EVOLUTION OF CONSISTENCY VALUES WITHIN AN OPTICAL COINCIDENCE SYSTEM: (A Case Study of Agricultural System)

BY

MODU MAINA

(B.L.S.) UNIMAID

A Thesis submitted to the Postgraduate School, Ahmadu Bello University, in partial fulfillment of the requirements for the Award of Degree,

MASTER OF LIBRARY SCIENCE

Department of Library Science,
Ahmadu Bello University,
Zaria - Nigeria.

August, 1986.

DECLARATION

I hereby declare that the work in this thesis submitted for the award of Master of Library Science degree has not been presented either wholly or partially for any degree, and is not being concurrently submitted for any other degree. It is therefore entirely my own work.

SIGNED: _____

MODU MAINA

DEDICATION

This work is dedicated with great affection to my Aunt, Ya Nana and my father Baba Maina Umar, who did for me what no one else could do.

ACKNOWLEDGEMENT

In the course of this research, many people have played one role or the other towards its success.

I would like to thank the entire staff of IAR library for making the facilities at the Institute readily available to me; I hasten to add that I am sincerely indepted to Dr. Ben Oviss Oruma, my principal supervisor for allowing me to tap from his immense wealth of knowledge in Information Science and the good guidance he has given me.

I am also grateful to the Head of Department, Professor Abdullahi Mohammed, Dr. M. Afolabi, Mr. Zia, Dr. D. O. Bozimo my second supervisor, Mr. F. C. Ekere and my course mates (at the top of whom were Adei Akulunye Abu, Sidiku Momoh, & Mohammad Maude Adamu) for showing concern on my progress.

I would like to acknowledge the excellent typing of Mallam Musa Sadiq, Institute of Education Ahmadu Bello University Zaria.

I am most grateful to my wife, Aji for her concern, encouragement and good care of our three children (Suleiman, Zara and Hauwa).

Thanks are not adequate to express my gratitude to Alhaji Ibrahim Guyo Kura, Mr. J. Y. Malgwi, the entire staff of Borno State Library Board and the many people who have enabled me to complete this task.

MODU MAINA.

ABSTRACT

This study is an attempt to show how inter-indexer consistency can be measured to see the evolution of consistency values within an Optical Coincidence System of the Institute of Agricultural Research Library at Ahmadu Bello University. This study is not only to evaluate the present level of performance, but also to find a way of improving the performance of the system.

The study tried to find answers to the following questions:-

(a) Is there indexing Consistency in the System over the period 1974 to 1986?

(b) If there is inconsistency, then identify some of the factors responsible for the inconsistency.

The hypothesis tested is that: Consistency in the system varies with the change of indexers over a period.

The study involved examination of the document cards of the Information Retrieval System in the IAR library. The indexing terms assigned by two groups of indexers to similar documents were compared.

The indexing period in the IAR system was divided into two Periods, P_1 (1974 - 1978) and P_2 (1979-1986). Pairs of similar index entries with one entry falling within either P_1, or P_2 were sorted out and compared for indexing Consistency Over the two periods. The cards were selected in such a way that they covered all the eleven main subjects

areas which the IAR collection covered. A total of 130 pairs pf entries for similar documents were found in the data base. Out of these, 64. pairs were found to satisfy the condition set for the final selection, that is each pair has one entry falling in P_1 and the other in P_2.

The degree of consistency was determined by totalling the number of identical terms assigned to the documents at each indexing, as compared to the total number of terms assigned to the two documents. Hooper's one-indexer consistency model was used in measuring inter-indexer consistency.

The findings showed that indexing consistency in the system is decreasing with the change of indexers. The consistency values ranged from 47.5% to 28.68% per indexer and pair-wise consistency values ranged from as high as 65.4% to as low as 51% per pair.

The study concluded that consistency in the system varies with the change of indexers over a period. This measure of inconsistency can partly be due to differences in subject background and experience of the indexers.

CERTIFICATION

This thesis entitled Evolution of Consistency values within an Optical Coincidence System: A Case Study of Agricultural System, by Modu Maina, meets regulations governing the award of the degree of Master of Library Science (MLS) of Ahmadu Bello University, and is approved for its contribution to knowledge and literary presentation.

Dr. Ben O. Oruma
(Chairman, Supervisory Committee) ---------- Sign and Date

Dr. D. O. Bezimo
(Member, Supervisory Committee) ---------- Sign and Date 23/12/86

Prof. Abdullahi Mohammed
(Head of Department) ---------- Sign and Date 23/12/86

Professor G. E. Osuide
(Dean Postgraduate School) ---------- Sign and Date

TABLE OF CONTENTS

TITLE PAGE	PAGE
Title Page:	i
Declaration:	ii
Dedication:	iii
Acknowledgement:	iv
Abstract:	v
Table of Contents:...	viii
CHAPTER ONE:	
INTRODUCTION:	1
1.1 Purpose of the Research:	1
1.2 Background to the Problem:..	1
1.3 A Statement of the Problem:	4
1.4 Hypothesis:	6
1.4.1 Assumptions:..	6
1.5 Significance of the study:..	6
1.6 Methodology:	7
1.7 Definition of Terms:..	9
1.8 Limitations and Scope of the Study:	11
1.9 Literature Review:	11
CHAPTER TWO:	
PROBLEMS INVOLVED IN AN OPTICAL COINCIDENCE SYSTEM.	
2.1 Background History:...	38
2.2 Present Level of Operation:.	42
2.3 Operational Problems: Hardware (Machine): ...	50
2.4 Operational problems: Software:...	52
2.5 Operational problems: non-machine:	54
2.6 Problems of inconsistency and Consistency Measure:...	56
CHAPTER THREE:	
DATA COLLECTION:	
3.1 Procedure for Data Collection:	67
3.2 The Cut-Off Date:	70

Title Page		Page
3.3	Consistency Values for Data Set:	72
3.4	Summary of Tables:	74
CHAPTER FOUR:		
RESULTS AND ANALYSIS OF RESULTS;		
4.1	Tabular and Graphical Presentation of Results:	77
4.2	Summary of Results:	77
4.3	Analysis of Data:	79
4.4	Speculations:	84
CHAPTER FIVE:		
SUMMARY AND CONCLUSION:		
5.1	Summary:	88
5.2	Inferences from Analysis:	90
5.3	Conclusions:	92
5.4	Recommendations:	92
5.5	Suggestions for Further Studies:	94
	Bibliography:	96
	Appendix I:	102
	Appendix II:	103

CHAPTER ONE

INTRODUCTION

1.1 Purpose of Research:

This study is carried out to improve on the retrieval performance of the optical coincidence system in IAR. In these days of budget cuts and austerity, no research library can afford to invest huge amounts of money on information retrieval system that cannot provide the right information at the right time due to the problem of indexing consistency.

A lot of time is used up in the search for specific information by researchers. The problem entailed in browsing through a heap of literature that has insignificant relation to need, and the low utility content materials cannot be over estimated. Therefore Scientists in research organizations like IAR should be aided with the means of getting quick and right information by using a system that is consistent in its indexing policies and practices.

This study is an attempt to show in simple terms how inter-indexer consistency can be measured to see the evolution of consistency values in indexing over a period of time. It is intended to throw light on the evolution of consistency values within the information retrieval system, of the Institute of Agricultural Research Library, Ahmadu Bello University, Zaria.

The purpose of such a study is not only to evaluate present performance level, but also to find a way of improving the performance of the system.

1.2 Background to the Problem:

The invention of printing press brought about publication (information) explosion and as a result of that handling of

information became extremely difficult.

Before Johann Gutenberg and his associates invented the technique for casting movable types with which books could be printed; books used to be copied laboriously by hand.

> "The printing press, however, did far more than merely relieve the pressure on the scribe. It shattered forever the boundries of a world of learning that hitherto had been the private dormain of the privileged elite. It made possible the rapid mobilization of new learning, first by making available for republication the entire corpus of classical scholarship, and subsequently by opening the doors of publication to contemporary writers. For the first time the written word could communicate pure science and its attendant technologies to an expanding and receptive audience. Science which previously had been the product of individual scholars whose contributions were derived from the age-old learning assimilated during years of apprenticeship, became cumulative, its pace was accelerated, and the scholar found himself compelled to read new books and thus keep up with the work and chought of his contemporaries if he was to advance. The world of scholarship suddenly became an 'open' society and it was doubtless no accident that the century and one-half that followed the invention of printing brought with it Galileo, Bacon, Gilbert, Harvey, the young Newton, and a host of lesser scholars who were influenced by the writings of these masters."[1]

Since then man began to devote serious attention to the bibliographic Organization of this out pouring of print.

In the mid 16th century Conrad Gesner realized that recorded knowledge has grown to such an extent that he set himself to the task of compiling a Universal bibliography of the world literature. In 1545 he compiled 'Bibliotheca Universalis'. Since then many people, like Melvil Dewey, Charles Ammi Cutter, Bliss, Ranganathan, etc have devised systems for bibliographic organization that would keep pace with the accelerated literary growth of the late 19th century.

> "Nowadays, the quantity of new information being generated is such that no individual can hope to keep pace with even a small fraction of it, and the problem that we have to face is that of ensuring that, individuals who need information can obtain it with minimum of cost (both in time and in money) and without being overwhelmed by large amounts of irrelevant matter....."[2]

In trying to find solution to the problems of handling information, people have deviced many ways and means of storing and retrieving information over the years. Among the means deviced are Pre-Coordinate Systems, such as Dewey Decimal Classification, Library of Congress, Universal Decimal Classification, Colon Classification etc. The Post-Coordinate Systems are term entry (Uniterm and Peek-a-boo) and item entry (edge-notched cards). With the invention of third generation computers the problem of information handling was reduced greatly because it enhanced capabilities of processing information. According to Vickery[3], the computer can be regarded as essentially a very fast and accurate clerk, that can read and write text, that can transfer information from one record to another, that can file, and match and sort; count; add, subtract, multiply and divide. Computers would have been the best solution to the problems of information storage and retrieval if not for the cost involved, the level of training required by personnel to run and maintain it. Apart from cost and manpower problems associated with the use of machine as an indexing tool, machine is never known to be the best indexer.[3] Vickery pointed out that, "in so far as human 'understanding' is involved, indexing will remain an art, feasible only to humans, at least until we know a great deal more about 'understanding' than we do today."[4]

There are other disadvantages of machine operated systems; they are less flexible than Manual Systems, less easily modified; they require an initial capital outlay; they cost money to maintain; they may break down, and put the whole system out of action; they can only be worked by specially trained Operators; and they are less flexible than the human brain. Copying, recording,

sorting and transmitting are all admittedly clerical tasks and fit to be mechanized, but indexing and searching, it is argued, require human judgement and should not be mechanized. Because of the resons stated earlier, most retrieval systems employe the services of human indexers; but human beings have problem of inconsistency, inaccuracy and at time human memory can fail. All this lead to inconsistency in indexing.

Studies in indexing and the consistency of indexing have shown that inconsistency is a major problem in manual indexing. A study found out that inconsistencies are due to indexers, rather than to indexing languages.[5] Based on the above reasons this study set out to measure consistency in order to establish continuity in indexing for an optical coincidence system. It will examine progress of consistency values within an optical coincidence system of the Institute of Agricultural Research Library, at Ahmadu Bello University, Zaria. In carrying out the research, index entries for similar documents indexed at different times by various indexers will be examined for inter-indexer consistency. The index entries to be examined will be right from 1974 to 1986 (i.e. from the date of inception of the system to date).

Although there are many studies on Optical coincidence systems (OCS) in libraries, there are no known studies that have delt with this particular area the study is trying to look into.

1.3 A Statement of the Problem.

Special libraries (i.e. research libraries) are frequently faced with the problem of quick (specialized) information retrieval. The

Problem doesn't end there, some researchers want all that the system has in their area of interest or ask for what is the latest in their area. Such a request and a host of them can best be handled by mechanized information retrieval systems.

The current trend in information retrieval is moving towards mechanized information retrieval systems, but mechanization means money, human resources and other facilities. These days of austerity, not many institutions have the resources to install and run computers. More so in a developing country like Nigeria where money is lacking to provide for the basic necessities of life, like food, cloth and shelter. Spare parts for many simple machines are lacking and computer a sophisticated machine cannot be an exception.

It has been pointed out earlier on, that the machine is not the best indexer so even the mechanized retrieval systems depend on human indexers for input.[3] The study seeks to discover how consistent or otherwise the human indexer has been in his indexing for the IAR system.

In short the study is trying to find answers to problems such as:-

(a) Is there indexing consistency in the system over the period 1974 to 1986; This is a way of measuring consistency to establish continuity.

(b) If there is inconsistency then identify some of the factors responsible for the inconsistency.

Hypothesis:

Consistency in the system varies with the change of indexers over a period.

1.4.1 Assumptions:

It is assumed that the following variables have no influence on consistency level in the system. The variables are Document type/form; change in research direction, language and revision of Thesaurus.

Document type/form is assumed not to affect consistency. According to Oruma[5(b)] document type/form does not affect consistency with which the document will be indexed. It could be a book, periodical article or newspaper cutting; all are considered not relevant to consistency level.

Change in research direction e.g. from Animal Science to Crop Science does not affect consistency and the samething applies to change in research language. Revision of Thesaurus is also assumed not to influence consistency in indexing.

It is also assumed that there is not necessarily a direct correlation between the number of terms applied and the quality of the indexing.

1.5 Significance of the Study:

The study will enable the IAR system operators to know the trend in its indexing. It is also aimed at providing guidelines for the study of the evolution of indexing consistency values within an optical coincidence system in Nigerian Special libraries. Another significance of the study lies in the fact that it is a pioneer in the field of OCS study in Nigerian Special Libraries.

1.6 Methodology:

The indexing described in this study is based upon an optical coincidence System i.e. OCS. Headings are chosen from the IAR thesaurus.

The study involves examining the Optical Coincidence System in the Institute of Agricultural Research Library which is currently in operation. Personal visits and observation of the library at work will no doubt afford assessment of how effective the OCS is, in retrieval of information by both researchers and staff members.

A close examination of the basic principles by which Optical Coincidence Systems operate will also reveal alot.

The study attempts to compare the index terms assigned by two groups of indexers to similar documents.

