UNITS**

There are fifteen study units in this course:

### ***Module1***

Unit 1 INTRODUCTION AND SCOPE OF DATA PROCESSING

Unit2 BASIC HARDWARE COMPONENTS AND MEMORY CAPACITY MEASUREMENT

Unit 3 BASIC SOFTWARE FOR DATA PROCESSING

Unit 4 DATA CAPTURING AND REPORTING DEVICES

### ***Module 2***

Unit 1 DATA STORAGE DEVICES

Unit 2 FILE ORGANIZATION

Unit 3 DATA PROCESSING TECHNIQUES

## Unit 4 TRADITIONAL FILE SYSTEM PROCESSING

### **Module 3**

Unit 1 DATA BASE PROCESSING

Unit 2 DATA PROCESSING PERSONNEL

Unit 3 TOOLS FOR CREATING DATA PROCESSING APPLICATION

### **Module 4**

Unit 1: INTRODUCTION TO INFORMATION SYSTEMS

Unit 2: INTRODUCTION TO BASIC SQL

Unit 3: SQL SYNTAX

Unit 4: SQL STATEMENTS

### **Textbooks and References**

Alexis Leon and Matthew Leon, Fundamental of Information Technology, L & L Consultancy Services Pvt. Ltd, 1999

Brain K. Williams, and Stacey C. Sawyer, Using Information Technology, A practical Introduction to Computers and Communications, McGraw Hill, 2005, N.Y

Carl Hamacher, Zvonko Vranesic, Safwat Zaky, Computer Organization, Fifth Edition, McGraw-Hill Higher Education, N.Y, 2002

David C. Kreines, Oracle SQL, O Reilly, USA, 2000

Gerald V. Post, Database Management Systems, Designing and Building Business Applications, Third Edition, McGraw-Hill, 2005

George Beekman, Eugene J. Rathswohl, Computer Confluence, Exploring Tomorrow s Technology, IT Edition, Prentice Hall, New Jersey, 2003

Jan L. Harrington, SQL Clearly Explained, Morgan Kaumann Publisher, An Imprint of Elsevier Science, 2003, USA

<b>Unit</b>	<b>Title of Work</b>	<b>Weeks</b>	<b>Assessment</b>
	<b>Activity (End of Unit)</b>		
	Course Guide	Week 1	
	Module 1		
1	Introduction and Scope of Data Processing	Week 1	Assignment 1
2	Basic Hardware Components and Memory	Week 2	Assignment 2
3	Capacity Measurement	Week 2	
4	Basic Software for Data Processing	Week 3	Assignment 3
	Module 2		
1	Data Storage Devices	Week 4	Assignment 4
2	File Organization	Week 5	Assignment 5
3	Data Processing Techniques	Week 5	
4	Traditional File System Processing	Week 6	Assignment 6

### Module 3

1 Data Base Processing

Week 7

Assignment 7

2 Data Processing Personnel

Week 8

Assignment 8

3 Tools for Creating Data

Week 9

Assignment 9

Processing Application

### Module 4

1 Introduction to Information Systems

Week 10

Assignment 10

2 Introduction to Basic SQL

Week 11

Assignment 11

3 SQL Syntax

Week 11

Assignment 11

4: SQL Statements

Week 12

Assignment 12

Revision

Week 13

TOTAL 13 weeks

## **MODULE 1 UNIT 1: INTRODUCTION AND SCOPE OF DATA PROCESSING**

### **1.0 Introduction**

The word data refers to raw facts about an object. Data is meaningless until it is subjected to manual or electronic processing by a computer to produce information used for decision making. Computers are used to process data because of their speed, reliability and other functionalities they offer. Computers can be distinguished by their sizes, types and generation.

### **2.0 Objectives**

At the end of this unit, you should be able to:

- Define and explain the terms; data and information
- Describe a computer system
- Explain the meaning of data processing
- Distinguish between the different types of computer
- Describe the basic operations of a computer
- List some benefits of using computer in business

### **3.1 Data and Information**

Data exist in a variety of forms such as numbers or text on pieces of paper, as bytes in computer memory or as facts stored in a person's mind. Data is the plural form of the word datum and refers to basic facts about an object. An object refers to a person, place, or a thing. Data consist of numbers, words, images, etc. Data is meaningless until it undergoes some forms of processing. The outcome of a processed data is information. Although the word data and information are often used interchangeably, there is an important distinction between them. In the strictest sense, data consist of the raw number that computers process to produce information. Data can be looked at as facts out of context.. Table below captures basic facts about a matriculated student of Fountain University, Osogbo. The data items consisting of Matriculation Number, Name, College, Department, Centre, Banker, Status and e-Email are used to hold the students details.

S/no	Data item	Actual Details or value
1	Matriculation Number	FUO/22/00054
2	Name	Ismail Toyibah Alake
3	College	CONAS
4	Department	Computer Science
5	Banker	Zenith Bank
6	Status	Returning
7	e-Email	toyibaalake@gmail.com

Similarly, a commercial bank may request for the following data before the same student can operate an account or enjoy some banking services

S/no	Data item	Actual Detail or value
1	Matriculation Number	FUO/22/00054
2	Name	Ismail Toyibah Alake
3	Date of Birth	14th April, 2003
4	Nationality	Nigerian
5	Sex	Female
6	Occupation	Student
7	School	Fountain University, Osogbo

The details of this student (Ismail Toyibah Alake ) if treated independently may not have specific meanings until it is subjected to some manual or computerized processing. The raw data in the tables may be further processed by a human or input into a computer, stored, processed and transmitted (output) to another human or a computer. A typical data processing function will be to locate a student record or data from a large file based on matriculation number. If this is to be carried out manually a lot of time and efforts will be required by the data processing personnel. When the method of data processing is manual, it is called manual data processing. On the other hand, when the method of processing is electronic we refer to it as electronic data processing. Computers are used to turn meaningless data by processing them into useful information, such as spreadsheet, graphs, and reports. In short, information is defined as either a meaningful answer to a query or a meaningful stimulus that can cascade into further queries. For example, with the data in the two tables, the bank official may be interested in the number of students from each school that have subscribed to a particular service its offers. In which case, a query will be run on the existing database. The result of the query will provide information which will help the management of the bank to discontinue or continue with the service. Some characteristics of information are: relevance, accuracy, conciseness, timeliness, and completeness.

### ***3.2 Electronic Data Processing Systems***

Electronic Data processing (DP) is any computer automated process that converts raw facts i.e data into output i.e useful information. Data becomes more useful only when they have undergone some computer processing and well-presented to allow for decision making. Data processing systems have become indispensable in managing information in modern day organizations. In recent times, data processing systems are quite often referred to as information systems to highlight their practicality. All the same, both terms are more or less the same, performing similar functions; data processing systems typically manipulate raw figures into information, and likewise information systems usually take raw figures as input to generate useful information as output. In many organizations, these systems exist as payroll systems, sales and marketing systems, finance and accounting systems, manufacturing and production system, human resources systems, University management systems etc.

### ***3.3 Scope of Data Processing System***

Conventional usage of the terms data processing and information systems restricts their use to refer to the algorithmic derivations, logical deductions, and statistical calculations that recur perennially in general business environments, rather than in the more expansive sense

of all conversions of real-world measurements into real-world information. In the real sense, although, data processing requires the use of computers, the level of computerizations is usually not too computing intensive as may be expected in scientific and engineering applications. Computing is a term restricted to number crunching that is arithmetical calculations. These include adding, multiplying, subtracting and dividing, and exponentiation (raising numbers to specific powers) etc.

