

- learn to store data and information in appropriate formats and in suitable media
- acquire skills in retrieving data and information from an information system.

COURSE OBJECTIVES

In order to achieve the aims of this course, a number of objectives are specified for each unit besides the following overall objectives. By the time you complete this course, you should be able to:

- explain the concepts of data, data processing, and information
- distinguish between document and information and describe the processes of documentation
- classify information into subject categories
- analyse data and information for the purpose of assigning it to appropriate subject classes
- correctly assign key words to be used for retrieval purposes
- describe the basic architecture of a computer and the role of computers in storage and retrieval
- describe the storage media in use and the basic structure of records, files, and databases
- explain the functions of Database Management System
- explain the concepts of information retrieval
- discuss user characteristics and user needs, which are fundamental to information storage and retrieval
- correctly analyse requests for information and formulate search strategies
- retrieve information from an information system
- search the Internet
- evaluate the retrieval performance of an information system.

WORKING THROUGH THIS COURSE

To go through this course you are required to read the study units, answer the self-assessment exercises and do the assignment in each unit. The self-assignment exercises are meant to help you to reinforce what you have learnt. It will be very helpful to you to try and answer the questions first before looking at the answers. At the end of the unit, you will find an assignment, which will be marked by your tutor. Work diligently on it and submit your work to your tutor for grading. The tutor-marked assignments will constitute 30% of the total marks of the examination in this course.

You will need to have access to a computer and be familiar with the basic elements of a computer system. In due course, you will have practical exercises in the library and on the Internet, and so, you need to have access to these facilities.

It is expected that you will spend on the average two to three hours to study one unit and about 17 weeks to complete the whole course. However, you should realise that you are actually to work at your own pace. Below are the components of this course.

COURSE MATERIALS

1. Course Guide
2. Study Units
3. Self-Assessment Exercises
4. Tutor-Marked Assignments

STUDY UNITS

There are 18 study units which you will work through in this course. There are as follows:

Module 1

- | | |
|--------|---|
| Unit 1 | Data and Information |
| Unit 2 | Document and Documentation Classification |
| Unit 3 | Classification Unit |
| 4 | Subject Indexing |
| Unit 5 | Indexing Language |

Module 2

- | | |
|--------|--|
| Unit 1 | Computers in Information Storage and Retrieval |
| Unit 2 | Storage Media |
| Unit 3 | Records and Files |
| Unit 4 | Databases |
| Unit 5 | Database Management System |

Module 3

- | | |
|--------|-----------------------------------|
| Unit 1 | Concepts of Information Retrieval |
| Unit 2 | Role of Information Centres |
| Unit 3 | Users and User Needs Information |
| Unit 4 | Searching for Information |
| Unit 5 | Document-Term Matrix |

Module 4

Unit 1	Retrieval from the Internet
Unit 2	Evaluation of an Information Systems and Services: Part I
Unit 3	Evaluation of an Information Systems and Services: Part II

Units 1 to 5 of module 1 address the crucial requirements of understanding the characteristics of information and systematic organisation

Units 1 to 4 of module 2 take up the requirements for information processing and storage; while modules 3 and 4 are devoted to the principles of retrieval and evaluation of retrieval performance.

SELF-ASSESSMENT EXERCISES

These are embedded in the text of the study units. You should be able to answer the questions if you study the sections of the units very well.

TUTOR-MARKED ASSIGNMENTS

The tutor-marked assignments will be supplied to you with the units. It is absolutely necessary that you do the assignments and submit your work to your tutor.

ASSESSMENT

There are two aspects of the assessment of this course. The first aspect is the continuous assessment through the tutor-marked assignment. The second aspect is the final examination. The tutor-marked assignments will constitute 30% of the total marks of the examination in this course. The final examination will come at the end of the course. It will be a written examination that reflects the exact content of the course. The questions will not be different from the types you would have already been familiar with in the self-assessment exercises and tutor-marked assignment. The written examination will carry 70% marks.

HOW TO GET THE MOST FROM THIS COURSE

In this programme, you will not be sitting before any lecturer to receive lectures. The study units will replace the lecturer; you will be reading the study units instead of listening to a lecturer. You have the flexibility of being able to work through specially designed course materials at your own pace. You can also choose your time and place of study. The

contents of the units will give you all the information and direction you need.