The indexing period in IAR System was divided into two periods, that is P_1 (1974 - 1978) and P_2 (1979 - 1986). This is done to have two distinct groups of indexers that fall within those periods without overlap. There is overlap in the period before 1978 and that after 1979.

Pairs of similar index entries or document cards each falling within either P_1 or P_2 were sorted out and compared for indexing consistency over the two periods. Everything available for comparison was selected in all the eleven main subject areas which the IAR collection covered. The eleven main subject areas covered are, Agriculture, Agronomy, Agric Economics, Agric Engineering, Animal Science, Crop Science, Environmental Sciences, Extension Services, Forestry, Plant Sciences and Sociology.

To start with 130 pairs of cards with entries for simila or exactly a like documents were selected from the data base and arranged according to the main subject areas covered by the IAR Information Retrieval System. After close examination

66 pairs were disqualified as possible entries because they failed to meet the exact criterion set for the study. At the end of the day 64 pairs survived the Weeding test and those 64 were subjected to inter-indexer consistency tes

Degree of consistency is determined by totalling the number of identical terms assigned to the documents at each indexing, as compared to the total number of terms assigned to the two documents.

In trying to measure consistency, various consistency models were considered. Among the models considered are Hooper's one-indexer model, Korotkin & Oliver Model, King & Bryant's Model, Rodger's Model, Slamecka & Jacoby's model, and Zadeh's Fuzzy set theory. Among all Hooper's model seems to be most suitable for application in this particular case because it can accommodate any number of indexers without problem. The calculation will not be cumbersome as it will be when other models are used.

Hooper's[6] consistency measure of one indexer with respect to a group of any number of indexers, N is given as:

$$CG_i = \frac{CP_{ij} + CP_{ik} + \ldots CP_{iN}}{N - 1} \quad \ldots (1)$$

where CG_{iN} = pair - wise consistency measure between indexers I_i and I_N where N is the total number of indexers in the group.

Therefore inter-indexer consistency was calculated using equation 1 above.

1.7 Definition of Terms:

The following terms are defined as follows to make for easy understanding and clarity of what is being discussed in this work.

- IAR - Institute of Agricultural Research, Ahmadu Bello University, Zaria.
- Information - Allen Kent[7] defined information as, "any recorded Knowledge that may be useful to some decision maker." All activities require decision and information is needed to support decision.
- Information Retrieval System - (IR) it means information storage and retrieval system. They can rightly be called document retrieval system since they retrieve references to documents containing the information.
- Mechanized Information Retrieval - according to Kent it is the use of Mechanized or non-Convensional tool in connection with any one or more of the following i.e. analysis, vocabulary control, storage of records, conducting of search, Question analysis, delivery of search results etc.
- OCS - Optical Coincidence System. It is basically a matter of alligning cards to see which positions let light through and such positions will give the numbers of the appropriate documents.
- Optical Coincidence Cards - are term cards which are known as peek-a-boo, battern cards, feature cards, peep-hole cards etc. They are the cards on which document numbers are punched.

- Document Cards - are cards with document numbers and bibliographical details of document. They are arranged either according to alphabetical author arrangement or according to document numbers.
- Retrieval - it is the act of recovering stored information when needed.
- Consistency - may be defined as the assignment of the same term or terms to a document each time it is indexed.[57]
- Inconsistency - includes differences in vocabulary usage as well as differences in the choice of subject content to be indexed (i.e. depth of indexing).[57]
- Inter-Indexer Consistency - the consistency with which two or more indexers tend to choose the same terms as descriptive of the same documents.
- Intra-Indexer Consistency - the frequency with which one indexer at different times will, for any given document, assign either the same or very similar sets of term. That is how consistent is the indexing done by one person on a particular document at different times.
- Hardware - e.g. reading equipment, drilling machine, typewriter, storage cabinets etc.
- Software - feature cards, thesaurus, document cards, indexing policy, query forms, etc.
- Input - recording information on feature cards by way of punching holes corresponding to document numbers.
- MEDLARS - Medical literature Analysis and Research Services.
- Indexing - it is an intellectual evaluation of the subject content of a given document with the aim of determining

the Key-words or terms which best express the subject matter of the published work.

- Thesaurus - it is "an authority file which can lead the user from one concept to another via various heuristic or intuitive paths. It may be manually operated or mechanized for assignment of index headings."[7.b] It can be described as an authority file which contains cross-references and can thus lead the user from one concept not in use to prefered terms.

1.8 Limitation and Scope of Study:

The study is limited to the information retrieval system in IAR Library which uses an optical coincidence system.

It is also limited to the number of entries for similar or exactly alike documents made within the two periods P_1 and P_2. It covered only those index entries input into the system.

This study does not attempt to evaluate quality of indexing, but merely consistency.

1.9 Literature Review:

Konrad Gesner - the father of modern bibliography compiled a bibliography of world writings in Latin, Greek and Hebrew in the 16th Century. Late in the 19th century Paul Otlet and Henri La Fontaine Started to compile world bibliography of recorded knowledge. All these were done in the name of information storage and retrieval.

According to Shera[8], when Otlet and La Fontaine started work on their great world bibliography of all recorded knowledge in the late 19th Century, they tried to make it comprehensive and attempted more penetrating subject analysis of bibliographi

materials than had been used in conventional Librarianship. In order to differentiate their activity from that of the librarians they gave it the name documentation. That was the begining of Schism in librarianship. Kondokov[9], pointed out that the desire to improve information Services to Science a industry and bring these up to the level of modern standards led to documentation.

Documentation came into being as a result of the need for improved bibliographic access to the records of scholarship. To support this point Shera[10], showed that Ancient, Medieval and even relatively modern libraries were vast storehouses of accumulated records with the most Crude and meager of bibliographic devices to reveal the Wealth of their resources.

The American Documentation Institute (ADI) was formed in 1937 and later during the second world,War, work on Documentation was suspended. A New era started after after the war. ADI was revived at the close of 1940's and brought in a number of people with interests other than microphotography; then documentation took a new dimension. Finally the influence of the microphotographers in the institute (ADI) Faded and they formed the National Microfilm Association which was composed largely of manufacturers and distributors of microfilm equipment and supplies.

According to Shera and Cleveland[11], from 1952 to the end of the decade, the institute and its activities became a battle ground for competing philosophies, methods and Techniques for the subject analysis of recorded knowledge. Calvin Mooer was inventing Zatocoding; Taube was pushing Uniterm and Coordinate indexing and eventually founded Documentation Incorporated. Though the fortunes of the ADI were faltering, in 1957 interest

in improving access to recorded knowledge particularly Scientific Knowledge was on the increase. Also Centre of activity was moving to governmental agencies.

Bradford[12], defined Documentation as, "the art of collecting, classifying and making readily accessible the records of all kinds of intellectual activity, ... the process by which.... is ... put before the creative specialist the existing literature, bearing on the subject of his investigation, in order that he may b made fully aware of previous achievements in his subject, and thus be saved from the dissipation of his genius upon work already done." Shera and Egan[13], also pointed out that a bread definition of 'Documentation' would include all those processes involved in the assembling of collections of written records and in providing guides to their use.

It was reported by Shera[14], that an American Librarian whose name has been forgotten was so hypersensitive to his professional status, that he proclaimed Documentation to be "Librarianship performed by amateurs." According to Mohrhardt[15], although both Documentation and Librarianship stand on the common basis of providing effective access to all types of information, their methods of operation and in some cases the varieties of information they handle have singular differences. He noted that it is discouraging to the leaders in both groups (Lib arianship and Documentation) to find that at times more emphasis is given to differencies rather than similarities and common bases.

Mohrhardt thought that it is fair to report that Documentalists and Librarians in the United States View each others

work with respect and also with an appreciation for the commonality of their responsibilities and objectives. In fact it is much easier to identify and describe points of similarity in these two fields than it is to find true points of difference either in theory or in practice. To support his arguement he pointed out that American Library Association in its 90th year had a broad base for membership which included all types of librarians, all ranges of Documentalists, Science Information Specialists, Research Workers, administrators and friends of libraries. This shows commonality.

Under the leadership of Dr. H. Borko, the ADI took ini iatives in identifying common elements of interest and areas for cooperation in various library and documentation groups. It is interesting to note that the President-elect of ADI in 1966 was a professional librarian, Mr. Bernard Fry. There is such a complete inter-relationship or overlap of the various library and Documentation activities throughout the United States, that you often find one individual a very active member in more than one of these groups. In addition to the inter-disciplinary character of the professional associations within the U.S., there is also an overlap of responsibilities within the research and special libraries themselves. Hence a library may be classified simultaneously as an information Centre or a documentation Centre.

Mohrhardt added that Documentalists are not happy to be included as a unit of librarianship, and librarians object strongly to being considered as a branch of Documentation. The author personally views Documentation as a development or extension of Librarianship because of his expirence in Librar nship.

According to him many who have not been associated with professional Librarianship tend to think of documentation as a discrete discipline, service or profession, which is not true since nearly all formal Academic training in documentation originated from library schools. He acknowledged the fact that a true documentation service certainly does go beyond convensional library activity both in scope and in depth of individual service to users, and added that the basic elements of both disciplines are similar if not identical.

He concluded that a joint meeting of officers of International Federation of Library Association and the International Federation for Documentation is a Landmark of Cooperation among the two.

Vickery[16], writing on the 'Concepts of Documentation' said documentation is very much wider than convensional librarianship because it includes publication and printing, distribution, some form of telecommunication, analysis, storage, retrieval and delivery to the user. Ranganathan[17], gave a model of documentation system. "The model replaces 'books' with the wider term 'documents' and 'readers' with 'recipients'. Instead of 'library' and 'staff' we find 'channel' and 'channel Agents'. The model further reminded us that there are Sources generating messages as well as recipients using them. Lastly the model is cyclic and emphasizes that documentary communication is a continuous flow process."

When the split in librarianship was widening many non-librarians (Scientists in particular) who are contemptious of librarianship invaded the field. At the time when documentation

was dying the emphasis moved to information retrieval and that was the begining of Information Science. Information Science and Documentation Co-existed for sometime before Information Science was accepted as a discipline. Many Documentalists and Librarians kicked against Information Science that it has nothing to offer, it is only Scientists, engineers and Mathematicians wanting to introduce mathematics to librarianship for prestige. They said, it is those who have interest in mathematics that want to bring their influence into librarianship for want of name.

Shera and Cleveland[18], Said attempts to persuade ADI and special libraries Association to form an American Counterpart of the British Aslib failed due to fear of loss of autonomy by both parties. In 1966 ALA replaced it earlier Interdivisional Committee on Documentation with the new Information Science and Automation Division, thus the begining of Information Science. Two years later, i.e. in 1968 Eugene Garfield and others succeeded in changing the name of ADI to American Society for Information Science (ASIS).

The Information Science Era. It is believed that Documentation movements in Europe and America led to the development of Information Science from 1950's to 1960's and it is believed to have emerged from documentation and Information retrieval and many other areas.

Shultz[19], Said that 1960 provided a climate that was exactly right for the development of information Science when interest in the problem of information on the part of the Scientific world and the American Federal Government reached an unprecedented level. Coupled with the accelerated technological development of that period, Information became a

focal point in space exploration, Science education etc, to regain Lost prestige. Transfer of information is an inseparable part of research and development.

The 1960's decade was a period of incredible technologic. advancement including of course the refinement of devices and systems in the field of communication and information processing itself. The third generation computers were introduced, utilizing the miniaturized components developed for use in space program. Massive data storage capabilities were now available to provide reliable high-speed access to billions of Units of stored information. By this time computers were considered as information processing machines rather than merely 'number crunchers' for mathematical application. This developments made it possible for information science to develop. According to Landau[20], without technology a significant proportion of information science activities would not exist.

Borko[21], Said that the existing institutions and proceedures dedicated to the accumulation and transmission of knowledge became inadequate due to the tremendeous growth in Science & technology, fast rate of obsolescence of technical knowledge, the large number of working Scientists, the increase in specialization and the short time lag between research and application. As a result of these pressures it became necessary to study information science.

A variety of definitions have been provided for this new field of information science. It has been described as the science which deals with the generation and transformation of information. The following definitions emerged from the Georgia Institute of Technology Conference held in October,

1961 and April, 1962, "The Sciences that investigates the properties and behaviour of information, the forces governing the flow of information, and the means of processing information for optimum accessibility and usability. The processes include the origination, dissemination, collection, organization, storage, retrieval, interpretation and use of information. The field is derived from or related to mathematics, logic, linguistics, psychology, computer technology, operations research, the graphic arts, communications, library science, management and some other fields."[22]

Information Science has also been defined interms of the problems it seeks to solve, or phenomina it investigates. Saracevic[23], pointed out that the basic problems addressed in information Sciences is the effectiveness in communication of public knowledge, by means of the 'literature'. He described information science in relation to library Science as trying to develop a theoretical base and a Scientific component with a tradition of theoretical/experimental inquires. According to Shera, "Librarianship is the generic term and information science is an area of research which draws its substance, methods and techniques from a variety of disciplines to achieve understanding of the properties, behaviour and flow of information. Information Science contributes to the theoretical and intellectual base for the librarians Operations."[23] Jean Tague[25], described the relationship between library Science and information Science as overlapping but not co-extensive disciplines. The former is concerned with all aspects of library Operation and practice; the latter with the characteristics of public information and the behaviour of its users. The boundry between the two is difficult to fix.

Taylor[26], was of the opinion that there are two misconceptions about information science. One of them is that librarianship and related professions tend to view information sciences as concerned solely with their information - processing activities. Information science is not concerned with the development of information handling systems, but rather with the explication of system and environment. A second misconception is that information science and engineering is generally viewed, in its relation to libraries, as directed only towards retrieval or management information systems, and serving a rather small segment of interests. He concluded that it is important to remember that the library is but one application of information science.

Goffman, Bracken, Belzer, Wersig and Neweling, all defined Information Science in various different ways. Some defined it by listing topics that fall under it.