The basic data processing operations performed on business data include the following:

#### Data Collection

- Data Capture
- Data Recording
- Data Entry
- Data Transmission
- Data Sorting
- Updating of Information
- Adding of Data
- Deletion of information
- Data Validation
- Data Verification
- Summarizing of data
- Printing of Results
- Summaries of data for management decision
- Calculating data
- Sorting data
- Classifying data
- Summarizing data
- Comparing data
- Statistical reports
- Data Mining

Hence, Data processing consist of those activities concerned with the systematic recording, sorting, computing, modifying, reporting displaying and printing of details relating to business transactions. While a data processing system can therefore be viewed as an administrative system superimposed upon the physical business systems such as banking, engineering, insurance etc. The ultimate goal of data processing is to obtain information with which to control financial and administrative aspects of the business. In addition, managers and administrators with up-to-date information are able to make good decision that will positively affect the growth of organization.

### ***Computer Systems***

A computer is a device which given a set of instructions or data can be used to perform given task or tasks. Also, a computer can be referred to as a programmable, multi-user device that accepts data, raw facts and figures, and processes, or manipulates, it into useful information Computer are used primarily to speed up the problems solving and increase the overall productivity of its users. The computer reads in data and instructions, does some processing, and stores or outputs desired results. The computer is designed with input mechanism for reading data into the computer, internal storage facilities, and mechanisms for communicating with the outside world (output for writing data out.) In order for data to be processed by a computer, the data needs first to be converted into a

machine readable format. Once data is in digital format, various procedures can be applied on the data to get useful information.

### ***The Basic Operations of Computing***

The computer processes data by performing the following five main operations

#### ***Input Operations:***

At this stage, data or instructions are captured electronically or entered by means of the available input device. An input device is hardware such as a keyboard that allows data from the external environment to be entered into the computer for processing. The data after it has been captured is transform into a form in which the computer can process it.

#### ***Processing Operations:***

At this stage, the already captured and transformed data is manipulated to generate desired result for the end user other processing systems. The data is worked on by the instructions in the form of programs or queries provided by the users. The instruction and the data determine what output is received from to the computer. The processing instructions may be to add, subtract, multiply, find total, summarize, group, select some data based on some condition etc.

#### ***Output Operations:***

At the stage the result or information obtained from the data is produced in a form acceptable by the user. An example of on output will be a list of all registered students in the department of accounting who are in their second year at National Open University, Lagos Centre. The output can be printed as printed text, played as sound, displayed as charts or graphs on the computer screen. Output is usually governed by the need to communicate specific information to a specific audience. The only limit to the different forms of output you can produce is the different types of output devices currently available.

#### ***Secondary-storage operations:***

At this stage of operations, data, information, and the instructions used for the processing of data are stored temporarily or permanently in primary or secondary storage devices. Data and instruction are stored in primary storage devices during processing to allow easy access and fast processing. Secondary storage devices are used for to store data or instructions more permanently. An example of a primary or internal storage device is RAM (Read Only Memory). While and example of a secondary or external storage device is a flash disk.

#### ***Communication Operation***

There may be a need to transmit already processed data to and end user or a output device in a remote location. Similarly, data may be gathered from a remote location. The communication operation of the computers performs these by using the communication hardware. This facilitates the connection between computer and between groups of connected computers called network. Computers linked together can share hardware,



programs and data. Though, computers can operate as standard alone machines i.e. they may not be connected to anything else. By connecting them in a network more benefits are realized.

## **MODULE 1 UNIT 2: BASIC HARDWARE COMPONENTS AND MEMORY CAPACITY MEASUREMENT**

### **1.0 Introduction**

To function properly, the computer needs both hardware and software. Thus a computer system is made up of hardware and software. The hardware consists of the electronic and electromechanical parts of the computer. This is the tangible or touchable part of the computer. These components include: input devices, central processing unit, main memory, secondary memory, and output devices. Two components handle processing in a computer: the central processing unit, or CPU, and the memory. Both are located on the computer system board, or motherboard, the circuit that connects the CPU to all other hardware devices. The processor works hand in hand with other circuits known as main memory and registers to carry out processing. The basic unit of information representation in the computer is the bit.

### **2.0 Objectives**

At the end of this unit, you should be able to:

- Identify and describe the basic hardware components of a computer system
- Explain the units of measuring computer memory

### **3.1 Input Device**

An input device is used to communicate data and instructions or programs to the computer. The input device accepts the data and converts them into machine understandable forms that can be processed by the computer.

### **3.2 The Central Processing Unit (CPU)**

The CPU is the brain of the computer and it is the place where data is manipulated within the computer system. In a microcomputer, the entire CPU is contained on a tiny chip called a microprocessor which is usually mounted on a piece of plastic with metal wires attached to it. The processor consists of two functional units: the control unit and the arithmetic-logic unit.

### **3.3 The Control Unit**

All the computer's resources are managed from control unit. The control unit of the processor coordinates all activities of the computer system. It directs the movement of electronic signals between arithmetic-logic units and main memory and also directs the electronic signals between main memory and the input and output devices.

### **3.4 Arithmetic and Logic Unit**

The computer can perform only two types of operations: arithmetic operations and logical operations. The arithmetic-logic unit is the electronic circuitry capable of performing these two basic logical operations. Arithmetic operations include addition, subtraction, multiplication, division and exponentiation. All data input into the computer system must

first be converted into numbers (0 and 1) before they are processed. In addition to arithmetic operations the computer is able to perform logical operations on two or more numbers. Logical operations are comparisons operations. The ALU compares two pieces of data to see whether one is less than < , equal to = , greater than ( > ), the other. Other comparison operations are greater than or equal to ( >= ), less than or equal to ( <= ) or not equal to ( )

### **3.5 Memory**

The CPU main job is to follow instructions encoded in programs but it does not have the capability to store entire programs or large sets of data permanently. However, the CPU has registers which are devices capable of holding a few bytes of data or instructions at a time. CPU needs to have millions of bytes of space where it can hold programs and the data being manipulated while they are being used. This area is called memory.

Computers use two types memory- primary and secondary memory. A computer s primary memory is an electronic device that store information necessary for a program to run. This consists of the volatile memory (RAM) and the non volatile memory (ROM).

#### **RAM**

RAM is an acronym for Random Access Memory. It is used for short-term storage of data or program instructions. RAM is volatile: Its contents will be lost when the computer s electric supply is disrupted by a power outage or when the computer is turned off. The purpose of RAM is to hold programs and data while they are in use. Physically, RAM consists of some chips on a small circuit board. A computer does not have to search it entire memory each time it needs to find data. Access to the data is usually direct using it memory address. The main functions of the RAM chip include:

- It holds data for processing
- It holds instructions (the programs) for processing
- It holds data that has been processed (useful information) and is waiting to be sent to an output device.

#### **ROM**

This is an acronym for Read Only Memory. This is a nonvolatile type of memory. Nonvolatile chips always hold the same data; the data in them can be read and used-it cannot be changed. One important reason a computer needs ROM is so that it knows what to do when the power is first turned on. Among other things, ROM contains a set of start-up instructions, which ensure that the rest of memory is functioning properly, check for hardware devices, and check for an operating system on the computer disks drives. Unlike the RAM, which is constantly being written on an erased, ROM cannot be written on or erased by the computer user. In addition ROM chips remember, permanently, information supplied by the manufactures such as the information about the manufacturer.

Three variations of ROM chips are used in special situation-PROM, EPROM and EEPROM

**PROM :**

This is an acronym for programmable read-only memory. These are blank chips on which the buyer, using special equipment, writes the programs. Once the program is written, it cannot be erased. Some microprocessor software packages come on PROM units. PROM chips are used by manufactures as control devices in their products.