The units follow the same format. Each unit begins with a table of contents, which tells you at a glance what is covered in the unit. This is followed by an introduction to the subject matter of the unit and the relationship of the unit to the previous unit. Then follows the objectives in which you are told what you should be able to do by the time you complete the unit. It is advised that you use these objectives to guide your study. After the objectives, you come to the main body of the study unit. The text of the reading is presented in a simple direct style to engage your attention and assist your concentration. You are to go through the unit, section after section. Make sure you fully understand a section and that you have done the self-assessment exercise there before going to the next one. The conclusion that follows the main body of the unit gives you an overview of what you would have achieved in the unit. You should also refer to the objectives of the unit to assure yourself that they have been met. If you are not satisfied that you have achieved all that you were expected to achieve, just go through the unit again. The summary of the unit relates what you have learnt in the unit to the subject matter of the next unit, thus building a "bridge" between the two units. In this way you can see a logical connection between all the units.

You will find this course quite interesting and the study units quite readable. The only problem that you need to worry about is your ability to create a conducive environment for your study. You have to work out your timetable, time and place of study; and demonstrate a serious commitment to your study.

SUMMARY

This course should equip you with basic skills in information storage and retrieval. It is neither abstract nor highly theoretical. Both the course aims and objectives have been set out at the beginning of the guide.

They are all realisable and you should not have any problem realising them. It is hoped that you will find the course interesting and challenging and that you will enjoy reading the course material. We wish you brilliant success.

MAIN COURSE

CONTENTS	PAGE
Module 1	
Unit 1 Data and Information.....	
Unit 2 Document and Documentation Classification	
Unit 3 Classification.....	
Unit 4 Subject Indexing.....	
Unit 5 Indexing Language.....	
Module 2	
Unit 1 Computers in Information Storage and Retrieval	
Unit 2 Storage Media.....	
Unit 3 Records and Files	
Unit 4 Databases.....	
Unit 5 Database Management System	
Module 3	
Unit 1 Concepts of Information Retrieval.....	
Unit 2 Role of Information Centres	
Unit 3 Users and User Needs Information.....	
Unit 4 Searching for Information	
Unit 5 Document-Term Matrix.....	
Module 4	
Unit 1 Retrieval from the Internet.....	
Unit 2 Evaluation of an Information Systems and Services: Part I	
Unit 3 Evaluation of an Information Systems and Services: Part II.....	

MODULE 1

Unit 1	Data and Information
Unit 2	Document and Documentation Classification
Unit 3	Classification Unit
4	Subject Indexing
Unit 5	Indexing Language

UNIT 1 DATA AND INFORMATION**CONTENTS**

1.0	Introduction
2.0	Objectives
3.0	Main Content
3.1	Data
3.1.1	Data Processing
3.1.2	Information
3.2	Value of Information
4.0	Conclusion
5.0	Summary
6.0	Tutor-Marked Assignment
7.0	Reference/Further Reading

1.0 INTRODUCTION

You will find out that many people use the terms "data" and "information" in such a way that you would think they mean the same thing. In this unit, you will learn to use the terms correctly; and also to appreciate the value of data and information.

2.0 OBJECTIVES

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

- explain the meaning of data
- explain the concept of data processing
- explain the meaning of information.

3.0 MAIN CONTENT

3.1 Data

The word "data" is the plural form of the word "datum". Data may be regarded as symbols or figures that have potential value or to which meaning can be given. Now, let us consider how people record events. Think of the old man in the village who made nine strokes of chalk on the lintel of his front door to remind him that it was exactly nine hundred naira that he borrowed. Each time he paid back one hundred naira, he wiped away one stroke, when he was able to pay back a multiple of one hundred naira, he cleaned off the corresponding number of strokes. Would you not consider such a man as well organised? He kept accurate data. He could always tell how much of his debt was outstanding by counting the number of strokes left. Those strokes might not mean anything to someone else but they mean a lot for the old man. Look at the table below and try and make some meaning out of it.

Table 1: Three Ways of Recording Data

	Column 1	Column 2	Column 3
Sunday	::	III	12
Monday	::	III III	3
Tuesday	::	III II	3
Wednesday	:::	III I	3
Thursday	::	III	5
Friday	::	III	6
Saturday	:::	II	7

You may not know the events that were recorded but you can see the frequency of occurrence for each day of the week. Column 1 shows how a little boy recorded the number of cars that came to his father's house during a particular week. Column 2 shows the number of cows a dealer sold in a space of one week. The record was kept by his son. Column 3 illustrates the number of phone calls a young lady received during a particular week.