Afolabi defined information science as "an interdisciplinary research-based subject field, that is, concerned with the theories of generation/production, transmission, storage, retrieval, dissemination and management of information, and draws its concepts and theories from a wide Spectrum of field such as computer science, library science, management, economics, semiotics and mathematics. It is related to library Science in the sense that it provides the theories on all aspects of information for the information work of libraries and information centres."[27]

Saracevic[28], Said that information science grow out of the concerns with three levels of (interlocke) problems, which are (a) communication problems (b) literature problems

and (c) library (or information system) problems - which gave rise to information retrieval and information networks.

Information Science emerged and grew after the second World War. Its growth coincided with developments in information technologies e.g. computing, telecommunication and reprographic technologies. Information Science is still at its developmental stage in most countries of the world.

In the 1950's Information Retrieval was born and the result was that a number of information systems were established with many claims and counter claims. The use of the term information science came about in the early 1960's when it was felt that there was a need for a Scientific dimension to all information activities, particularly information retrieval. According to Afolabi[29], it was during the early 1960's that information retrieval systems came into existance; a number of large national systems were established (some government -based, some society-based, some commercial based) while those in companies and institutions were on the increase. The following innovations also emerged in the practical aspect: new services (e.g. selective Dissemination of Information), new tools (e.g. KWIC indexes, citation indexes, thesaurus), new computer file organizations (e.g. Combined files, lists and chain files) and Work on library automation also began.

In 1958 Keppler[30], hypothesized that a bibliographic searching machine using Hollieth equipment and punch card for subject and author indexing that would replace the conventional library catalogue and other traditional bibliographic publications will appear.

Meadous[31], pointed out that since early sixties, systems for the mechanical retrieval of information have greatly expanded. This period has seen for example the implementation of SDI as a routine activity. The role of computers in the production of various types of index has for example, been highly significant. New sorts of literature guide, such as the science citation index (SCI) could not have been produced in the absence of computers.

Mechanization of information retrieval systems started with the post-coordinate indexes such as Uniterm, edge-notched and peek-a-boo, then moved to punch cards and finally to computers.OF the semi-mechanized systems (Uniterm, edge-notched, and peek-a-boo); Peek-a-boo or Optical coincidence system is more advantageous in information retrieval. Johoda, arguing in favour of optical coincidence system said that, "there are no mechanical restrictions on the total number of index units in the indexing vocabulary or on the number of index units selected for a document in the peek-a-boo system, such a limitation occur with edge-notched cards and machine sorted punch cards systems since one is restricted to capacity of the cards and sometimes has to resort to codes that are difficult to encode and/or search."[32]

According to Costello[33], both data and information can be handled by optical coincidence system. He went on to say that retrieval display in OCS is instantaneous and Coordinatio are not limited to two terms at a time as they are in manual cards systems. The potential for false retrieval is small, even with storage of 50,000 capacity cards. It might be feasible to store Microimages of documents, and small original

documents or condensations of larger original source documents Storage and retrieval devices, such as Optical coincidence card have been especially designed for information storage and retrieval operations. Foskett[34], pointed out that an optical coincidence system (OCS) can be used for both small personal collections and large information centres depending on the size of the card. He said there is no limit to the number of descriptors that may be used and searching the file is simple no matter how large it is. The use of natural language and ease of searching have led to a tendency to adopt peek-a-boo rather than edge-notched cards in libraries and information centres. He argued that convensional card file requires a card for each method of approach; one or more for the author/s and one for each subject in order to conduct a successful search. There is also the problem of choosing a suitable system of subject headings or classification to arrange the subject cards. To avoid these difficulties many people turned to more recent forms of index for example Optical Coincidence System. Optical Coincidence Systems are cheap and require cheap materials to operate with. Lancaster[35], pointed out that post-coordinate retrieval systems (OCS inclusive) permit multi-dimensional access to any subject from any point of view economically. So far it has been seen that OCS are suitable and easy to use as information retrieval systems compared to other systems.

After information retrieval systems have been deviced, the problem that came with it was the problem of evaluation. The Systems have to be evaluated to see how far they have

achieved the purpose they are set to achieve or how do they compare with other systems in their performance.

According to Fairtherne[36], the requirement of all retrieval systems is to aid document retrieval. He pointed out that, if at the time of input the document description is fine the retrieval also will be fine and less irrelevant materials will be retrieved; and if document description is bad, retrieval results will also be bad. Cleverdon[37], considered two variables which he termed "recall ratio" and "relevance ratio" (precision ratio) in system evaluation. He emphasized that there is inverse relationship between the two whereby, when recall ratio increases, relevance ratio decreases and vice versa. Therefore highly restrictive queries would lead to a relatively high relevance ratio with a relatively low recall ratio; a less restrictive querry would enhance the recall ratio at the expense of the relevance ratio. Swets[38], pointed out that during the period of intensive study of information retrieval systems, ten different measures for evaluating the performance of such systems have been suggested. He was of the opinion that the various measures have much in common, that is, eight of them evaluate, only the effectiveness (accuracy, sensitivity, discrimination) of a retrieval system and are derived completely, in one way or another, from a 2-by-2 contingency table of pertinence and relevance. The other measures assess efficiency as well as effectiveness, by including such performance factors as time, convinence, Operating Cost and product form.

Cleverdon[39], in the cranfield project 1, attempted to compare the performance of four index languages: UDC, alphabetical subject catalogue, Uniterms and special faceted

classification. The study indicated that there is no significant differences in the efficiency of retrieval of the four indexing schemes. They gave results ranging from 60 to 80 percent retrieval with 25 percent relevance under test conditions. It was noted that Human error in indexing and searching were more serious than errors due to file organization. It was concluded that file organization is relatively unimportant in the performance of information retrieval systems. The specificity of the vocabulary and the exhaustivity of the indexing are much more important factors affecting performance.

Swets[40], said a desirable measure of retrieval performance is that it would express soley the ability of a retrieval system to distinguish between wanted and unwanted items. He argued that it is difficult to compare two systems e.g A and B and say which system performs better. He said the use of emprical recall-precision curve as a measure of system effectiveness provides partial solution in that it does not tell by how much one system is better than the other. He concluded by saying that the present retrieval methods leave considerable room for improvement. Van Rijsbergen and Jones[41], said a basic assumption in retrieval systems is that the document relevant to a question are separated from those which are not relevant. They said, "if initial separation between relevant and non-relevant documents is poor, our chances of being able to widen the gap materially are low."

Salton and Young[42], studied the existing practice in automatic indexing and showed that the standard theories for the specification of term values (or weights) are not adequate.

Vickery[43], defined the effectiveness of a system to mean the degree to which it achieves its stated objectives. If the objectives can be related to a quantitative characteristic e.g. recalling all relevant documents - then effectiveness can be measured as the ratio of the two numbers. That is number of relevant items retrieved in search, to the number of relevant items in system. He added that people evaluate because they are not getting the desired results or when they want to compare one situation with another one or more others. Most systems perform more than one function, and the practice of measuring a single characteristic leads to considerable danger of sub-optimisation. At the end the effectiveness of a system according to one characteristic may be maximized, but its performance in relation to other characteristics (of equal or greater relevance to the system use) may be worsened.

Lancaster[44], reporting on MEDLARS said it is a multipurpose system with a prime purpose of producing index medicus and other recurring bibliographies. The system was found to be operating on the average, at 57.7% recall and 50.4% precision. Causes of system failure were found to be due to indexing or searching subsystems; or a combination of the two subsystems. Index language leads to system failure, likewise user-system interaction contributes to failure. Lancaster and Jenkins[45], showed that the evaluation of MEDLARS (1) has demonstrated that it is possible to undertake an investigation to measure the effectiveness of a large mechanized retrieval system and to identify the factors that are most important in controlling and limiting the effectiveness of the system. They noted that a one-time evaluation, however can only measure the performance of a system at a particular point in time. When changes are

made to the system, as a result of the investigation, we would like to be able to estimate the effects of these changes. Moreover a one - time evaluation, however comprehensive, cannot identify all sources of system failure. It can locate the most important sources and areas of greatest weakness but it cannot identify, for example all specific instances of vocabulary inadequacy in the system.

The effectiveness of any information and/or document retrieval system is dependent upon several factors: the documents or information included in the Original input, the manner in which this is indexed, the limitations of the program and machines involved, the search strategy, and the degree to which the system actually meets the needs of its users. The indexing factor can further be divided into the indexing system itself and the manner in which that system is applied to a body of information and/or group of documents by the indexers. All the above factors must be studied before any total evaluation of a given system can be made. This study, however, will be limited to only one aspect of such an evaluation: the consistency with which different indexers apply one given indexing system to a collection.

According to Fairthorne[46], indexing is the basic problem as well as the Costliest bottleneck of information retrieval. Laurent[47], showed that indexing is a frequent source of search failure; and the compatibility of term selection between indexers and Searchers is important. Zull and Baumanis[48], pointed out that lack of agreement in interpretation of terms between indexers and searchers is a source of system failure. Lancaster[49], reported that some wide variations in assigning

certain index terms were noted during the evaluation of MEDLARS. Lufkin[50], investigated aspects of subject indexing and found out that consistency in indexing increased with increasing length of the document.

King[51], discussing Manual indexes said that Quality human indexing is a tedious, time consuming, and difficult task, requiring the use of highly skilled personnel. Often, a set of rules or an authority list has been provided to an indexer for use in selecting the index terms, but, as is well known, two indexers using these guidelines are quite likely to select different index terms to represent a given document. There is also a high likelyhood of one indexer selecting different index terms to represent the same document at different times. Many attempts have been made over the years, with only limited success, to sufficiently refine the indexing criteria so that a unique set of index terms will be selected, independent of indexer, to represent a document.

Tinker[52], argued that consistency of indexing measures the precision with which meaning is understood by the indexers. His study suggests that a drastic reduction in the number of allowed indexing terms would increase consistency with which the terms would be used (applied). It also suggests that differences of opinion as to the meaning or definition of a term leads to inconsistency.

Tarr and Borko[53], discovered that two groups of indexers proven statistically to be highly consistent, may nevertheless be quite inconsistent in their indexing of one particular document. Thus consistency must be considered to be

independent of any particular document. They also reported that the use of a thesaurus improves consistency only when the indexing process is separated into two independent but interrelated steps; these are 1st, formulating verbal labels for concepts to be indexed, and 2nd using a thesaurus to convert the verbal labels into controlled indexing terminology. The performance of these two steps by separate people may hold the promise not only to improving consistency, but also of lowering cost since the second step could be done by less highly trainned people. Tinker[54], proved that a given term is applied consistently by a group of indexers is no guarantee that the group will apply some other terms with equal consistency. He sited a group of indexers who were consistent in the application of the term "electrophotographic plates", and not consistent in the application of term "conductivity." Rayward and Svenonius[55], reporting on interindexer consistency at various indexing depth, that indexers agree on surprisingly small number of index terms. This led the writers to believe that the low degree of consistency is due to the indexers' difference of opinion as to what length of phrase should be in order that it be selected as an index term.

According to Loukopoulos[56], human inconsistency in distinguishing among a given set of eligible terms under which a document may be indexed is inevitable. Huruitz[57], pointed out that it does not seem possible for a system to retrieve documents effectively if they are not indexed consistently. Once this point is recognized, the problem becomes not to eliminate inconsistencies, for this is impossible, but to discover what types of inconsistencies

occur and the best ways to minimize the effects of these on document retrieval.

Savage[58], suggested that one need not expect that two indexers (or one indexer at different times) will, for any given document, assign either the same or very similar sets of terms. What one should expect is that for any homogenous collection of documents any two indexers (or one, etc) will assign terms with roughly the same relative frequencies (or because of differences in depth, a high correlation between rank orders of relative frequencies) and that a document set retrieved on the basis of one indexing will be equally satisfactory to that set retrieved by the other.

Zunde and Dexter[59], said that measures of indexing consistency should reflect not only the formal agreement of indexers on a number of terms, but also the significance of terms on which the indexers agree or disagree.

According to Cooper[60], it is often assumed that inter-indexer consistency is somehow indicative of the qualify of the indexing; but his study led him to conlude that until equations describing this relationship has been derived, measurements of interindexer consistency will have little meaning as clues to indexing quality. Leonard[61], in his doctoral thesis concluded that inter-indexer consistency is mainly dependent on three factors - indexers experience, term type and document length. The study also confirmed that there is a high correlation between interindexer consistency and high retrieval effectiveness, implying that efforts to maintain high levels of consistency are probably quite justifiable in term of assuring good retrieval performance.

According to Slamecka and Jacoby[62], interindexer consistency improves significantly with the use of prescriptive indexing aids which contain a minimal display of the variable semantic relationships among terms. This conclusion imply tha greater consistency in coordinate indexing, and hence an improvement in the quality of indexing, lie in the direction of controlled indexing vocabularies which formalize the relationships of terms so as to uniformly and invariably prescribe the choice of indexing terms.

Hooper[63], cited consistency value ranging from 10% to 80% depending on the condition under which the indexing was performed and the measure of consistency used. Jacoby[64], reported a mean consistency of 10% when chemical patents were indexed by three experienced and three inexperienced indexers and no indexing aids were used. Consistency values of 35% to 45% were obtained by Slameck and Jacoby[65], for experienced indexers using certain indexing aid such as controlled vocabulary. In their report Jacoby and Slamecka[66], arrived at a consistency of 16.3% for experienced indexers and 12.6% for inexperienced indexers. No indexing aids other than indexing rules were used.

Painter[67], reported consistency values of 40%, 42% 48% and 70% with varying indexing systems and types of documents. Rodgers[68], came up with an average consistency of 24% for combination of two indexers in a group. Korotkin and Oliver[69] reported consistencies ranging from 36% to 59% in an experiment in which five psychologists and five nonpsychologists indexed abstracts.

Consistencies up to 80% were reached when indexers were required to use classification schedules and thesauri and had very limited freedom or no freedom at all to use terms not contained in the above indexing aids (67, 69, 70, 71, 72).

It can be seen from the reviews so far, that any two indexers, indexing one and the same document individually, will select sets of indexing terms which are most unlikely to be identical. In other words, if one compares indexing terms assigned by any two indexers to the same document, one discovers that, as a rule, the indexers differ considerably in their judgement as to which terms reflect the contents of the document most adequately. It is clear that this difference in the judgement of indexers introduces a great deal of uncertainty in any information retrieval system based on human indexing. Although an information system normally would not contain identical documents indexed by different indexers, the implication is that documents which are similar may and often would be indexed so differently that their similarity would not be properly reflected in the sets of indexing terms assigned to those documents. Therefore it is important to measure consistency in order to find ways of minimizing inconsistencies in our indexing systems.