### ***EPROM***

EPROM stands for electronically erasable read-only memory. They are like PROM chips except that the contents can be erased, using special equipment and new materials can be written. EPROM chips are used for intelligent device control, such as in robots, where the program may have to be modified on regular basis. Programs in EPROM chip can be erased and reprogrammed.

### ***EEPROM***

This is an acronym for electronically erasable programmable read-only-memory. The memory chips can be reprogrammed can be erased by exposing it to an electrical charge. The advantage of EEPROM chips is that they need not be removed from the computer to be changed.

***Registers*** These are special, high-speed storage area within the CPU. All data must be represented in a register before it can be processed. For example, the control unit might load two numbers from memory into the registers in the ALU. Then it might tell the ALU to multiply the two numbers (arithmetic operations), or to see whether the number are equal (a logical operation). The number of registers that a CPU has and the size of each help determine the power and speed of a CPU. For example a 64-bit CPU is one in which each register is 64 bits wide. Therefore, each CPU instruction can manipulate 64 bits of data.

## ***3.6 Measuring the Memory Capacity of Computer System***

### ***Bits***

The term bit is a short form of binary digit. A bit is the smallest possible unit of data. To represent anything meaningful the computer needs groups of bits.

### ***Bytes***

A group of eight bits is called a byte. It is the next larger unit of data representation within a computer system. With one byte, the computer can represent up to 256 different values because it is possible to count from 0 to 255 with 8 binary digits. The byte is an important unit because it can be adequately used to represent any character on the keyboard, including all the letters (uppercase and lowercase), number, punctuation marks, and other symbols.

### ***Kilobyte***

A kilobyte, abbreviated K or KB represents approximately 1000 bytes (or characters). The actual value of 1 Kilobyte is 1024 ( $2^{10}$ ) bytes

### ***Megabyte:***

A megabyte, abbreviated M or MB and sometimes called meg is used to refer to about 1 million bytes of data.

### ***Gigabytes:***

A gigabyte, G or GB, often pronounced gig-a-bite, is used to refer to about 1 billion bytes of data.

### ***Terabyte:***

A terabyte, T or TB, is used to refer to about 1 trillion bytes, or 1000 gigabytes of data.

### ***Petabyte:***

This is a new measurement which accommodates the huge storage capacities of modern database- It is used to refer to about 1 million gigabyte data

## ***MODULE 1 UNIT 3: BASIC SOFTWARE FOR DATA PROCESSING***

### ***1.0 Introduction***

Apart from the hardware of a computer discussed in Unit 2, a computer system requires software to make the best use of its capability in solving problems. Computers do not have intelligence of their own; they rely on the instructions and data supplied by a user in order to perform a task. These instructions are called software. Software consists of a group of related programs written in a specific code called programming language. Software can be purchased as Commercial, off-the-shelf (COTS) or developed in house for data processing purposes. Software purchased to perform a general business functions is often referred to as a software package.

### ***2.0 Objectives***

At the end of this unit, you should be able to:

- Explain the term software
- Describe some basic software for data processing

### ***3.1 Software systems***

Software consists of a group of related programs written in a specific code called programming language. Software is created by professional software programmer and comes on disk, CD-ROM, or downloadable online across the Internet. Software can generally be divided into two categories.

- System software
- Application Software

#### ***3.1.1 Systems Software***

This is software designed to allow the computer to manage its own resources and run the hardware and basic operations. They may be complex to develop and are provided by the manufacturers of computer systems to complement computer hardware. However, others may be purchased from vendors. This software runs the basic operations; it lets the CPU communicate with the keyboard, the screen, the printer, and the disk drive. However, it does not solve specific problems relating to a business or a profession as applications software. Some examples of system software are operating systems and Language translators

##### ***a. Operating System***

An operating system is the most important software for a computer. It is made up of many component programs and manages the overall operations of a computer system. Some of the tasks performed by the operating system are:

- It controls various input and output devices and coordinates input and output operations
  - It Manages the systems resources such as available memory space in the primary and secondary storage devices
  - It allocate memory and processor time to programs
  - It validates users and ensure that the resources a user is requesting is made available
- Some popular operating systems are Windows, Operating System 2, Macintosh Operating System, Novell Netware, UNIX, and Linux.

### ***b. Language Translator***

This is software which translates a computer program written in a high-level programming language, or assembly language to a machine understandable form. If you write a program say in BASIC language, it cannot be executed directly without a translator. A program that is written in high level language is a source program. A source program in its translated form made up of machine codes is called object program. High level languages are either translated from source code to object code by an Interpreter or a compiler. An interpreter translate one line of a high-level instruction and immediate execute the code, proceeding to the next until the entire program is executed. A compiler on the other hand, scans the entire program first and then translate it into machine (object) code.

### ***c. Utilities***

These are programs which are often required by man y application programs. Examples include programs for formatting or defragmenting a disk. Also we have the sort/merge utilities which are used for sorting and merging large volumes of data.

## ***3.1.2 Application Software***

Application software includes programs that are developed using systems software in order to achieve some tasks. There are two types of applications software: application programs that you purchase for solving special classes of problems and application programs that you write to solve your own problems. Some commercially available end user programs (application software) which are developed by professional programmer teams can be used with little or no additional programming skills. Examples of application software include word processing and desktop publishing software; spreadsheet; database programs; graphics programs; communication software; and special-purpose programs suitable for accounting, scientific and engineering application, education, and entertainment and so on.

## ***3.2 Software for Data Processing***

The computers in a data processing environment would run one or more of the following software to facilitate data processing activities.

### ***a. Word processing software***

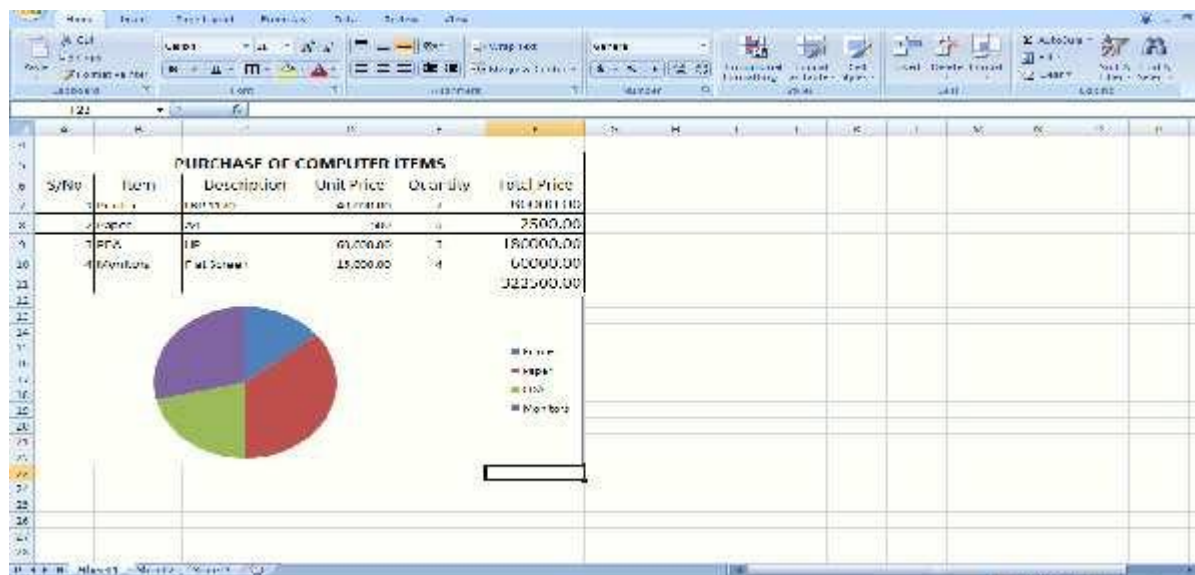
This allows you to use computers to create, edit, sore and print documents. You can

easily insert, delete, move word, sentences, and paragraphs without ever using the eraser. Of all computer applications, word processors are the most common. Word processing also offers a number of features for dressing up documents with variable margins, type sizes, and style. The user can do all these manipulations on a document or report on the screen before printing it. For example, if you make a typing mistake, you simply move the cursor to the position and correct the mistake. You can move a section of your document to another easily and apply different kind of formatting on your document. Word processing software also features for spelling and grammar, Insert, delete, copy, pastes, find, replace, search, cut, copy, format, and printing. Some commonly used word processor are Microsoft Word, WordPerfect, AmiPro, etc

### ***b. Spreadsheet Software***

Traditionally, spreadsheet was simply grid of rows and columns on special paper that was used by accountant and others to produce financial projections and report. Electronic spreadsheets application allows users to enter data in rows and column, calculate means perform statistical analyses, create tables and produce other financial schedules. Spreadsheet software also has features that allow the creation of analytical graphics.