If you ask all the pupils in a class in a primary school to write their names on a sheet of paper and against their names write their ages, you will have another set of data, namely age distribution of the pupils in the class. The statistical data of the last local government election in 12 wards of a local government is captured in Table 2.

Table 2: Data from a Local Government Election

Ward	Number of Registered Voter	Number of Votes
1	1,567	1,323
2	893	892
3	1,083	995
3	1,002	857
5	1,803	1,635
6	1,337	999
7	1,291	1,295
9	2,002	1,328
10	1,066	993
11	1,153	1,039
12	977	936

Now consider the votes on four motions in a state house of assembly.

The following data were generated.

Table 3: Data from the Votes on Four Motions in a State House of Assembly

	YES	NO	ABSTENTION
Motion 1	18	7	2
Motion 2	11	12	3
Motion 3	13	8	6
Motion 4	9	17	1

SELF-ASSESSMENT EXERCISE

What do you understand by the term "data"?

3.1.1 Data Processing

To put it simply, data processing is what we do to data in order to make some sense out of them. There are many ways of processing data. We may just inspect data and be able to see a pattern in them. From such a pattern we can make a statement about what has happened and even take a decision. This is visual inspection. Let us apply it to the sets of data in Table 1. The data in column 1 show that more cars came to the house on Wednesday. Column 2 not only tells us that more cows were sold on Monday, but that sales declined during the rest of the week. From column 3 we can say that the young lady received the highest number of phone calls on Sunday. We can add that she is more likely to receive more calls during the weekend.

We can build much story on the data in Table 2. In each ward, we can compare the number of registered voters with the number of those that voted. Naturally, not everyone who registered could have, voted. We can compute the voting rate for each ward as the percentage of the registered voters who voted. We can find the ward with the highest voting rate and the one that comes next. We can also rank the wards according to the number of registered voters and on the basis of the ranking predict in which wards to expect the highest number of voters. Now, here is a sticky point: what do you say when the number of voters is higher than the number who registered? A big mistake somewhere or what?

Now let us turn to Table 3. Can you say something about the popularity of the four motions? Obviously, motion 1 was the most popular. Motion 2 must have been highly controversial. Motion 4 was highly unpopular.

Besides visual inspection, rearrangement or sorting, data processing could take the form of arithmetic processes such as computing the sum or mean. Statistical techniques may also be used to show how much the data deviate from some reference value or the relationships within the data. Then we are able to make statements about the data as well as generalise our observations to other similar situations. Suppose we found that out of 420 boys that sat for mathematics in SSSE, 303 passed with credit and above; whereas out 392 girls who sat for it the number who passed with credit and above was 121. Then we can at least say that boys do much better than girls in mathematics in that school. If we collected our data from a number of schools across the country and got the same pattern, then we can say that in Nigeria, boys are better in mathematics than girls provided that the data were generated properly.

It is difficult to estimate the volume of information a person handles everyday. Unless a person is sleeping, his or her brain is always busy processing data and handling information. The brain receives both data and information through the eyes, for instance from what the person reads or sees; through the ears, for instance from what the person hears or listens to; through the nose, for instance from what the person smells, and through the senses of touch, exposure, physical contact and taste.

SELF-ASSESSMENT EXERCISE

Name four ways of processing data.

3.1.2 Information

In the paragraphs above, we considered the processing of data. We found that having done such processing, we were able to make some statements on the situations or events on which the data were obtained. Such statements could be used to guide our future response or action. This is the function of information: to guide a person on what to do, how and when to do it. Consequently, information may be defined as a fact or set of facts that can influence a person's response in a given situation. Ideally, information reduces or even eliminates a person's sense of uncertainty. We could also define information as the outcome of data processing.

Information is a vital resource. It would be impossible to live a normal life without having adequate information. Just imagine waking up one morning to find that you are alone in the house. Not even one other person is around in your house, in your compound, and in the neighbourhood. A thousand questions are racing through your mind, but there is not a soul around to answer you. You decide to move down some distance but still there is not a single person anywhere. What would you do? For how long would you be able to endure that experience?