According to Oruma and Saracevic, studies in indexing and the consistency of indexing have shown that inconsistency is a major problem in manual indexing. According to them much of the inconsistencies associated with manual indexing may be attributed to differences in human assessment.

References:

1. Shera, Jesse H. (1963). The Propaedenticeof the New Librarianship. In: Simonton, W. (Ed.) Informatior Retrieval Today. University of Minnesota, Minneapolis, pp. 5 - 20.

2. Foskett, A. C. (1977). The Subject Approach to Information, 3rd ed. Clive Bingley, London, p. 1.

3. Vickery, B. C. (1968). On Retrieval Systems Theory, 2nd ed. Buttermorths, London, 191p.

4. Vickery, B . C. (1968) ibid.

5. Case Western Reserve University Centre for Documentation and Communication Research (1968). An inquiry into testing of information Retrieval Systems Part I: In: Caudra C. A. (Ed.) Annual Review of Information Science and Technology - Encyclopedia Britanica Inc., Chicago, 4:pp. 81-82.

6. Hooper, R. S. (1965). Indexer Consistency Tests - Origin, Measurement, Results and Utilization. In Oruma and Saracevic (n.d) Consistency Studies: Eliminating the Disparities Resulting from Consistency Models and Definitions. pp.1 - 11.

7a. Kent, Allen, (1966). Textbook on Mechanized Information Retrieval, 2nd ed. Interscience Publishers, New York, 371p.

7b. Homerton, Paul W. (1965). "Organic and Functional Concepts Authority Files." In: Simon M. Nemman (Ed.) Information Systems Compatibility. Spartan Books, Washington.

8. Shera, Jesse H. (1968). "Of Librarianship, Documentation and Information Science." Unesco Bulletin for Libraries, 22(2): pp. 58 - 65.

9. Kondaker, I. P. (1966). "The Library and Documentation." Libri, 16(3): pp. 216 - 220.

10. Shera, Jesse H. (1951). "Documentation: Its scope and Limitations." Library Quarterly, 21(1): pp.13 - 26.

11. Shera, Jesse H. and Cleveland, D. B. (1977). History and Foundations of Information Science. In Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, - Encyclopedia Britanica Inc, Chicago, 12: pp. 249 - 275.

12. Bradford, S. C. (1948). Documentation. Crasby Lockwood, London, p. 11.

13. Shera, Jesse H. and Egan, M. E. (1950). "Documentation in the United States." American Documentation, 1(1): pp. 8 - 12.

14. Shera, Jesse H. (1951). Op.cit. pp.13 - 26.

15. Mohrhardt, Foster E. (1966). "Librarianship and Documentations: Relationships in the United States." Libri, 16(3): pp.211 - 215.

16. Vickery, B. C. (1978). "Concepts of Documentation." Journal of Documentation, 34(4): pp.279 - 287.

17. Ranganathan, S. R. (1957). Five Laws of Library Science. Blunk, London.

18. Shera, Jesse H. and Cleveland, D. B. (1977). Op.cit.

19. Schultz, Claire K. (1976). "ASIS notes on its Founding and Development." Bulletin of the American Society for Information Science, 2(8): pp.49 - 5 .

20. Landau, Robert M. (1974). Impact of Information Technology. In Anthony, Debons (Ed.) Information Science: Search for identity: Proceedings of the NATO Advanced Study Institute in Information Science; 1972 August 12 - 20; Champion, P. A. Marcel Dekker Inc., New York, pp. 191 - 196.

21. Berko, H. (1968). "Information Science: What is it?" American Documentation; 19(1): pp.3 - 5.

22. Conference on Training Information Science Specialists, Oct.12 - 13, 1961 and April 12 - 13, 1962. Proceedings. Institute of Technology, Atlanta, Georgia.

23. Saracevic, T. (1979). "An essay on the past and future of Information Science Education:1. Historical Overview." Information Processing and Management, 15: pp. 1 - 15.

24. Shera, Jesse H. (1972). Foundation of Education for Librarianship. Becker and Hayes, New York, p.293.

25. Tague, Jean (1979). "Information Science in graduate Library programs." Canadian Library Journal, 36(3); pp.89 - 99.

26. Tayler, Robert S. (1967). "Interface Between Librarianshi and Information Science and Engineering." Special Libraries, 58: pp.45 - 48.

27. Afolabi, Michael (1985). Information Science: Its definitions, evoluation and Place in Nigerian Departments of Library Science. Paper read at the Department of Library Science Seminar, Jan. 23.

28. Saracevic, T. (1979). Op.Cit.

29. Afolabi, Michael (1985), Op.cit.

30. Kapple, Frederic J. (1939). Looking Foward - A Fantasy. In Donton, Emily Miller (Ed.) The Library Tomorrow, American Library Association, Chicago, p.5.

31. Meadow, A. J. (1978). "Documentation and the Scientific Producer/User." Journal of Documentation, 34(4): pp.324 - 332.

32. Jahoda, G. (1963). Indexing with Edge - notched and Internally punched cards. In Simonton, W. (Ed.) Information Retrieval Today, University of Minnesota, Minneapolis, pp.37 - 47.

33. Costello, John C. Jr. (1966) Coordinate Indexing Volume VII. Rutgers University Press, New Jersey, 219p.

34. Foskett, A. C. (1967). A guide to Personal Indexes. Clive Bingley, London, 80p.

35. Lancaster, Wilfred F. (1968). Information Retrieval Systems: Characteristics, testing and evaluation. John Wiley, London, 222p.

36. Fairthorne, Robert A. (1955). "Essentials for Document Retrieval." Special Libraries, 46: pp.340 -353.

37. Cleverdon, C. W. (1962). Interim Report, Association of Special Libraries and Information Bureaux, Cranfield, England.

38. Swets, J. A. (1963). Information Retrieval Systems. Science, 141: pp. 245 - 250.

39. Cleverdon, C. W. (1962). Report on Testing and Analysis of an Investigation into the Comaprative Efficiency of Indexing Systems. College of Aeronautics, Cranfield, England.

40. Swets, John A. (1969). "Effectiveness of Information Retrieval Methods." American Documentation, 20(1): pp. 72 - 89.

41. Van Rijsbergen, C. J. and Sparck K. Jones (1973). "A Test for Separation of Relevant and Non-relevant Documents in Experimental Retrieval Collections." Journal of Documentation, 29(3): pp. 251 - 257.

42. Salton, G. and Young, C. S. (1973). "On the specification of term values in antomatic Indexing." Journal of Documentation, 29(4): pp. 351 - 372.

43. Vickery, B. C. (1973). Information Systems. Butterworth, London, 350p.

44. Lancaster, F. W. (1969). "MEDLARS: Report on the Evaluation of its Operating Efficiency." American Documentation, 20(2): pp. 119 - 142.

45. Lancaster, F. W. and Grace Jenkins (1970). "Qaulity Control applied to the Operations of large Information System." JASIS, 21(5): pp. 370 - 371.

46. Fairthorne, R. A. (1958). "Automatic retrieval of recorded information." Computer Journal, 1(1): pp. 36 - 42.

47. St. Lawrent, M. C. (1967). A review of the literature of indexer Consistency. In: Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, Britanica Inc., Chicago. 3:p .82.

48. Zull, Carolyn G. and Baumanis George J. (1967). Textual Correlates of relevance: Preliminary Investigation. In: Caudra C. A. (Ed.) Annual Review of Information Science and Technology, Britanica Inc., Chicago. 3: pp. 84 - 85.

49. Lancaster, F. Wilfred (1968). Evaluation of the MEDLARS demand search service. U.S. Dept. of Health Education and Welfare. Public Health Service, National Library of Medicine, Bethesda, Md., 276p.

50. Lufkin, Richard C. Determination and Analysis of some parameters affecting the subject indexing (Thesis B.S). In: Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, Britanica Inc., Chicago. 7: pp.

51. King, Donald W. (1968). Design and Evaluation of Information Systems. In:Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, Britanica Inc., Chicago, 3: p. 119.

52. Tinker, John F. (1966). "Imprecision in Meaning Measured by Inconsistency of Indexing." American Documentation 17(2): pp. 96 - 102.

53. Tarr, D. and Berko, H. (1974). "Factors Influencing interindexer consistency." In Zunde, Pranas (Ed.) Information Utilities: Proceedings of the 37th Annual meeting of the ASIS, Volume II, Oct. 13 - 17, Atlanta G. A. Washington, D.C.

54. Tinker, John F. (1966). Op.cit.

55. Rayward, W. Boyd and Svenonius, Elaine (1967). "
"Consistency, Concensus sets and Random deletio " In:Caudra, C. A. (Ed.) Annual Review of Informa on Science and Technology, Encyclopedia Britanica Inc., Chicago, 3: pp. 61 - 103.

56. Loukopoulos, Loukas (1966). "Indexing Problems and Some of their solutions." American Documentation, 17(1): pp. 17 - 25.

57. Huruitz, F. I. (1969). "A study of Indexer Consistency." American Documentation, 20(1): pp. 92 - 94.

58. Savage, Terry R. (1965). "A Boyle's Law of Indexer Consistency." American Documentation, 16:p. 33.

59. Zunde, Pranas and Dexter, Margaret E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3): pp. 259 - 267.

60. Cooper, Williams S. (1969). "Is Interindexer Consistency A Hobgoblin?" American Documentation, 20(3): pp. 268 - 278.

61. Leonard, Lawrence Edwards (1975). Inter-indexer Consistency and Retrieval Effectiveness: Measurement of relationships (Ph.D. Thesic). In: Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, Britanica Inc., Chicago, 12:pp. 97 -98.

62. Slamecka, V. and Jacoby, J. E. (1965). The Consistency of Human Indexing, in the coming of Age of Information Technology. Documentation Incorper ed, Washington, D.C., pp. 32 - 56.

63. Hooper, R. S. (1965). Indexer Consistency Tests - Origin Measurements, Results and Utilization. In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

64. Jacoby, J. (1962). Methodology for Indexer Reliability Tests. In: Zunde P. and Dexter, M. E. (1969) "Indexing Consistency and Quality." American Documentation, 20(3): pp. 259 - 267.

65. Slamecka, V. and J. Jacoby (1963). Effect of Indexing Aids on the Reliability of Indexers, In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

66. Jacoby J. and V. Slamecka (1962). Indexer Consistency under Minimal Conditions. In: Zunde P. and Dexter, M. E. (1969)."Indexing Consistency and Quality." American Documentation, 20(3): pp. 259 - 267.

67. Painter, A. F. (1963). An Analysis of Duplication and Consistency of Subject Indexing Involved in Report Handling at the Office of Technical Services, U.S. In: Zunde P. and Dexter, M. E. (1969)."Indexing Consistency and Quality." American Documentation, 20(3): pp. 259 - 267.

68. Rodgers, D. J. (1961). A study of Intra-Indexer Consistency. In: Zunde P. and Dexter, M. E. (1969)."Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

69. Korotkin, A. L. and L. H. Oliver (1964). The effect of Subject Matter Familiarity and the use of an Indexing Aid upon Inter-Indexer Consistency. In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

70. Hooper, R. S. (1963). A Facet Analysis System, Automation and Scientific Communication. In: Zunde P. and Dexter, M. E. (1969)."Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

71. Stevens, M. E. (1965). Automatic Indexing: A state of the Art Report. In: Zunde P. and Dexter, M. E. (1969)."Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

72. Bryant, E. C., D. W. King and P. J. Terragno (1963). Some Technical notes on coding Errors. In: Zunde P. and Dexter, M. E. (1969)."Indexing Consistency and Quality." American Documentation, 20(3): pp. 259 - 267.

S.b. Oruma, Ben Oviss (1983). The Relationship between Exhaustivity and consistency at the Optimum Indexing level. Unpublished Ph.D. Dessertation, Case Western Reserve University,

CHAPTER TWO

PROBLEMS INVOLVED IN AN OPTICAL COINCIDENCE SYSTEM.

2.1 History of IAR Library with its Information Retrieval System:

The Rate at which world population is increasing as compared to the quantity or amount of Agricultural products that man in his present level of Activity can produce are not comparable. As a result man has turned to natural science for solution to this ugly situation as it is with his other problems hence the need for Agricultural Research Institutes. It was in an effort to combat such unpleasant situation that Agricultural Research Institutes were established.

The history of Agricultural Research Institutions in Nigeria dates back to the colonial era when the colonialists encouraged Agriculture to satsify their selfish interest of producing the needed raw materials for industries in Europe.

According to Momoh[1], in a developing country like Nigeria where the economy depended on Agriculture in the 1950's and 60's the need for Agricultural Research Institute can not be ov r emphasized; at least for the following two reasons:

(a) Nigerian economy before the oil boom was tied to Agriculture with over 70% of the population engaged in farming and animal rearing, but not enough is being produced to feed the teeming population because of strict adherence to traditional farming methods.

(b) The country has been a victim of drought and pest which forced the Federal Government to substitute what is locally produced with imported food.

The Institute for Agricultural Research Samaru refered to as IAR was established in 1922 as the administrative Headquarters of the then Department of Agriculture in the Northern Provinces of Nigeria. The Agricultural Research In titute was established to conduct research into various diseases that affect animals and crops and to improve both their quality and quantity.

Research started there in 1924 and soon after that a small collection of books and periodicals was started. The activities of these scientists have not only led them to demand the numerous records both published and upublished. Their activities also generated alot of publications out of their researches. The number of publications in that station became too large for one researcher to cope with so the need for a library became obvious. There existed a collection of books needed for research since the early days but untill 1956 when a librarian was appointed that there was systematic effort to gather the stock of Older documents which were scattered in files and offices. The documents have been pr served, and every effort has been made subsequently to gather all documents on Nigerian Agriculture as they have beer produced. They were conventionally catalogued and classifie by the Universal Decimal Classification Scheme. For some tir the need for a semi-mechanized retrieval system to provide speedier and more efficient access to these documents was felt but that did not materialise until in 1974 when a staff was available to undertake the establishment of such a system.