When viewed on a monitor or printed out, analytical graphics, or business graphics, help make data and the generated report easily to comprehend and analyzed for decision making by management. Electronic spreadsheets have features that allow charts and graphs to be created from a table of numbers to show the significance of a selection of data which can be displayed in a number of ways: bar graphs, line graphs, and pie charts etc. A spreadsheet document is called a worksheet. Some commonly used Spreadsheet software are VisiCalc, Lotus 1-2-3, and Microsoft Excel etc.



**Figure 1: Microsoft Excel Chart**

### ***c. Database Software***

A database management system (DBMS) is a collection of programs that allows you to store modify and extract information from a database. A database is a collection of interrelated files in a computer system. With a DBMS, activities such as updating, deleting, adding, and amending of records are easily performed. Databases can also be seen as electronic filing cabinets that allow systematic storage of data for easy access,

reporting, and retrieval of records. Some commonly used database software are MS Access, Informix, Sybase, Oracle, DB3

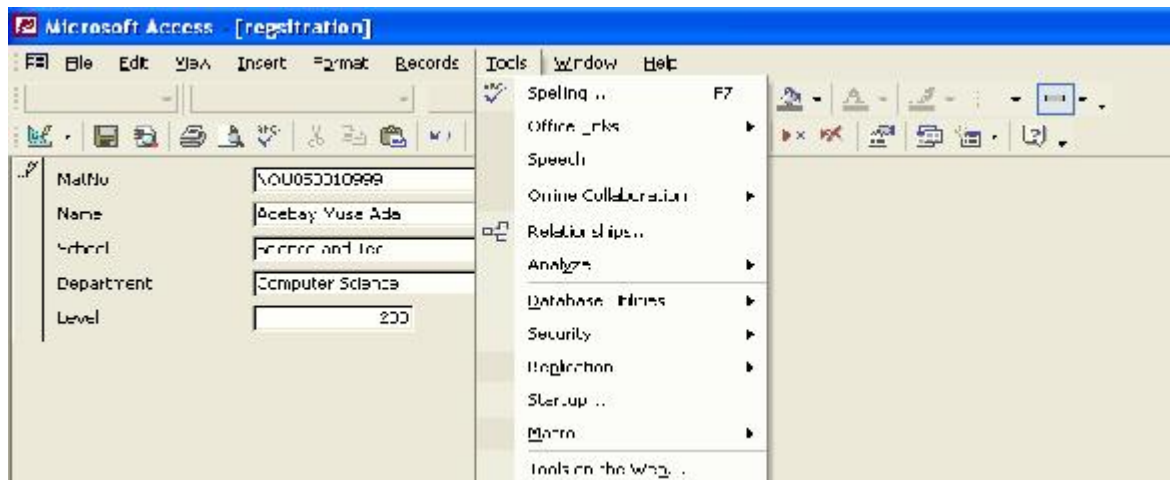


Figure 2: Microsoft Access Database

#### d. Presentation graphics

Presentation graphics enable users to create highly stylized graphics for slide show and report. Presentations software may make use of some analytical graphics (bar, line, pie chart) and also allow you to use electronic painting and drawing tools for creating lines, rectangles and others forms of shape. Most presentation software have features that allows you to add text, animated sequences, and sound to your report. Your report or presentation can be sent to the screen or printed on transparency acetates. Some presentation software packages provide artwork called clip art that can be electronically cut and pasted into the graphics. Some commonly used presentation graphics are Examples of well known presentation packages are Microsoft PowerPoint, Aldus Persuasion, Lotus Freelance Graphics, and SPC Harvard Graphics.

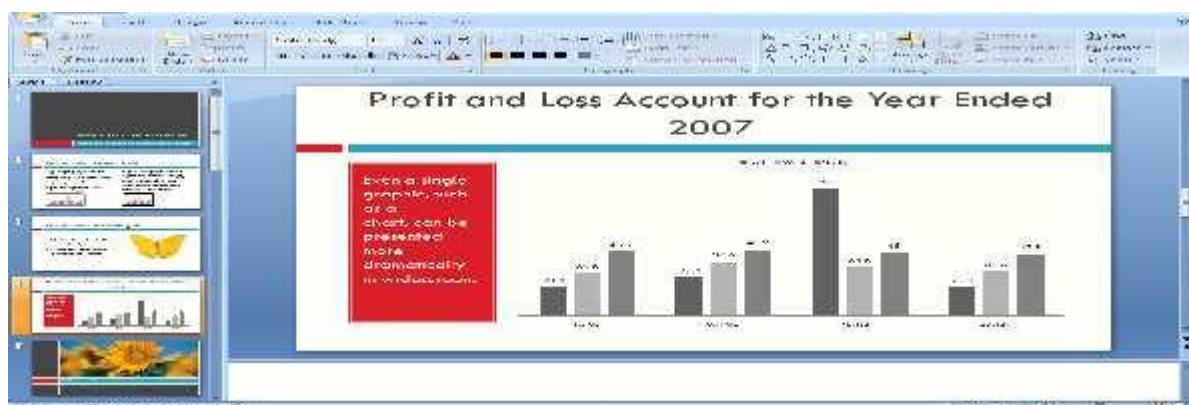


Figure 3: Microsoft Power Point Presentation

## **MODULE 1 UNIT 4: DATA CAPTURING AND REPORTING DEVICES**

### **1.0 Introduction**

Computers interact with its external environment and users via the Input / Output devices. Inputs are the signals or data sent to the system, and outputs are the signals or results received from the system to the outside. Input devices are hardware equipment by which a computer system receives data while output devices are hardware equipment by which results are communicated to the outside world.

### **2.0 Objectives**

At the end of this unit, you should be able to:

- Identify and describe data capturing devices
- Identify and describe data reporting/output device

### **3.1 Input Devices**

For data to be processed, they are first to be captured via input devices and translated to computer understandable form. Input devices are use to capture and translate data into a form the computer can process. Though the data may be captured in the forms of text, characters, letters, images, sound, pictures and video, the translated data will be in the form of 0s and 1s, or off and on signals which can be easily processed by the computer. Input devices exist in the following categories:

#### **3.1.1 Keyboard**

Keyboard is used for entering text and numeric values into the computer. It is the most popular means of data entry in microcomputers. The keyboard may look like a typewriter but with some special keys added. The keyboard has facilities for converting letters, numbers and other characters into electrical signals that are machine-readable the processor of computers. Data entry functionality is also available on via keys on a banks automatic teller machine or the keypad of a personal digital assistant or a cell phone.