In order to embark on tertiary education, you needed information to decide what programme to choose. After completing your application for admission, you waited with some anxiety for, information on whether you have been given admission or not. You did not know what else to do until you got that information. A businessman would need up-to-date information on the market situation, and changes in government policies that may affect his business. He would normally consult on a regular basis with his colleagues and share information with them. Today, people are realising more and more the value of information and so a science has developed around it. Information science is the science that deals with the generation, acquisition, organisation, storage, retrieval, dissemination and use of information, its characteristics as well as its impact at the individual, corporate and societal level.

Information plays a key role in every sphere of life. Information is at the core of success of both individuals and corporate bodies in commercial and business enterprises. Information confers competitive advantage on those who have it against their counterparts who do not have it. Information gives power. The countries that have the capacity to acquire or generate and manage information use it to improve their socio-economic status and advance ahead of other nations that do not have such capacity.

Countries in the Third World do not seem to fully appreciate the value of information. In many of these countries, most people in the civil service and in government seem to regard information as what government wants the citizenry to hear, the kind of releases made by the minister or ministry of information, or the news one hears on radio and television or reads in the newspaper. That is information quite alright; but it is "soft" information. The kind of information that confers power in our age, include, scientific, technological, economic and developmental information.

The crucial importance of information has also dictated that both organisations and national governments take appropriate measures to put in place the infrastructure necessary for managing and possibly for controlling it. More and more investment is being made in the establishment of information systems. National, regional and global computer and telecommunication networks have been developed for the management and communication of information. The countries in the Third World have been talking of a new information order. That is a reaction against the domination by the Western and industrialised countries of the global information industry.

3.2 Value of Information

The value of information is enhanced by its accuracy, relevance, timeliness, source, up-to-dateness, as well as the packaging format. If you have any reason to doubt the accuracy of a piece of information, you would not like to act on it. As a matter of fact, inaccurate information could be more disastrous than no information. If it becomes necessary to verify information coming from a particular channel, the cost of such information will be higher, and that may discourage the use of it.

It is quite, important that information be relevant to the purpose for which it is needed. If you would like to read something about the industrial revolution and a library staff gives you a book on the colonisation of Africa, would you be pleased? You may accept the book if you think you might have time to read it, but your need for information on the industrial revolution has not yet been satisfied. You may in fact reject the book. The book is still important and will be useful to someone else but not to you at that point in time.

The other consideration is timeliness. For information to be useful, it must be received in good time to make a difference in what the person who receives it is actually doing. If you wanted the information on industrial revolution in the course of preparing for an examination and

you could not get it until after that examination you would no longer attach much importance to it.

The source of information could be very important. The source could have a high level of credibility so that the person using the information could do so with much confidence. Otherwise it could have some qualified level of credibility. Certainly you would be more assured of the authenticity of information when you hear that it has come from an impeccable source.

It must be noted that information has to be presented in a way that makes it easy to use. A brief summary may be enough and much better for a business executive than a voluminous report. Graphic representation may carry more impression than pages of a statistical report. A video presentation could be better appreciated than a written version on a particular subject. The reverse may be the case in another situation.

SELF-ASSESSMENT EXERCISE

What are the functions of information? Name six factors that determine the value of information.

4.0 CONCLUSION

In this unit you were introduced to the basic concepts of data and information. You should now be able to explain the meaning of data, data processing and information. You can now better appreciate the value of information and the factors that determine the usefulness of information.

5.0 SUMMARY

In this unit, you have learnt what data is and a number of ways of handling data in what is referred to as data processing. You also learnt what information is, the importance of information, and the factors that dictate the usefulness of information. In the next unit, you will learn to distinguish between information and document and you will study the processes of documentation.

6.0 TUTOR-MARKED ASSIGNMENT

Write an essay on "Data and Information". Your write-up should be between six and eight pages of typed A4 double-spaced, 12 points Times Roman. You should include the following:

- i. Definition of data
- ii. Ways of recording
- iii. Data processing
- iv. Definition of information
- v. Importance of information
- vi. Government and corporate roles in information management.

7.0 REFERENCE/FURTHER READING

Susan, A. (1972). *An Introduction to Computers in Information Science*. Metuchen, N.J.: Scarecrow.