The semi-mechanized information retrieval system of IAR provides an alternative method of searching for the documents which contain information on your subject. Normally when you come to the library seeking information on your field, using the conventional system you have to look up your subject in the subject index to the classification scheme to find the classification numbers which express that subject; then you have to search through the subject catalogue under combinations of those numbers to find the documents which contain information on your subject.

With the semi-mechanized system search is quicker and allows more flexible combinations of subjects. The system has been refered to as an "Information retrieval system" but should more accurately be called a "document retrieval system," in that it retrieves only references to the documents which may supply the required information. The documents theselves have to be obtained from the library stock in the normal way.

Information retrieval system does not provide information on the subject of inquiry, it rather directs inquirer to source of information that can answer the question. F. W. Lancaster[2], defined information retrieval system as,"it does not inform (i.e. change the knowledge of) the user on the subject of his inquiry. "It merely informs him on the existance (or non-existance) and where about of document relating to his request."

According to Oruma[3], the then librarian of IAR initiated the move to establish IR system for the institute library in 1956 and much later an information scientist studied the situation and confirmed the need for the IR system.

After the scope of other IR systems in Nigeria have been studied it was recommended that the system at IAR be designed to cover documents on Nigerian Agriculture. It has been estimated then, that collections on Nigerian Agriculture grow at the rate of 500 documents per annum and based on that estimation it was recommended that the feature card system be employed in the IR, following the recommendations the academic Professional Board of the ABU approved the establishment of Information retrieval system in IAR in late 1972.

The academic Professional Board of ABU approved the purchase of a machine capable of holding not less than 10,000 documents on a set of centrally punched feature cards. Later after thorough market survey Anson equipment was accepted and purchased for economy and compactness. It was acquired in 1974 at a cost of £1,500. It comprises of feature cards in a storage cabinate and a drill mounted on a reader.

Mrs. Helen Nation, who is an information scientist was sent in by Canadian University Services Overseas on a two year term. In 1974 a Nigerian graduate became available and the system started.

Operation of the system involves two procedures; namely input and output. To input information into the system a controlled vocabulary of indexing terms which is called Thesaurus is used to replace the classification Scheme used in the main library.

The subject content of each document is recorded by assigning the appropriate indexing words/terms from the thesaurus.

Each indexing term in the thesaurus is represented by a large plastic card, which is called a feature card. Thus there is one card for "Wheat", one card for "Irrigation" another card for "fungal diseases," etc. On these cards are recorded the numbers of all documents which contain information on that subject; they are recorded by drilling a hole through the space on the card which corresponds to the number of the document. Each card has spaces for 10,000 numbers, or 10,000 documents. (The drilling of the holes by a precision drilling machine is the "mechanized" element of the system). Refer to figure 2.1.

2.2 Present Level of Operation:

2.2.1 Input to IAR System.

It is the process of storing information or feeding information into the system. The procedure of inputting information into the system is as follows:

When a document is first received, it is assigned an accessions number to begin with, then the indexer reads throug the publication to extract the indexable terms. With the help of a thesaurus (controlled vocabulary) the document is assigned term/s that represents its subject content. A document card is then made; it will contain all information on a book/article i.e. bibliographical details of a document, index terms, Botaincal terms and other details of that document. The document cards are made in two, one filled under author (author file) and the other under document number (document file).

Before filling the document cards the document numbers are punched on the feature cards (term cards with numbered

HL4 OPTICAL DRILL.

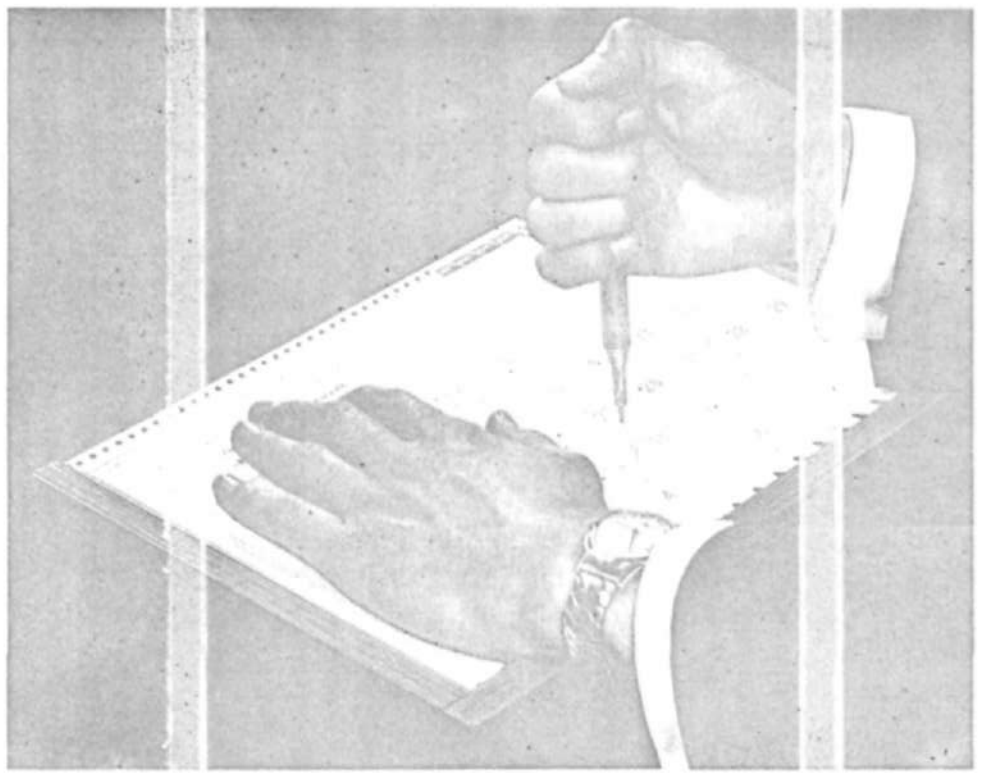


Fig. 2.1

Drilling with precision drill.

Numbers corresponding to the document Numbers are punched on the relevant term cards (feature card) to store or input information into the system.

positions on them). Normally all the index terms in the thesaurus have their feature cards already made for them so at the input stage the appropriate cards are drawn out and the document numbers punched out on them. Then the cards will be filed back. Refer to figure 2.2.

2.2.2 Retrieval From IAR System:

Output is the information that a reader gets out of the system in response to a query. Recalling information is called retrieval.

When you come to the information retrieval Unit seeking for information, you are required to fill an enquiry form (Fig. 2 specifically designed to give the required details necessary for question analysis and subsequent identification of the required materials. After providing the required details of your query, the information officer analyses your question. The next step is selection of the feature cards which express the subject you are interested in, with reference to the thesaurus and with the assistance of the librarian (information officer). The holes which Coincide on all the selected cards represents documents which contain information on your topic, refer to figure 2.4. For easy identification of the holes, the selected feature cards are allinged and placed on a light box (a box with a source of light with one of its sides translucent) and light will show through the Coincidence holes. The hole(s) through which light has shown gives numbers to documents relevant to the query. Refer to figure 2.4.

The document cards relating to the numbers revealed earli are sorted out from the file. The document card will provide full bibliographical detail of the document, its library accession number, classification number, and all the index

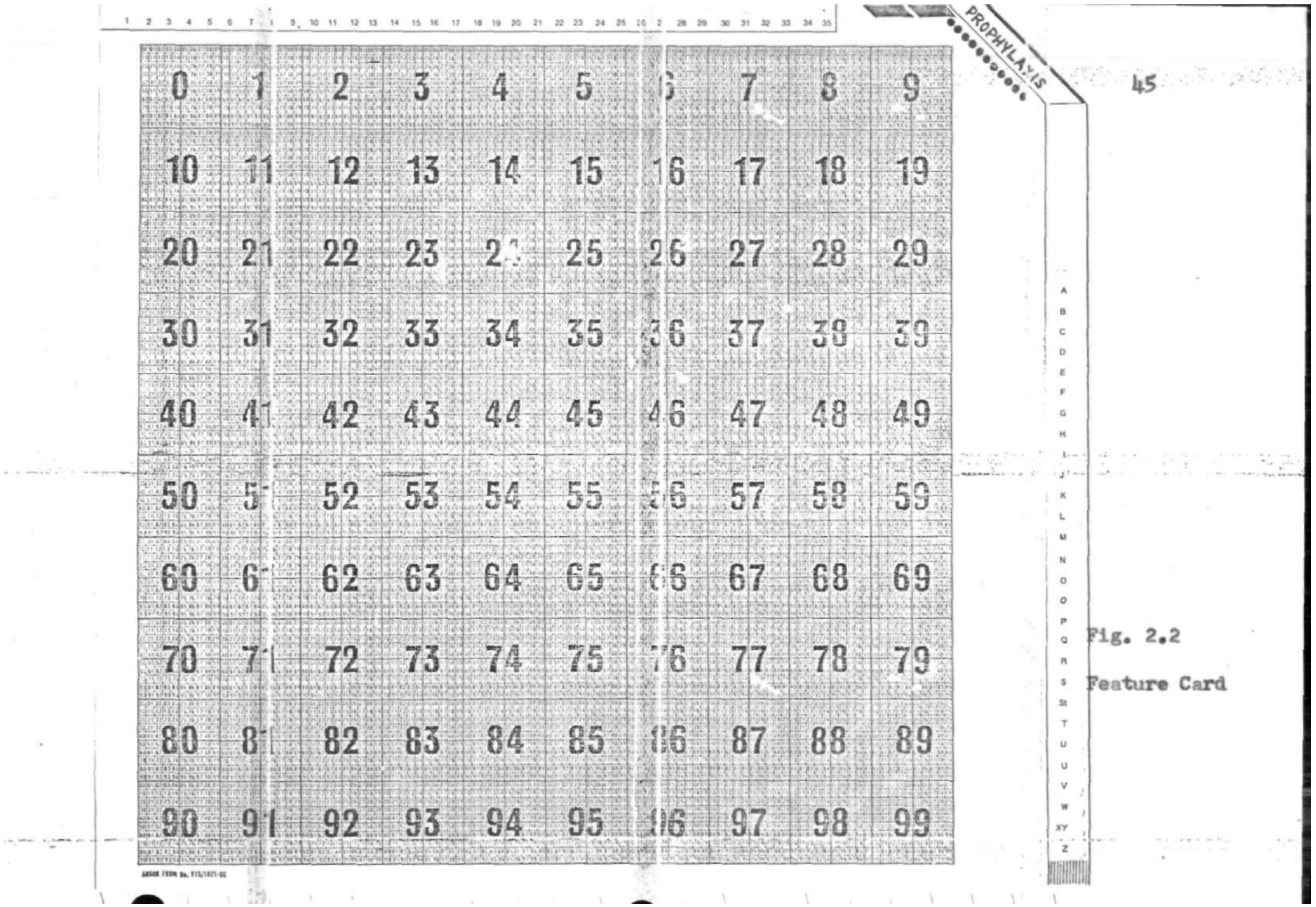


Fig. 2.2

Feature Card

INFORMATION RETRIEVAL SERVICES
INSTITUTE FOR AGRICULTURAL RESEARCH LIBRARY
AHMADU BELLO UNIVERSITY, ZARIA

ENQUIRY FORM
FOR INFORMATION ON NIGERIAN AGRICULTURE

Fig. 2.3

Enquiry Form.

From: ....

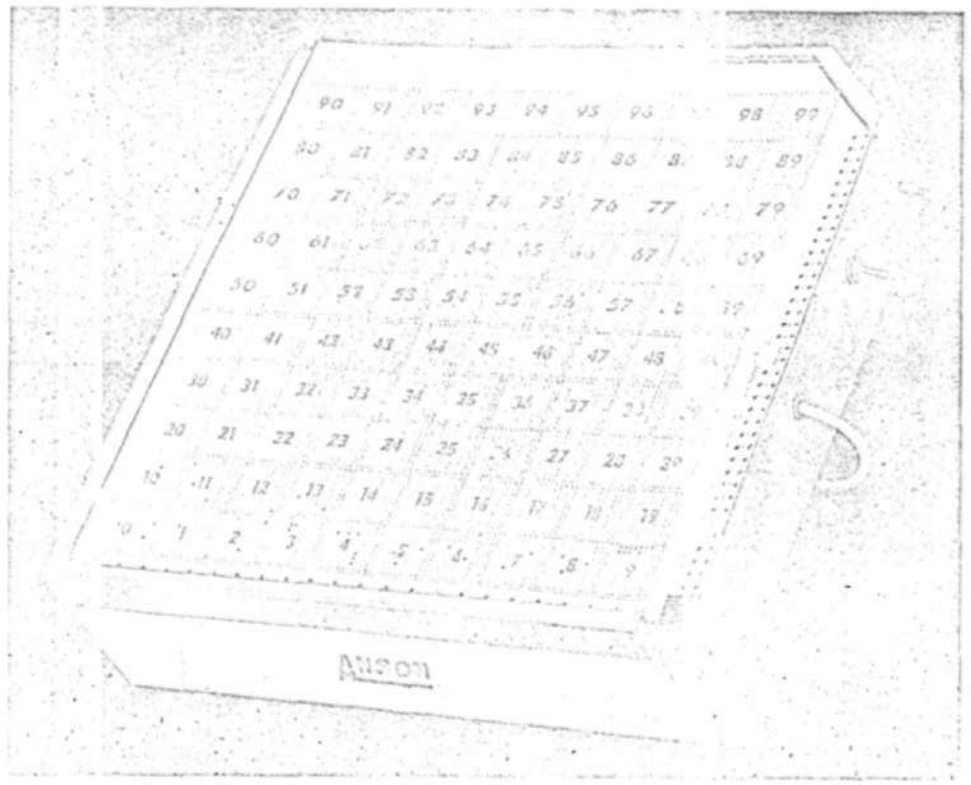
.....

.....

.....

ENQUIRY Please state question precisely ..... (e.g.: What information have you on "the spread of rosette diseases in the Sudan zone"?)	
SUBJECT AREA a) Broad subject area under which enquiry may be grouped ..... (e.g. Crop Rotation)	
b) More specific field of interest. This should reflect the particular subject interest of enquirer ..... (e.g. Insect vectors of plant diseases and resistance to control).	
RELATED KEY TERMS Please use any specific terms that may further describe the Query ..... (e.g. Aphis craccivora rosette diseases viral diseases Arachis hypogaea)	
GEOGRAPHICAL COVERAGE Is enquirer interested in any geographical area (within Nigeria)? ... (e.g. Nigeria (North) particularly Kano State).	
MISCELLANEOUS Any other helpful details/comments ... (e.g. current materials (1970-76 only).	