#### **3.1.2 Pointing Devices**

These constitute the second category of input hardware. These devices control the position of the cursor or pointer on the screen. Pointing devices include:

##### **a.) Mouse**

A mouse is an input device that is popularly used with microcomputers. When used with desktop computers it is rolled about on a desktop to direct a pointer and select commands on a computer display screen. The point-and-click actions of the computer mouse is fast making it an alternative to keyboard and text-based commands since most of the commands executed with a keyboard can almost be done with a mouse. In addition, the mouse can be used to draw images on the screen.

##### **b.) Trackball**

This is a variant of the mouse. It can be moved on top a stationary device that is rotated with the fingers or palm of the hand. Trackball are commonly used with hand held devices.

##### **c.) Touchpad**



Touch pads allow a user to control the cursor/point on the computer display screen with his fingers. It is about the same size as a mouse, but is usually, rectangular and flat devices that use very weak electrical field to be activated.

#### ***d.) Joysticks***

A joystick is a pointing device that consists of a vertical handle like a gearshift lever mounted on a base with one or two buttons. Joysticks are used principally in video games, in some computer-aided design systems, and in computerized robot systems.

#### ***3.1.3 Pen-Based Devices***

Pen-based computer devices use a pen-like stylus to capture a user's handwriting and marks into a computer. This is one of the means by which data is entered into small handheld computers such as personal digital assistants (PDAs).

#### ***e.) Light Pen***

This is a light-sensitive stylus, pen-like device that can be connected by a wire to the computer system. A stylus is a pen-like device with which a user sketches an image. When used to capture data, the user brings the pen to desired point on the display screen and presses the pen button, which identifies that screen location to the computer.

***f.) Digitizing Tablets*** A digitizing tablet consists of a tablet linked by a wire to a stylus or puck. A puck is a copying device with which the user copies, or traces, an image. Digitizing tablets are used to capture data directly into the computer for processing.

#### ***3.1.4 Source-Data Entry***

Keyboard entry requires typing; errors could be introduced by the operator during data entry. Non-keyboard source-data entry devices such as scanner are used to minimize data entry errors. The following categories of devices are used:

#### ***Scanning Devices***

Scanners use laser beams and reflected light to capture and translate hardcopy images of text, drawings, photos, and the like into computer understandable form for processing. Scanning devices include Mark and character recognition devices, Fax machines and Imaging systems. Specifically, the Mark and character recognition devices are usually referred to by their abbreviation OCR, OMR and MICR,

#### ***g.) Optical mark recognition***

Optical mark recognition (OMR) uses a device that reads pencil marks and converts them into computer-understandable form. This technology is widely accepted by many examination bodies such as Joint Admission and Matriculation Board (JAMB), West African Examination Council (WAEC), NECO (National Examination Council Organization) and other Scholastic Examination bodies across the world as a convenient way of processing their candidates results.

#### ***h.) Optical Character Recognition***

Optical character recognition (OCR) devices are used to translate designed marks character, and code into digital form for processing by the system. The most widely used optical code is the bar code, which is used in point-of-sale systems in supermarket, shopping mall, hospital, libraries, military operations, utility bills and price tags on departmental-stores merchandized and transportation facilities. The code can include time, date, location, and identification data for deriving appropriate information for

decision making. It is usually more convenient and faster to use OCR than retyping a document into the computer.

### ***i.) Magnetic Ink Character Recognition (MICR)***

This Technology is used in cheques processing by most banks and other financial institutions. An MICR device has facilities to interpret and translates magnetic ink character on Bank cheque which contains characters identifying the bank, cheques number etc into computer understandable form.

### ***j.) Fax Machine***

A fax machine also referred to as facsimile transmission machine has facilities that scan an image on paper into electrical signal, transmit same over telephone lines and re-creates the image to a receiving fax machine on paper.

## ***3.2 Other Input devices***

Sensors are devices that collect data directly from the environment for input into a computer system. For instance, atmospheric data can captured and transmitted via sensor network to a data processing centre for weather forecasting.

### ***h.) Voice input***

These devices convert spoken words into computer understandable digital form for processing. When voice recognition systems are used to capture data for processing, the user documents can be created by speaking words into a computer rather than keying them in. The system will have audio facilities such as sound card, microphone, and speakers.

The physical component or materials which data is stored are called storage media. The hardware components that write data to, and read it from, storage media are called storage devices. For example, a diskette is storage medium, whereas a diskette drive is a storage device. Storage media and derives have evolved dramatically since computers were in their infancy, and this pace has accelerated since the introduction and growing popularity of PCs.

## ***3.3 Output Devices***

Outputs are the outcome of processed data. An output device is any piece of computer hardware equipment used to communicate the results of data processing carried out by a computer to the outside world. Some common output devices include: Visual display unit A visual display unit (also called VDU, monitor, or screen), Printers and Plotters.

### ***a.) Monitors***

The first is the typical monitor that looks like the television screen and used uses a large vacuum tube, called a cathode ray tube (CRT). Its operation is very similar to that of television picture tube, with an electronic gun shooting a beam of electrons to illuminate the pixels (picture elements) on the screen. The resolution is determined by the number of pixel per screen. The resolution of a monitor indicates how densely the pixels are packed. The second type, known as flat-panel display, is used with laptops or notebook computers. While the earlier monitors could only display one color i.e black, grayscale, etc modern monitors have good support for color and graphics display though this may require more memory.

### ***b.) Printers***

Printer is a device that prints text or illustration on paper and in many cases on

transparencies and other media. They are the most popular output devices beside monitors. The speed of early printers was measured in units of characters per second while that of modern printers are measured in pages per minute. There are different types of printers based on the technology they use. We shall only describe a few that you will find in a data processing center:

#### ***i.) Dot-Matrix Printer***

This operates by striking pins against an ink ribbon to create the required characters. Each pin makes a dot, and combination of dots form characters and illustrations. Dot matrix printers are cheap, relatively fast but they do not produce high-quality output.

#### ***ii.) Line Printer***

This type of printer is able to print an entire line at one time with as much speed as 3,000 lines per minute. The print quality is relatively low and they are very noisy.

#### ***iii.) Thermal Printers***

This type of printer produce images by pushing electrically heated pins against special heat-sensitive paper. Thermal printers are economical and are used in most calculators and many fax machines. They produce low-quality print and paper requires special handling of papers.

#### ***iv.) Ink-jet Printer***

This work by spraying ionized ink at a sheet of paper. Magnetized plates in the ink's path direct the ink onto the paper in the desired shapes. They are popular with as portable printers because they could exist in small sizes. Ink-jet printers require a special type of ink that may smudge on low-cost copier paper. They provide an economical way of way to print full-colour documents.

#### ***v.) Laser Printers***

Laser printer utilizes a laser beam to produce an image on a drum. The light of the laser alters the electrical signal on the drum wherever it makes a hit. The drum is made to roll through a reservoir of toner, which is picked up by the charged portion of the drum. After which, the toner is transferred to the paper through a combination of heat and pressure. LaserJet printer produces very high quality output. This is as a result of the resolution that is, how many dots per inch (dpi) they lay down. The available resolutions range from 300 dpi at the low end to 1,200 dpi at the high end. The color laser printers are able to produce colored outputs though they are usually more expensive because of the addition toners required.

Laser printers are non-impact printers; they are not as noisy as dot-matrix or line printers. They are relatively fast, although not as fast as some dot-matrix printers. Their speed ranges from 4 to 20 pages of text.

#### ***c.) Plotter***

A Plotter, though mainly used for engineering applications uses pen to draw picture or lines on paper based on commands from a computer. The can produce continuous lines, whereas printers can only simulate lines by printing a closely spaced series of dots. Plotters are not as fast as printer but are useful in producing large-size chart, maps, or drawings even in color.

#### ***d.) Speaker***

A voice output device such as a speaker converts digital output data back into intelligible

speech. They are use to deliver audio output from animation, multimedia application from the web and music to the user.