Fig. 2.4



For easy identification of document numbers, the selected feature cards are allinged and placed on the light box. When the light is switched on,,some holes will let light through. Such holes reveal numbers of the documents relevant to our query.

terms assigned to it. An abstract or summary will be added when available, and when a document is very short, such as newspaper article, the 'full text' will be included on the document card. Refer to figure 2.5.

Fig. 2.5 Document Card.

ABALU, G.O.I. 6402

Supply response to producer prices: a case study of groundnut supply to the Northern States Marketing Board, Zaria, Institute for Agricultural Research, 1976: 8p. 1 table. (Samaru Res. Bull. 268) Reprinted from Savanna, 1975 V.4(1) p.33 - 40 (P. 18762 - 18764) 338.17: 634.58

Agricultural economics
Economic conditions
Legume crops
Marketing

Marketing boards
Mathematical models
Prices
Producer prices
Supply & demand.

Arachis hypogaea

abstract
I.A.R.
Northern
Periodical
articles

The reader or inquirer will be allowed to go through the document cards to see the bibliographical details of the book and where applicable an abstract also is included so that it will enable the reader decide on the suitability or otherwise of the material. If it is found to be relevant, then the information officer will obtain the material from the shelve. Alternatively the reader can go to the shelve for the material.

The Search can easily be varied by adding or removing feature cards, to broaden or narrow the topic, or change its emphasis.

2.2.3 Thesaurus Development:

Thesaurus is a name given to a controlled vocabulary of indexing terms. It is used to replace the classification scheme used in the library.

The first draft of the thesaurus was made in 1975 when the system took off. In building the thesaurus, potential indexing terms were initially collected from the IAR library. These were obtained mainly from the subject index to the UDC which has been in use in IAR Library from its foundation. Others were supplied from IAR research programmes. Free indexing was done for sometime. Consultation with potential users and experts also played a significant role in the assembly of index terms. Established agricultural thesauri were also used in the thesaurus build up. The thesaurus was revised in 1976 and 1981.

2.2.4 Equipments Used in the Retrieval System:

The machine approved for IAR System was to be capable of holding not less than 10,000 documents on a set of centrally punched feature cards. According to Oruma[3], after thorough market survey Anson Equipment was accepted and purchased for economy and compactness. It was acquired in 1974 at a cost of £1,500. It comprised of feature cards of 28cm by 30.4cm, size in storage cabinet and a drill mounted on a reader.

2.2.5 Progress and Plans:

Efforts are being made to up date the input of current materials as much as possible and work has began on the backlog materials from the IAR Library Subject catalogue (Nigerian Section) and the subject catalogue of periodical literature on Nigeria. However, for a more thorough search users would

have to look in both the catalogue and the Retrieval System.

Readers use of the system has increased[4]. Comparing the number of enquiries recieved between January and March 1984 t(that of previous years, it can be seen that only 16 enquiries were attended to in 1980, the number increased to 32 in 1983 and from January to March 1984, 44 enquiries were attended t "It is satisfying (encouraging) to note that most users of th system appreciate it very much because it is faster than going through the catalogue and it is more precise."[4]

It is intended that the system will be expanded to include materials on Tropical Agriculture and West African Agriculture in general.

It is intended that somebody with B.Sc. Agric, or Botany will be given training in librarianship so that he replaces holders of Diploma in Library Science who operate the system now. This change will enable expert analysis of queries so that they fit into the most appropriate thesauric language. It will also facilitate better indexing.

Researchers are now being encouraged to use the system through monthly accessions lists and at times they are allow i to try their hands in information retrieval.

2.3 Operational Problems: Hardware:

There is problem of spare parts for the machine. Spare parts have to be imported in the event of parts needing replacement. These days of austerity it is not easy to get Foriegn exchange for the importation of spare parts. Out of the three reading/drilling machines only one performs a combination of the two functions of reading and drilling. The

other two can only perform one of the two functions (i.e either to drill or to be used as reader only). The drilling section of the only machine that performs the two functions as of now does the drilling for a short time, that is, it automatically ceases to work after drilling about fifteen cards and resumes after about three hours. This has slowed down the input process into the system and it has resulted in a large backlog.

The machines require electricity supply to be operated and IAR has no standby generator, it depended on NEPA for its power supply. When there is black out or power failure from NEPA the Information Retrieval Unit neither inputs information into the system nor does it attend to quiries from users. That unit remains out of use for as long as there is no power supply from NEPA.

There are no known technicians that handle the repairs of these machines in the country. In the event of brake down or servicing the technician has to be brought in or the machines have to be sent abroad. This takes us back to the problem of Foriegn exchange.

The machine performs one function at atime, that is you can't read and drill with it, all at a time. It can only be used by one person at a time. When there is a rush of readers the process of attending to one reader after the other has made information retrieval very slow and has defeated the aim of giving quick information.

When using the machine it reveals document numbers only and not the information or bibliographical details of the document containing the information.

Not all the punching machines are precise. One of the punching machines is not a precise punching machine. It can not be used in punching more than one card at a time. If more than one card is punched with it at a time there is tendency that one hole will be larger than the other and may likely not be on the same line. This imprecision has led to so many problems of information retrieval.

The system presents problem of searching in many places before one gets information. The system is made in such a way that, you search for the term cards representing a query, then allign the cards to get document numbers. Using the document numbers revealed you sort out the document cards and subsequently you retrieve the documents. The searching stage are so many that the whole process becomes slow at the end.

2.4 Operational Problems: Software:

The thesaurus needs to be up-dated either by issue of supplements or by way of revision to accommodate new terms. The thesaurus was first compiled in 1975 and was last revised in 1981. Many changes have taken place from 1981 to date and these changes need to be reflected in the thesaurus if it is agreed that thesaurus is dynamic. Many inconsistencies found their roots from the thesaurus, more especially in geographical location.

There is problem at the Geographical Section of the thesaurus in giving a good relationship between Related Terms (RT), Broader Terms (BT) and Narrower Terms (NT). This has led to having different entries for the same subject using different terms. Typical examples are entries for Adamawa Division and Adamawa Province. Adamawa Province is a broader term for Adamawa Division and ideally entries for the Division

should appear on the term card for the Province; but this is not the case. Therefore Faulty thesaurus leads to duplication, false retrieval and other problems in the system.

Inconsistency in indexing - at the initial stage of the information retrieval system of IAR, there was no thesaurus specifically built for the system to work with, except some thesauri meant for other collections elsewhere. Later a thesaurus was specifically built for the system to work with. Therefore the little indexing done at the initial period is not consistent with that done at a later date.

Feature cards are costly and not easy to get locally so they have to be imported.

The size of the numbers to be punched on the feature cards are so small that it makes it almost impossible to read and record the numbers without much error. Numbers can at times be misread or miscopied. A typical example of miscopying was when 7212 was copied instead of 7213 and the wrong document was retrieved. This anamoly was rectified by going back to the term card to read once more.

Holes for withdrawn documents or those punched by mistake are normally patched with a light piece of paper. Many such pieces, used for patching the holes have been torn off due to constant use. This problem has led to many false retrievals. When such numbers are revealed during a search and their document cards checked, it will be found indicated that the document is withdrawn from the collection or lost. There are cases where document cards for withdrawn documents have been removed from the tray but the holes on the cards still remain. It also leads to false retrieval.

In some instances during a search the points of coincidence where light will show through the holes are so many, up to Sixty or more. It becomes difficult for one to record all such numbers with any degree of accuracy. Where it is done the reader may not require that much documents to get his querry answered. This large document numbers revealed will make it difficult for the information officer to decide which materials to present to the reader and which ones not to present. There is lack of strict adherance to terminal indexing point. That is the maximum or minimum index terms to use per document is not observed by indexers as a rule. In some instances the indexing is highly exhaustive or too broad. Where it is highly exhaustive it leads to the retrieval of materials of low relevance to queries. Where it is too specific, relevant documents may be missed at the time of retrieval and that means the information is lost in the system.

There are duplications in the entries made. There are times when two terms meaning almost the samething have feature cards made for them each. Different documents are entered under each term. When searching if you ignore one term and look under the other term, you are bound to miss alot of relevant documents. Examples are cards for Adamawa Province and Adamawa Division, discussed earlier. This problem stems from thesaurus.

2.5. Operational Problems: Non-machine:

Input to the system gives alot of problems due to lack of enough hands to index the documents. Presently there is a large backlog of materials awaiting to be indexed and current

materials also filing. This problem of indexing has made the system not up to date and not comprehensive on documents on Nigerian Agriculture. There are only two indexers doing the job with one assistant who cross-checks typed document cards for correction, and assists in information retrieval and files back all cards drawn out for consultation. The assistant also does the job of drilling holes on feature cards for input. At times the only assistant will be on afternoon duty and this slows down the already slow process in the system. There are no enough full time typists to type the document cards of the few documents indexed. The present typists combine the job of typing official letters and document cards, this too slows down the process of input.

On many occassions wrong cards have been pulled out for correctly copied numbers due to human error.

There is problem of wrong punching of holes. The holes on some of the feature cards are badly done and this leads to problem in information retrieval. Some holes are too large while others are too small and in some cases the points punched are not in a straight line. Where the holes are not in a straight line, at the time of retrieval light will not show through even though the correct cards are alligned. This simply means many relevant documents may not be retrieve due to wrong punching.

Readers are not allowed to use the system on their own except with the assistance of information staff or system operator. The system is only accessible to readers during the working hours of 7.30 a.m to 3.30 p.m. It is not accessible after 3.30 p.m and Saturdays even though the library is open.

305035

The information contained in the database on Nigerian Agriculture is neither up to date nor comprehensive. For a more thorough search users would have to look in both the catalogue and the retrieval system.

2.6 Problems of Inconsistency and Consistency Measure.

2.6.1 Problems of Inconsistency:

According to Oruma and Saracevic[5], studies in indexing and the consistency of indexing have shown that inconsistency is a major problem in manual indexing. While much of the inconsistencies associated with manual indexing may be attributed to differences in human assessment, most is due to the variation in the models, the measures and the definitions adopted by the different investigators.

As a rule any two indexers, indexing one and the same document individually, will select sets of indexing terms which are most unlikely to be identical. That is, if one compares indexing terms assigned by any two indexers to the same document one discovers that, the indexers differ considerably in their judgement as to which terms reflect the contents of the document most adequately. Therefore indexing consistency is essentially a measure of the similarity of reaction of different human beings processing the same information. It is clear that thi differences in the judgement of indexers introduces a great deal of uncertainty in any information retrieval system based on human indexing.

Hooper[6], cited consistency values ranging from 10% to 80%, depending on the condition under which the indexing was performed and the measure of consistency used. The disparity in consistency values recorded for different indexing exercises

has remained the principal problem associated with manual indexing. There is no study that is known to have produced 100% consistency.

Some Researchers have come to accept inconsistency in indexing as human. According to Loukopoules[7], human inconsistencies in assigning index terms is inevitable. Savage[8] also is of the Opinion that one need not expect that two indexers will for any given document, assign either the same or very similar sets of terms. Rather what one should expect is that for any homogenous collection of documents any two indexers (or one, etc) will assign terms with roughly the same relative frequencies and that a document set retrieved on the basis of one indexing will be equally satisfactory to that set retrieve by the other.

According to Huruitz[9], it does not seem possible for a system to retrieve documents effectively if they are not indexed consistently. He pointed out that once this point is recognized, the problem becomes not to eliminate inconsistency, for this is impossible, but to discover what types of inconsistencies occur and the best ways to minimize the effects of these on document retrieval.

Different authors use different formulae for measuring degree of interindexer consistency. In some circumstances indexers were observed to be almost totally inconsistent with one another, while in other situations, they were only moderately inconsistent. In no case was perfect consistency observed. According to Oruma and Saracevic[10]. The measures and definitions adopted in any indexing consistency research

can very easily influence the results and the reliability of the results derived from such experiments.

From the foregoing it can be seen that inconsistency in manual indexing is a necessary evil that men have to do with. The best that can be done about inconsistency is to minimize its effect on document retrieval. The measures and definitions adopted in various consistency studies are substandard measures and definitions and this is partly responsible for the disparities in consistency values. According to Leonard[11], we do not yet have any standard measure of consistency. This absence of standard is partly responsible for the disparities in consistency values.

2.6.2 Consistency Measures:

A review of the various consistency measures used in earlier research by authorities like Hooper, Korotkin and Oliver, King and Bryant, Rodgers and the rest of them will enable one decide which of the models is most appropriate in this situation.

Hooper's One-indexer Model.

Using pair-Wise consistency measures, Hooper derived a one indexer consistency measure CG_i with respect to a group of two or more indexers[12].

He defined CG_i with respect to a group of three indexers I_i, I_j and I_k as

$$CG_i = \frac{CP_{ij} + CP_{ik}}{2} \quad \ldots\ldots (1$$

Where CP_{ij} = pair-wise consistency measure between indexer I_i and indexer I_j ;

and CP_{ik} = pair-wise consistency measure between indexer I_i and indexer I_k.

Hooper's consistency measure of one indexer with respect to a group of any number of indexers, N is given as:

$$CG_i = \frac{CP_{ij} + CP_{ik} + \ldots CP_{iN}}{N-1} \quad --(2)$$

where CP_{iN} = pair-wise consistency measure between indexers I_i and I_N, where, N, is the total number of indexers in the group.

Hooper's measure is actually the mean of all pair-wise consistency values between one indexer and all other indexers. This measure can be applied to any number of pairs without any problem. However large the number of authors may be, they can be comfortablly accommodated by this model.

Korotkin & Oliver.

They "used an indexer - pair concept in their computation of consistency measures. This measure is the ratio of number of pairs achieved' to 'total number of pairs"[13]:

$$\text{Consistency} = \frac{\text{Number of pairs achieved}}{\text{Number of possible pairs}} \quad \ldots(3)$$

Where the numerator is the number of pairs of indexers assigning the same terms to a given document.