***e.) Microfilms and Microfiche***

Microfilm and microfiche were used to store large quantities of output as microscopic filmed document, but they are now being replaced by optical disk technology.

## MODULE 2 UNIT 1: DATA STORAGE DEVICES

### 1.0 Introduction

The computer primary memory is limited in the amount of information it can store at a time. In addition, the content of the RAM is temporary i.e., once the power of the computer is turned off, all the data and program within it simply vanishes. This is why data must be stored on secondary storage devices which are able to hold data and programs on a more permanent basis. External storage devices consist of large-capacity, slow-access data storage attached to a digital computer and used to store information that exceeds the capacity of main storage. Two main technologies are used to store data today: magnetic and optical storage. Although devices that store data typically employ one or the other, some combine both technologies. The primary types of magnetic storage are:

### 2.0 Objectives

At the end of this unit, you should be able to:

- Identify and describe data storage devices
- Parameters for disk performance

### 3.1 Diskette

A diskette or floppy disk is a data storage medium that is composed of a disk of thin, flexible magnetic storage medium encased in a square or rectangular piece of mylar plastic. Floppy disks are read and written by a floppy disk drive or FDD, the initials of which should not be confused with "fixed disk drive", which is another term for a hard disk drive. Floppy disk exist in the following sizes; 8-inch (200 mm), 5¼-inch (133 mm), and the most common 3½-inch (90 mm). Though floppy disks are still used in some data processing environment, they are now being largely superseded by flash and optical storage devices while some users consider emails as a convenient way of exchanging small to medium size digital files. Floptical drives combine magnetic and optical technologies to store about 21MB of data on a media similar to 3½-inch floppy disks. The name is a portmanteau of the words 'floppy' and 'optical'. This device was introduced in 1989 by Insite Peripherals of San Jose, but did not become popular, because it of the limited storage capacity it offer. A similar technology was used in the Laser Servo-120 drive introduced in 1996 with 120MB capacity.

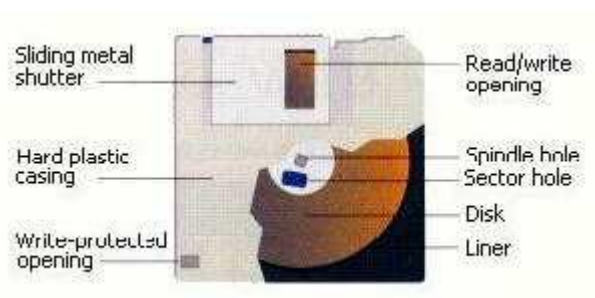


Figure 1: Floppy Disk

### 3.2 Hard disks

In a data processing environment, a hard disk drive (HDD), commonly referred to as a hard drive, hard disk or fixed disk drive, serve as permanent storage device for large amount of data. Originally, the term "hard" was temporary slang, substituting "hard" for

"rigid", before these drives had an established and universally-agreed-upon name. The hard disk drive (often shortened to "hard drive") and the hard disk are not the same thing, they are packaged as a unit and so either term is sometimes used to refer to the whole unit. A hard disk is really a set of stacked "disks," each of which, like phonograph records, has data recorded electromagnetically in concentric circles or "tracks" on the disk. A "head" (something like a phonograph arm but in a relatively fixed position) records (writes) or reads the information on the tracks. Two heads, one on each side of a disk, read or write the data as the disk spins. Each read or write operation requires that data be located, which is an operation called a "seek." (Data already in a disk cache, however, will be located more quickly.) Modern computers come with a hard disk that contains several billion bytes (gigabytes) of storage.



Figure 2: Hard Disk

### ***3.3 Removable hard disks***

This is a variation of the hard disk in which hard disks enclosed in plastic or metal cartridges are easily removable like floppy disks. It combines the best features of hard and floppy disks. They are used to provide large economical, high, fast, and portable storage facility for data processing.

### ***3.4 Magnetic tape***

Magnetic tape has historically been found more convenient means of large data storage over disk where media portability or removability is required for backup. Magnetic Tape uses the same read/write techniques as disks. Data is stored on flexible mylar tape covered with magnetic oxide. Data is stored in parallel tracks of 9, 18 or 36. Data on tapes are accessed sequentially. Tapes provide slow, very cheap, large capacity backup for data. The rapid advances in disk storage technologies resulting in and the improvement in disk storage density, and reduced price, coupled with arguably declining innovation in tape storage technology, has reduced the market share of tape storage devices.



Figure 3: Tape

### 3.5 Optical Storage

The optical storage devices such as CDs and DVDs are means in which data is written and read with a laser for archival or backup purposes. The optical storage devices are fast replacing both hard drives in computers and tape backup in mass storage. This is because optical media are more durable than tape and less vulnerable to environmental conditions lasting up to seven times as long as traditional storage media. However, at present optical media are slower than typical hard drive speeds, and offer lower storage capacities. Optical disk capacity ranges up to 6 gigabytes i.e. 6,000,000 bytes which is far more compared to the 1.44 megabytes (MB) i.e., 1,440,000 bytes offered by a floppy disk. A newer technology, the digital versatile disc DVD, has about 4.7 gigabyte storage capacity on a single-sided, one-layered disk compared with 65 gigabyte of storage for a CD-ROM disk. Invariably, they can be used to hold large amount of data.



Figure 4: Optical Storage

A USB flash drive is a NAND-type flash memory data storage device integrated with a USB (universal serial bus) connector. USB flash drives are typically removable and rewritable, much shorter than a floppy disk (1-4 inches or 25-102 mm), and weigh less than 56g. Their storage capacities typically range from 64MB to 32GB or more. They have 10-year data retention. USB flash drives offer potential advantages over other portable storage devices, particularly the floppy disk. They are more compact, faster, hold more data, are more reliable for lack of moving parts, and have a more durable design.



Figure 5: A Flash

### 3.6 Accessing Data from Disk

Bits of data (0 s and 1 s) are stored on circular magnetic platters called disks. A disk rotates rapidly (& never stops). A disk head reads and writes bits of data as they pass under the head. Often, several platters are organized into a disk pack (or disk drive). Disk contains concentric tracks. Tracks are divided into sectors. A sector is the smallest addressable unit in a disk.

When a program reads a byte from the disk, the operating system locates the surface, track and sector containing that byte, and reads the entire sector into a special area in

main memory called buffer. The bottleneck of a disk access is moving the read/write arm. So it makes sense to store a file in tracks that are below/above each other in different surfaces, rather than in several tracks in the same surface.

A cylinder is the set of tracks at a given radius of a disk pack. i.e. a cylinder is the set of tracks that can be accessed without moving the disk arm. All the information on a cylinder can be accessed without moving the read/write arm.

### **3.7 Disk Performance**

In measuring the performance of a storage device, we may consider the following three parameters.

#### **a.) Rotational delay/Latency**

This is the time it takes to position the proper sector under the read/write head. In general, it is used to refer to the period of time that one component in a system is spinning its wheels waiting for another component. Latency, therefore, is wasted time. It makes sense to separate read latency and write latency, and in case of sequential access storage, minimum, maximum and average latency. Consider a hard disk which rotates at about 5000 rpm i.e. one revolution per 12 msec. The average latency can be calculated as follows:

Min latency = 0

Max latency = Time for one disk revolution

Average latency (r) = (min + max) / 2

= max / 2

= time for ½ disk revolution

Typically 6 8 ms average

#### **b.) Throughput**

This is the rate at which information can be read from or written to the storage. It is expressed in terms of megabytes per second or MB/s. A media accessed sequentially, as opposed to randomly, typically yield maximum throughput.