The denominator is the number of pairs possible if there were perfect agreement.

Oruma and Saracevic pointed out that, "with a higher level of exhaustivity, the derivation of consistency by the Korotkin - Oliver model could be computationally cumbersome."

This model can best be used in a system where inter-index consistency for a number of indexers over a given document is to be measured. That is, the indexers must index a given document and assign a fixed number of terms. The other disadvantage of this model is that, it works best when the exhaustivity level is so low, i.e. it will be easy to compute consistency.

This model will not be a good measure when applied to this particular situation because we are not measuring consistency over a given document rather consistency within sets of similar documents of which the level of exhaustivity is not fixed.

King and Bryant's Model.

Their approach is similar to Korotkin-Olive formula which is founded on indexer-pair consistency measures[14]. It is based on the frequency of assignment for a given term over N, documents. "It is represented Symbolically as:

$$C = \frac{N_{11}}{N_{01} + N_{10} + N_{11}} \quad \text{--- (4)}$$

The variables are represented by a 2-by-2 contingency table:

	2nd Indxer Term not Assigned	2nd Indexer Term Assigned
1st Indexer Term not Assigned	Noo	No1
1st Indexer Term Assigned.	N_{10}	N_{11}

Table 1. 2-by-2 contingency table representing frequency of Assignment for a given term over N documents.

N_{10} = number of times the term was assigned by 1st indexer only.

N_{01} = number of times the term was assigned by 2nd indexer only.

N_{11} = the number of times the term was assigned by both indexers.

King and Bryant's measure of consistency which is consistency by reference to one term at a time is not a very useful measure; it does not hold much practicality for every day indexing routines[15]". This model also like the earlier Korotkin-Oliver model it can not be a good measure to be applied to the case at hand. This model can best be used in a situation where consistency by reference to one term at a time is being measured.

Rodger's Model.

Rodger[16], defined consistency of a pair as the ratio of "words in common' to 'total words assigned'. Rodger's definition was expressed by Hooper as:

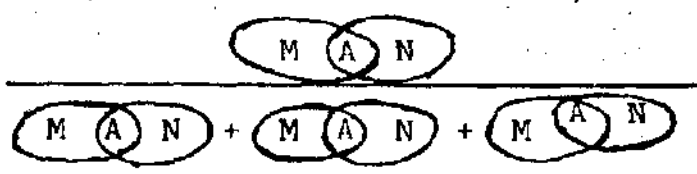
$$CP(\%) = \frac{100.A}{A + M + N} \quad \ldots (5)$$

Where A = N^{o} of term agreement between indexers I_m and 1 ,

M = N^{o} of terms used exclusively by indexer I_m

and N = N^{o} of terms used exclusively by indexer I_n

The relationship has been expressed graphically by Boolean representation as:

CP =  ... Fig. 2.6.

The sum, A + M + N, is infact the union of M and N,

Therefore Boolean representation in Fig. 2.6 above reduces to:

$$CP(\%) = \frac{100 \times M \cap N}{M \cup N} \quad \ldots (7)$$

This model is good for consistency pair measurement between two indexers and at $_{\wedge}$ times even more than two, only that the operations became cumberson with more indexers. It can not be easily used in measuring the consistency of one-indexer compared to the whole group of indexers in the system. It is therefore not suitable in our case.

Slamecka and Jacoby Model.

Two separate variables were specified in their consistency measures: (a) percentage of matching terms, and (b) depth of indexing which is defined as the number of descriptors selected per article by each indexer[17].

Symbolically, terms matched is:

$$\frac{N(T_i\ T_j\ T_k)}{N(T_i\ T_j\ T_k) + N(T_i) + N(T_j) + N(T_k) - N(T_iT_j) - N(T_iT_k) - N(T_jT_k)} \quad ..(8)$$

$N(T_i)$, $N(T_j)$ and $N(T_k)$ are the number of terms used by the three indexers I_i, I_j, and I_k respectively.

$N(T_iT_j)$, $N(T_iT_k)$ and $N(T_jT_k)$ are the number of terms matched between the three pairs of indexers.

$N(T_iT_jT_k)$ = number of terms matched among all three indexers.

For a two - indexer exercise, percent consistency by the above measure would be:

$$C_{ij}(\%) = \frac{100.\ N(T_i\ T_j)}{N(T_i) + N(T_j) - N(T_iT_j) + N(T_jT_i)} \quad \ldots(9)$$

$- N(T_i\ T_j)$ and $+ N(T_j\ T_i)$ cancel out so that equation (9) becomes:

$$C_{ij}(\%) = \frac{100.N(T_i\ T_j)}{N(T_i) + N(T_j)} \quad \ldots(10)$$

This model also is like the Rodgers models when the number of indexers increase or becames large it is difficult to use. This model is good for measuring consistency of indexers on a given document and not all that suitable for measuring consistency of one-indexer with respect to a group of indexers.

Pair-wise Consistency Measure.

According to Zunde[18], the consistency measure used in most previous studies was the ratio of the number of terms selected by all indexers in the group to the total number of different terms selected for the document. For a group of two indexers the consistency measure on a given document is thus defined as:

$$C_{ij} = \frac{n\ (T_i\ n\ T_j)}{n\ (T_i\ u\ T_j)} \quad \ldots(11)$$

Where T_i and T_j denote the sets of terms selected by indexers I_i and I_j respectively and n(T) denotes number of elements in set T.

This measure can be extended to K> 2 indexers by taking

$$C_{ij...k} = \frac{n\,(T_i \cap T_j \cap ..nT_k)}{n\,(T_i \cup T_j \cup ...uT_k)} \quad ...(12)$$

The main disadvantage of such a consistency measure is the underlying assumption that all the indexing terms selected are equally significant and relevant for the representation of the information content of the document. It is suitable for measuring inter-indexer consistency over a given document It is not most suitable for measuring one-indexer consistency with respect to a group of indexers.

According to Oruma and Saracevic[14], in general inter-indexer consistency tests involving three or more indexers require the introduction of a fuzzy set modification. The pair-wise consistency measure can however be applied to any number of indexers without the need for modification.

From the consistency models reviewed so far it can be seen that non-fits our particular situation better than Hooper's one indexer consistency model. Therefore in this research we are going to use Hooper's model to measure the consistency of one-indexer with respect to a group.

References:

1. Momoh, Sidiku (1979). A Comparative study of Readers Services in the Libraries of National Veterinary Research Institute, Vom and Institute of Agricultural Research. Unpublished B.L.S. Project, Ahmadu Bello University, Zaria, Nigeria.

2. Lancaster, F. W. (1968). Information Retrieval Systems: Characteristics, Testing and Evaluation, Wiley, New York, 222p.

3. Oruma, Ben Oviss (1978). The Information Retrieval System In Nigerian Agriculture (An Evaluative Study of the Feature Card System of the IAR Samaru). Unpublished M.LS Thesis. Ahmadu Bello University, Zaria, Nigeria.

4. Unpublished guide to the Information Retrieval System in IAR.

5. Oruma, Ben Oviss and Saracevic Tefko (1983). Consistency Studies: Eliminating the Disparities Resulting from Consistency Models and Definitions. CWRU, Cleveland Ohio 44106. PP 1-11.

6. Hooper, R. S. (1965). Indexer Consistency Tests - Origin Measurements, Results and Utilization. In Zunde P. and Baxter, M. E. (1969). "Indexing Consistency and Quality." American Documentation 20(3): pp. 259 - 267.

7. Loukopoulos, Loukas (1966). "Indexing Problems and Some of their Solutions." American Documentation, 17(1): pp. 17 - 25.

8. Savage, Terry R. (1965). "A Boyle's Law of Indexer Consistency." American Documentation, 16:p. 33.

9. Huruitz, F. I. (1965). "A Study of Indexer Consistency." American Documentation, 20(1):pp. 92 - 94.

10. Oruma, Ben Oviss and Saracevic Tefko (N.d). Op.cit.

11. Leonard, Lawrence Edwards (1975). Inter-Indexer Consistency and Retrieval Effectiveness: Measurement of relationships (Ph.D. Thesis). In: Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, Britanica Inc., Chicago, 12:pp.

12. Hooper, R. S. (1965). Op.cit.

13. Korotkin, A. L. and Oliver, L. H. (1964). The Effect of Subject Matter Familiarity and the Use of an Indexing Aid upon Inter-Indexer Consistency. In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation 20(3):pp. 259 - 267.

14. Oruma, Ben Oviss and Saracevic Tefko (n.d.). Op.cit.

15. King D. W. and Bryant E. C. In Oruma, B. O. and Saracevic, Tefko, Consistency Studies: Eliminating the Disparities resulting from Consistency Models and Definitions. pp. 1 - 11.

16. Rodgers, D. J. (1961). A study of Intra-Indexer Consistency. In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3):pp. 259 - 267.

17. Jacoby, J. and Slamecka, V. (1962). Indexer Consistency Under Minimal Conditions. In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3): pp. 259 - 267.

18. Zunde, Pranas and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3): pp. 259 - 267.

CHAPTER THREE

DATA COLLECTION

3.1 Procedure for Data Collection:

After almost all the entries for similar or identical documents in the system have been pulled out, they were compared. Cards with entries for similar or identical documents were paired; and a total of 130 pairs cutting a cross all the eleven main subject areas covered by the system were chosen. After it has been certified that the pairs are for similar or identical documents, they were checked for the period of indexing. That is, when was the indexing done and was it done within the Period P_1 (1974 - 1978) or was it in Period P_2 (1979 - 1986).

Some of the cards have no date of indexing written on them except the initials of the indexers. Where date of indexing is given the date is used and where no date but initials are given, the initials were used in sorting out entries according to Periods P_1 and P_2. The initials helped in determining the period of indexing because the period within which the various indexers did the indexing in the system is known.

At the end of the sorting exercise, pairs having entries for similar or identical documents with either of the cards falling in period P_1 and P_2 were retained; while pairs with both entries falling within the same period or pairs found to have been indexed by the same indexer were dropped. After this exercise 64 pairs were found to have met the criteria set for this study. This pairs were then compared for inter-indexer consistency.

Degree of consistency is determined by totalling the number of identical terms assigned to each pair of documents by the indexers, as compared to the total number of terms assigned to that pair of document.

The cumulative consistency of a pair of indexers is measured as follows and the results expressed in percent.

Author IAR
Title: Notes on the Socio-econo-
mic and extension programme.
Cropping Scheme Meeting, 1978.
Zaria, 1978. 78p. 17 tables.

Author IAR
Title: Notes on the Socio-
economic and extension
programme. Cropping Scheme
Meeting, 1980. Zaria, 1980.
83p.

	Ti
1.	agricultural economics
2.	annual report.
3.	bibliographies.
4.	economic conditions.
5.	extension.
6.	sociology.
7.	body (IAR)
8.	form (reports)

	Tj
1.	agricultural economics
2.	research.
3.	sociology
4.	body (IAR)
5.	form (reports)

$T_i = 8$

$T_j = 5$

$T_i \cap T_j = 8$

$T_i \cup T_j = 13$

$$\therefore C_{ij} = \frac{T_i \cap T_j}{T_i \cup T_j}$$

$$= \frac{8}{13} = 61.5\%$$

In this situation the lines joining the terms in the two entries indicate identical terms assigned by the two indexers to the pair of documents.

From the above entries it can be seen that the total number of identical terms assigned to that pair of documents by the indexers is 8; and the total number of terms assigned to that pair of documents is 13. Therefore the degree of consistency would be $8/13 \times 100 = 61.5\%$.

Therefore the degree of consistency of the pair of indexers in this case is 61.5%.

Consistency for each pair of indexers in the system was calculated using the pair-wise consistency measure for all possible pairs by using the following formula[1]:

$$C_{ij}(\%) = \frac{T_i \cap T_j}{T_i \cup T_j} \quad \ldots(1)$$

where C_{ij} = Consistency pair of indexer i and indexer j;

$T_i \cap T_j$ = Number of terms selected by indexers i, and indexer j, in common.

$T_i \cup T_j$ = total number of terms selected by both indexers i and j in indexing the documents.

After percent consistency value for all possible pairs have been calculated, the results were tabulated for each pair. Later average percent consistency measure for each pair was taken and Hooper's[2], one-indexer consistency measure was applied to the data.

Hooper's consistency measure of one-indexer with respect to a group was applied to get the consistency of each indexer with respect to the whole group of indexers in the system.

Hooper's one-indexer consistency measure for any number of indexers used the following formula:

$$CG_i = \frac{CP_{ij} + CP_{ik} + \ldots CP_{iN}}{N - 1} \quad \ldots(2)$$

where CG_i = One - indexer consistency measure with respect to a group of two or more indexers,(i.e indexer I_i).

CP_{ij} = Pair-wise consistency measure between indexer I_i and indexer I_j.,

CP_{ik} = Pair-wise consistency measure between indexer I_i and indexer I_k.,

CP_{iN} = Pair-wise consistency measure between indexers I_i and I_N where N, is the total number of indexers in the group.

Agroup of five indexers are involved in this study; that is entries made by Five indexers are used.

3.2 The Cut-Off Date:

For the purpose of this study, the indexing period was divided into two. That is period P_1 (1974 - 1978) and Period P_2 (1979 - 1986). The period 1974 to 1978, was from the inception of information retrieval system in IAR to the end of 1978 when both, the two founding members of the system (indexers) left IAR. Even though the two indexers did not leave at a time, but by the end of 1978 they had both left IA

The period 1979 to 1986 is the period when another set of indexers come to the system and started indexing. The indexers falling within this period came at different times, some have left while others are still on with the system.

The cut-off date is fixed at this period because it is more convinent and easy to tell when one indexer stopped and the other started. The period when the first two indexers stopped and the others started has Coincided with the end of the year. All other periods when indexers changed hands fall within the year (i.e. some months in the year). In all the remaining periods with exception of the one chosen as cut-off point, there is an overlap of indexers (i.e. indexers have to work with other indexers for Long before they leave). There is a case where an indexer left for school and came back to continue the indexing and due to lack of indexing dates on the entries it is difficult to say when entries for two documents indexed by such an indexer were made. The above reasons made it difficult to approximate the cut-off date for any other period, other than the one chosen. If a.y other period were chosen the approximation will be a poor one and that will affect the validity of the findings.