#### **c.) Seek time**

This is the amount of time between when the CPU requests a file and when the first byte of the file is sent to the CPU. Seek times between 10 and 20 milliseconds are common.

## **MODULE 2 UNIT 2: FILE ORGANIZATION**

### **1.0 Introduction**

It is not enough to acquire excellent computer system; hardware and software. Files within the system must be well maintained, managed and organized. Good file organization leads to productivity. When files are well arranged and maintained, users can easily access and retrieve the information they need in good time. Different ways of file organization are: Serial File Organization, Sequential File Organization, Indexed Sequential Organization, Direct or Random File Organization.



## 2.0 Objectives

At the end of this unit, you should be able to:

- Explain the meaning of a computer file
- Identify and describe file organization techniques
- Describe the type of data files used in a Data processing Environment

### 3.1 What is a Computer file?

A file is collection of data or information that has a name called the filename. Most of the information stored in a computer system is stored as files. A file is often stored with a user s given name and a system supplied extension. The name of the file should reflect the content of the data stored in the file. For example, payroll or result File. The extension of the file should reflect either the type of file (i.e. program file, image file, audio file etc) or the software used to create the file (e.g MS WORD document, MS EXCEL Worksheet, BASIC Compiler). For example;

#### S/No Type of File / Software Extension

1	Microsoft Word	.doc Document
2	Audio file / Video	.wav , avi, mp3 3 BASIC compiler .bas
4	Text	.txt
5	Microsoft Excel	.xls
6	Images	.png , jpg , gif

Table 1: Computer Files

The different types of file stored in a computer system are as follow:

#### a.) Program File

These are files that store sets of instructions written in a programming language. A source program file for examples contains the instructions written in a high level language such as BASIC or FORTRAN programming language by a programmer while the object file is the translated form of the source file in machine code after. The files that contain the machine code are called executable files (or binary files)

#### b.) ASCII File

ASCII stands for American Standard Code for Information Interchange. ASCII files are text-based files. That is the characters are represented in ASCII code (without formatting such as underline, italics, boldface or graphics). Files stored in this format are used to to transfer documents between incompatible computer platforms, such as IBM and Macintosh.

#### c.) Image File

Documents containing digitized graphics or images are stored in this format.

Image files: Image files contain digitized graphics

Audio and video files: Audio files contain digitized sound while video files contain digitized video images and animation.

#### d.) Audio and Video File

This is a file that is used to store digitized sound or digitized video images and animation.

### ***e.) Data File***

This refers to document files, contain data, not programs. Their content are using application software.

## ***3.2 File Organization***

File are created, arranged and maintained in data processing systems for a purpose of easy retrieval. Computer systems store files permanently on secondary storage devices. Records or files are arranged in several ways on the storage media, and the arrangement determines the manner in which individual records can be accessed or retrieved. Four common ways of file organization and access are:

### ***a) Serial File Organization***

In this method of file organization records are not arranged in any specific order. If magnetic tapes are used for the storage of data, it would be necessary to wind the tape forward and backward to access a given record since access can only be made in the sequence in which the records were physically stored on the tape, i.e serially. And if records are stored on disk, a full index will be required to access any given record. This method of file organization is therefore inefficient.

### ***b.) Sequential File Organization***

In this type of file organization, data records are normally stored in ascending order of key field. Data must be retrieved in the same physical sequence in which they are stored. It is the only file organization method that can be used on magnetic tapes. Magnetic tape is a sequential storage device. That is, records and files are stored in magnetic tape in a sequential order. They are also read in sequential order. Note that records may also be stored sequentially on disk if desired. Serial and sequential file access mean the same thing in respect of files stored on tapes when stored in sequence but this may not be the same case with disk files as the records accessed serially may not be defined in a key sequence. This sequential file organization method is no longer a popular approach of storing or access record in a file.

### ***c.) Indexed Sequential Organization***

This technique of file organization uses both the sequential and direct access methods. It is widely applied to the storage of record on magnetic disk. It allows sequential file to be manipulated serially as the record are stored in ascending order of key field. In addition it allows direct access storage devices, to be accessed directly using the indexed sequential access method (ISAM). This access method relies on an index or key fields to locate individual records. An index to a file is similar to how of a book can be used to locate it physical position //on a library shelf. The method requires that data are stored in magnetic or optical disk. For //example, a university could index certain ranges of student s matriculation number 0000 to 1000, 1001 to 2000 and so on. For the computer to find the record with the key field 8888, it would go first to the index, which would give the location of the range in which the key field appears (for example, 8001 to 9000). The computer would then search sequentially (from 8001) to find field 8888

### ***d.) Direct or Random F ile Organization***

This file organization is utilized with magnetic disk technology. Most computer applications today use this approach for storing records in computer files. In this

approach, individual records are stored in a particular sequence of key fields. Thus allowing users to access records in a sequence they desire, without regards to actual physical order on magnetic tapes or disk. With this approach, every record has an address that makes it possible to locate it independently of other records on the storage media. To allow easy access and retrieval of information an index or table of the key are maintained with the relative record number of the record in storage. In order to retrieve a record, the actual key is looked up in the index with the corresponding record number of the record that matches the key. Once this is found, the address in storage is worked out and the records accessed. Records stored with the technique are much faster to be accessed than records store with the sequential file organization. Though they may be more expensive because optic or magnetic disk that ma y be involved for their storage.

### ***3.3 Classification of Storage Devices***

Storage devices can be classified generally as sequential access or random access. For example, a tape drive is a sequential-access device because to get to record 5 on the tape, the drive needs to pass through points 1, 2, 3 and 4. A disk drive, on the other hand, is a random-access device because it allows the record at any position on the disk to be accessed without passing through all intervening positions

Figure 1: Storage Device Access

### ***3.4 Types of data files***

Data stored in a data processing centre could be considered as transaction file and master file

#### ***a.) Transaction file***

This refers to a collection of transaction records. The transaction file is a temporary holding file that stores records that normally have a limited useful lifetime. For an employee file for a payroll processing, a transaction file would hold, the name, contact information, hour worked, pay rate, tax, utility bills etc for a staff for a particular month. At the end of every month or so, there will be a need to compute the salaries of the staff from the information in the transaction file. After the transactions are successfully carried out, the information in the transaction file will be use to update the master file. In a data processing system, transaction records ma y be retained online for some period and later achieved off-line on permanent storage devices. Transaction files can serve as audit trails and history for the organization.

#### ***b.) Master File***

Master file is a collection of records that are relatively permanent records that are updated periodically. Thus once a record has been added to a master file, it remains in the system indefinitely. The value of fields for the record will change over its lifetime, but the individual records are retained indefinitely. Master files contain descriptive data, such as name and address, as well as summary information, such as students Cumulative, Grade Point Average in an examination processing system or total net pay, total tax deductions in a payroll system. The changes to be made to a master file could be addition of records, deletion of record or update of record. In an organization this will occur when a new staff joins the work force or when a staff resigns his appointment. Furthermore, the master file of a payroll system may be composed of discrete pieces of information (such as a name,

address, or employee number etc) called data elements. Data are keyed into the system, updating the data elements periodically. The elements on the master file are combined in different ways to make up periodic reports of interest to management and government agencies or to generate paychecks sent to the staff at the end of the month. Other examples of master file include: Customer, Product, Result or Supplier file.

## **MODULE 2 UNIT 3: DATA PROCESSING TECHNIQUES**

### **1.0 Introduction**

The processing of data need not be restricted to a particular environment. Data processed in remote locations can be accessed via telecommunications infrastructure. Different kind of processing techniques are made possible by telecommunications. The techniques are discussed in this unit.

### **2.0 Objectives**

At the end of this unit, you should be able to:

- Definition and explanation of the terms; data and information
- Describe a computer system
- Identify the types of computer
- Describe the basic operations of a computer

### **3.1 Meaning of Data Communication**

Data communication refers to the process of transmitting data / information from a source via a telecommunications system in coded forms to a destination for the purpose of interchanging details of business transaction. The transfer of data require some forms of electromagnetic energy such as electricity, light, radio or waves which is transmitted through a physical medium such as the atmosphere, cable or wire

A telecommunication system refers to a collection of compatible hardware and software required to exchange information from one location to another. A telecommunication system is able to transmit text, graphic images, voice or video information. The major components of a telecommunication system are.

- Computer: This is required for data processing
- Terminals: These are any input/output device that send or receive data
- Communication channel: This is the link by which data, graphic images, voice or video are transmitted between sending and receiving devices in a network.

Communication channels use various communications media, such as telephone lines, cable, coaxial cable, fiber-optics and wireless transmission

- Communications software: This controls input and output activities and manages other functions of the communication network

Question: Describe the basic components of a telecommunication system

### **3.2 Batch Processing**

This used to be the dominant form of data processing. It is a method of data processing, in which transaction data such as data from time card is collected over a period say, a month for payroll processing, and then processed all at once as a batch to update the master file. A payroll system keeps tracks of the money paid to employees. With this kind of processing, a user may not be able to receive a reliable response from the data processing system until the batch is processed. For instance, if the pay of a staff is computed at the end of the month, it will not be possible to know exactly how much the staff will earn until all the data elements for computing his pay are collated and

processed. Another example is a banking system where reconciliation of accounts is done only after the close of day operations and the entire transactions is processed as a batch. Any queries performed before the batch is processed will certainly not be accurate. However, many banks in Nigeria have begun to integrate real-time modules into their applications.

### **3.3 Online processing**

This is a method of data processing method whereby data about a single transaction is processed immediately it is captured. This method of processing allows transactions to be entered directly to the system via terminals, PCs or workstations as they take place thereby updating the master file immediately as the transactions occur. The point of entry may be remote from the location at which update his made. For example when you withdraw cash from an automated teller machine, your withdrawal is instantly processed and your account balance updated. Other online processing systems include:

- Visa processing
- Result checking
- Banking (account enquiry)
- Air seat reservation
- Application processing
- Examination

Because of the prevalent of PC in data processing we rarely hear the term online processing. The term client / server computing is more popular where the PC is the client.

### **3.4 Transaction Processing :**

Transaction processing system (TPS) is an automated system that performs and records the daily or routine transactions necessary to conduct a business. A transaction may include an order or payment confirmation. Transaction processing entails the updating of appropriate database records as soon as a transaction is captured into the computer. The processing of a transaction may require that a confirmation be sent to the user or a customer immediately. Transaction processing systems are the backbone of an organization because they update constantly. The instant update is a requirement because at any given moment, a customer may need an inventory balance, an account balance or the result of an examination Transaction processing is also referred to as "online transaction processing" (OLTP). This means that the system should be available 24x7 on reliable computers and networks. Common examples of transaction processing systems include: Sales order entry, hotel reservation systems, payroll, examination processing systems, and shopping systems.

### **3.5 Real-time Systems**

These are systems designed to deal with dynamic situations in order to control a critical operation such as an airline reservation system which must be continually updated as events occur. Seat reservation in flight operations require communication-oriented Server computers supported by network of terminals or PCs serving as clients. These facilitate response to enquiries on seat reservations and ensure that the master file is updated as soon as transactions are completed. The systems ensures that enquiries on available seat are responded to instantaneously and prevents double, or overbooking of seat in the aircraft.

### ***3.6 Distributed Processing***

Many organizations that were used to centralized systems for data processing are now able to adopt Distributed data processing because of advances in computing technologies. A centralized system consists of a central multi-user computer (usually mainframe) which hosts all components of a data processing system. The users interact with this host via terminals or PCs serving as client, but virtually all of the actual processing and work is done on the host computer. All the devices in the centralized system such as PCs, terminals, network devices, and printer converge on one central computer, even though the users may work from remote locations via terminals. All processing and storage take place at the central location. On the other hand, a distributed system allows the components of data processing system to be made available at multiple locations in a computer network. Which means that, the processing workload required for supporting the components is also distributed across multiple computer on the network. In addition, the computers, storage devices, and even some computer personnel may need to be distributed to separate locations throughout the organization for the efficiency of the system. Distributed data processing allows data processing and storage to occur at several locations in the computer system. There are advantages and disadvantages associated with adopting distributed data processing in an organization. And these are as follow:

#### ***Advantages of Distributed Data Processing***

- Modularity                      Easy Integration
- Better response time Data processing is closer to the end user.
- Ability to share data
- Greater reliability
- Local control of data
- Lower cost
- Direct Users interaction

#### ***Disadvantages of Distributed Data Processing***

- Technical problem of connecting and maintaining dissimilar systems
- Need for sophisticated communication system may incur addition cost
- Data integrity and security problem resulting from possible duplication of data on several systems along the network
- Interoperability issues- technical problem of connecting dissimilar machines
- Lack of skilled professional for systems support

## ***MODULE 2 UNIT 4: TRADITIONAL FILE SYSTEM PROCESSING***

### ***1.0 Introduction***

File systems were really an early attempt to computerize the manual systems that were prevalent in the early days of data processing. The manual system works well while there are not many items, but it breaks down when we have a lot of items and we want to cross reference them. So the file based system was a response to industry needs for more efficient data access. However, rather than have everything centralized, each department would have their own set of files. File system which is also referred to as a traditional approach of file processing is inadequate for many organizations because it involves creating, organizing, storing, manipulating and maintaining records within an organization that leads to each functional area or department creating and maintaining its

own data files and programs.

## **2.0 Objectives**

At the end of this unit, you should be able to:

- Explain the concept file system
- Identify and explain the problems with the file system

### **3.1 Traditional File Systems / Processing**

As an organization grows, computer systems and application become more complex. Take an example, a University's computer system that handles student's information. If the data of students are to be kept independently by the units that each student interacts with, for example, Registration, Hostel, Accounts, Examination and Records, Students Affairs, Health Centre etc. Worse still, if each department is allowed to keep students information independently in their application, a time comes, when multiple files containing the same records of students will exist in the different units.

As this process goes on for more than five years depending on the number of student's enrollment, the university is saddled with hundreds of programs and applications without adequate control on the data. The University will be collecting the same information in more than one file. Some of the resulting problems are data redundancy, lack of data integrity and lack program-data dependence, inflexibility, and inability to share data among applications.

### **3.2 Problem with the Traditional File System**

Some problems with the traditional file environment are: Data Dependence, Program-Data Dependence, and Difficulty of Data-Sharing. These are discussed as follows:

#### **a) Data Redundancy:**

This means the presence of duplicate data in multiple data files and often in different format. This is often the result, when different departments are allowed to collect the same piece of information about an object. For instance, within the university environment, the hostels and student registration department might collect the same student's information (Name, Mat No, Level and Address). Because it is collected and maintained in so many different places, the same data items may be repeated in different departments. When data fields are repeated in different files, storage spaces are wasted and a lot of time is spent trying to update the records.

**b.) Program-Data Dependence** Program-Data Dependence is the tight relationship between data stored in files and the specific programs that process the information in the files. Where computer programs become so data specific, any changes in data would also mean any modification of the program that processes them. Such changes could be very expensive in terms of the time and cost of re-programming.

#### **c.) Difficulty of Data-Sharing**

It is difficult to share data in a file environment because it may be difficult to relate the data in one file with that of another within one or several departments where files are kept. In addition, there is no control over the access to data which makes it further more difficult to retrieve the desired information.