The amount of indexing done within the period 1974-1978 is about half of what was done in the period 1979 - 1986. From 1974 - 1978 a total of 2160[3], entries was made in the system, while from 1979-1986 a total of about 4,400 entries was made in the system. If any other period was fixed as a cut-off date, it will affect the balance of entries made into the system.

Since the study is dealing with an existing system it is good a cut-off date is fixed to be able to measure inter - indexer consistency.

3.3 Consistency Values for Data Set:

Pair of Indexers

Percent Consistency Values.

HN & LA	HN & TK	HN & RM	BO & LA	BO & TK	BO & RM
%	%	%	%	%	%
60	44	62.5	83	72	55.5
60	72	53	61.5	60	42
43.5	54.5	38	51.6	60	55.6
66.7	66.7		70.6	59	80
63	66.7		68.9	50	64.5
73.7	54		47	60	97
55.6	55		60	66.7	54.5
61.5	35.7		89		60.6
95			63		62.5
74			64.5		
57			57		
93			80		
60			64		
76			61.5		
25			61.5		
31			46		
66.7			83		
68					
66.7					
78					
1273.9	448.6	153.5	1112.1	127.7	572.2
63.7	56	51	65.4	61	63.6

Table 3.1. Percent Consistency Pairs for All Possible Pairs in the System.

Summary of Percent Consistency Pairs.

HN & LA = 63.7%

HN & TK = 56%

HN & RM = 51%

BO & LA = 65.4%

BO & TK = 61%

BO & RM = 63.6%

To calculate one-indexer Consistency with respect to a group we use Hooper's formula which is as follows:

$$CG_i = \frac{CP_{ij} + CP_{ik} + \ldots CP_{iN}}{N - 1} \quad ..(3)$$

This formular is used because the Number of indexers in the group is more than three.

Key

H - HN

B - BO

L - LA

T - TK

R - RM

$CP_{HL} = 63.7\%$

$CP_{HT} = 56\%$

$CP_{HR} = 51\%$

$CP_{BL} = 65.4\%$

$CP_{BT} = 61\%$

$CP_{BR} = 63.6\%$

i. $CG_H = \frac{CP_{HL} + CP_{HT} + CP_{HR} + CP_{HB}}{N - 1}$

$\therefore CG_H = \frac{63.7 + 56 + 51 + 0}{5 - 1}$

$= \underline{42.68\%}$

$CG_H = \underline{42.68\%}$

ii. $$CG_L = \frac{CP_{LH} + CP_{LB} + CP_{LT} + CP_{LR}}{N - 1}$$

$$\therefore CG_L = \frac{63.7 + 65.4}{4}$$

$$= \underline{32.28\%} \quad \therefore CG_L = \underline{32.28\%}$$

iii. $$CG_T = \frac{CP_{TH} + CP_{TB} + CP_{TL} + CP_{TR}}{N - 1}$$

$$\therefore CG_T = \frac{56 + 61}{4} = \underline{29.25\%}$$

$$\therefore CG_T = \underline{29.25\%}$$

iv. $$CG_R = \frac{CP_{RH} + CP_{RB} + CP_{RL} + CP_{RT}}{N - 1}$$

$$\therefore CG_R = \frac{51 + 63.6}{4} = \underline{28.65\%}$$

$$\therefore CG_R = \underline{28.65\%}$$

v. $$CG_B = \frac{CP_{BL} + CP_{BT} + CP_{BR} + CP_{BH}}{N - 1}$$

$$\therefore CG_B = \frac{65.4 + 61 + 63.6}{4} = \underline{47.5\%}$$

$$\therefore CG_B = \underline{47.5\%}$$

3.4 Summary of Tables:

Summary of Percent Consistency Pairs.

HN & LA = 63.7%

HN & TK = 56%

HN & RM = 51%

BO & LA = 65.4%

BO & TK = 61%

BO & RM = 63.6%

It can be seen from the table that the pair of indexers HN and LA have an average Consistency value of 63.7%, indexers HN and TK, 56% and indexers HN and RM with 51%. The pair of indexers BO and LA scored 65.4%, indexers BO and TK got 61% while the pair of indexer BO and RM scored 63.6%.

The consistency values ranged from 51% to 65.4% for a pair. The pairs of indexer BO and LA scored an average of 65.4% making the highest score while the pair of indexers HN and RM scored an average of 51% making the least score in the study.

Generally consistency values of pairs with indexer BO are higher than those pairs with indexer HN, which falls within the same period with BO.

Consistency values decrease with the change of indexers in the system.

References:

1. Zunde, Pranas and M. E. Dexter (1969). "Indexing Consistency and Quality." American Documentation, 20(3): 259 - 267.

2. Hooper, R. S. (1965). Indexer Consistency Test-Origin, Measurements, Results and Utilization: In Oruma and Saracevic (n.d.) Consistency Studies: Eliminating the Disparities Resulting from Consistency Models and Definitions. pp. 1 – 11.

3. Oruma, B. O. (1978). The Information Retrieval System in Nigerian Agriculture (An Evaluative Study of the Feature Card System of the IAR Samaru). Unpublished MLS Thesis. Ahmadu Bello University, Zaria, Nigeria.

CHAPTER FOUR

RESULTS AND ANALYSIS OF RESULTS

4.1 Tabular and Graphical Presentation of Results:

Indexers	Percentage Consistency Values
HN	42.68%
BO	47.5%
LA	32.28%
TK	29.25%
RM	28.65%

Table 4.1 - One-Indexer Consistency with respect to a group using Consistency Measure CG_i.

For a graphical presentation of the data on Table 4.1 you refer to figure 4.1.

4.2 Summary of Results:

It can be seen from either the table or the graph that indexer HN has consistency value of 42.68% with respect to the group of indexers. Consistency of indexer BO with respect to the group is 47.5% while that of indexers LA, TK and RM are 32.28%, 29.25% and 28.65% respectively.

The consistency values of the indexers ranged from 28.65% to a maximum of 47.5%. Indexer BO had the highest consistency value of 47.5%, followed by HN with a value of 42.68%, then indexer LA with 32.28%, indexer TK had a score of 29.25% and indexer RM scored the least, that is 28.65%.

The average consistency of the group is found to be approximately 36% with a standard deviation of 8.5.

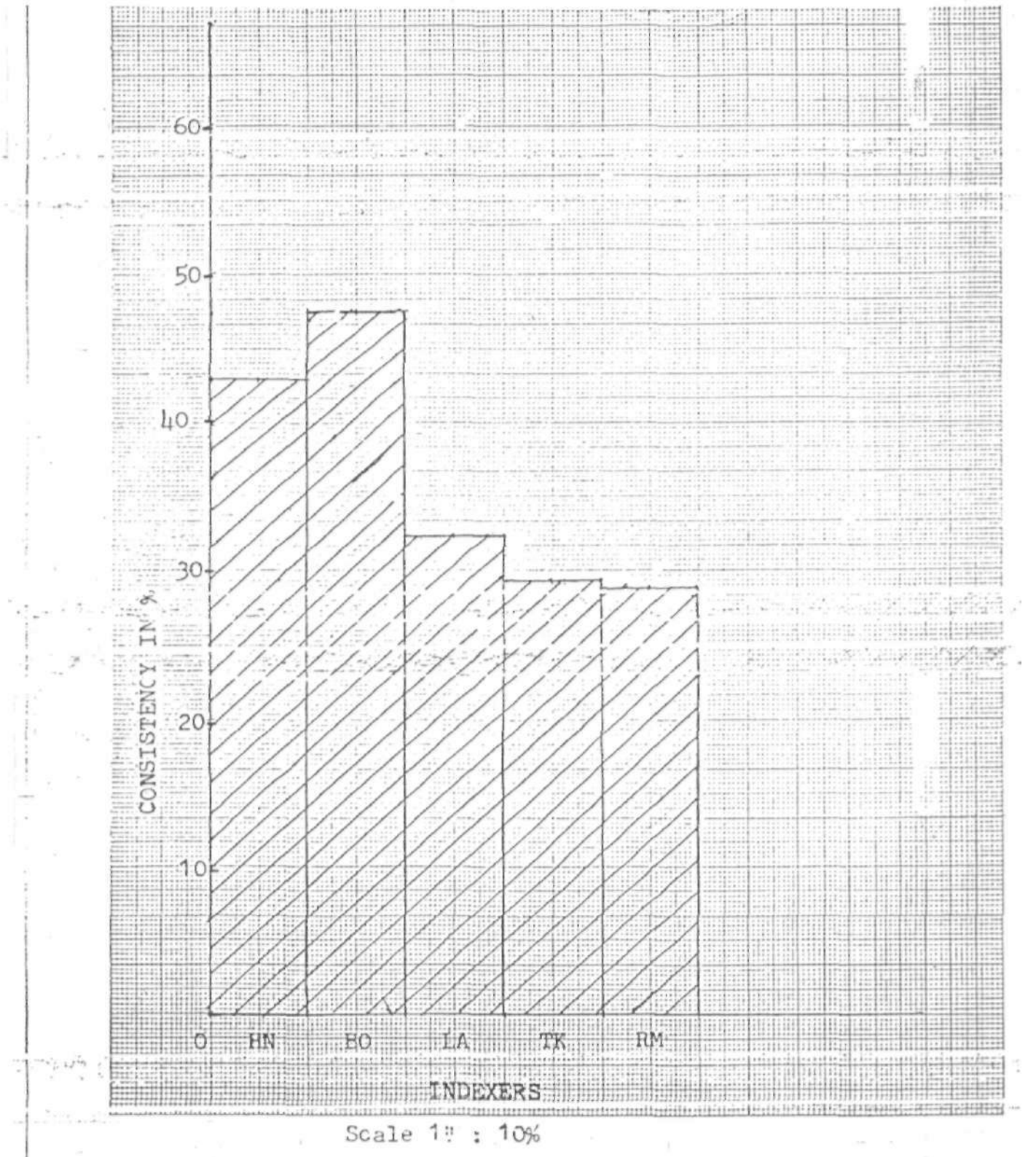


Fig. 4.1 Graphical Presentation of one-indexer consistency values with respect to a group based on the Data in Table 4.1.

Consistency of indexers within the system is on the decrease from period P_1 through P_2. It has decreased (dropped) to an unacceptable level.

4.3 Analysis of Data:

The analysis will be done under two headings. The first part will analyse data on the consistency pair of indexers while the second part will analyse the data on one-indexer consistency with respect to the group of indexers (that is 5 indexers in the system).

4.3.1 Data Analysis of Consistency Pair of Indexers.

It was found that there are six possible pairs of indexers in the system. The pairs are that of indexers HN and LA, HN and TK, HN and RM. The others are pairs of indexers BO with LA, BO and TK, and finally BO and RM.

The average consistency value of the possible pairs are:

HN and LA = 63.7%

HN and TK = 56%

HN and RM = 51%

BO and LA = 65.4%

BO and TK = 61%

BO and RM = 63.6%

The average of all pairs in the group is $360.7/6 = 60\%$.

The Percentage Consistency of the pairs of indexer HN with LA ranged from as low as 25% to a maximum of 95%. The average consistency value of the pair is 63.7% which is above the group average of 60%.

The 95% consistency value obtained for the pair is never known to have been reported in any experiment on consistency Measure. The highest consistency value so far recorded in any study was 80% (1, 2, 3, 4, 5). They indexed the highest number

of identical materials compared to all in the group. That is a total of 20 pairs of identical materials were indexed by the two indexers.

(b) The consistency values of the pair of indexer HN and indexer TK ranged from 35.7% to a maximum of 66.7. The gap between the two extremes is wide but not as wide as that of indexer HN and LA, where there is a difference of 70%. The average consistency of the pair HN and TK is 56%, which is below group average of 60%. The pair indexed 8 pairs of identical materials in the study.

(c) The pair of indexer HN and Indexer RM indexed a total of 3 pairs of identical documents in the system, the smallest number in this study. Their consistency values ranged from 38% to 62.5% with an average of 51%. The average of this pair is well below group average of 60%.

(d) The consistency values of pair of indexers BO and LA ranged from 47% to a maximum of 89%. The maximum score of 89% consistency value of the pair also is above values reported for earlier studies of consistency[6]. The pair indexed a total of 17 pairs of identical documents and scored an average of 65.4%. The average of this pair is above group average of 60% and it is the highest pair average in the whole study. This pair also indexed the second largest number of identica documents (i.e. 17) after the pair of indexer HN and LA with 20.

(e) Consistency values of the pair of indexer BO and TK ranged from 50% to 72%. The difference between the two extremes is about 22%. They indexed a total of seven pairs of identical documents and scored an average consistency value of 61% which is about the same with the group average.

(f) The pair of indexer BO and indexer RM scored consistency values ranging from as low as 42% to 97%. The gap between the highest and lowest score of the pair is very wide and the consistency value of 97% is far above what has been reported in earlier studies on consistency. The pair had an average consistency value of 63.6% which is above the group average. The average consistency value of this pair is the third largest score in the study.

(g) The pair of indexer BO and LA had the highest average consistency value of 65.4%, followed by pair of indexer HN and LA with 63.7%, then by the pair of BO and RM - with 63.6%, then the pair of BO and TK - 61%, HN and TK with 56% and the pair of HN and RM with the least score of 51%.

The differences between the consistency values of the pairs is not much. They all ranged from an average of 51% to 65.4%.

The variability within the pairs is found to be

x	$(x - \bar{X})^2$		
63.7	$(3.7)^2$	=	13.69
56	$(4)^2$	=	16
51	$(9)^2$	=	81
65.4	$(5.4)^2$	=	29.16
61	$(1)^2$	=	1
63.6	$(3.6)^2$	=	12.96
360.7			153.81

$$\bar{X} = 60$$

$$S = \sqrt{\frac{(x - \bar{x})^2}{n - 1}}$$

$$\therefore S = \sqrt{\frac{153.81}{5}}$$

$$= \sqrt{30.762} = \underline{5.54}$$

This variability of 5.54 tells us that the differences between the consistency pairs is not much and they are all clustered around the mean. That is the consistency values are scattered very close to the mean.

4.3.2 Analysis of Data On One-indexer Consistency with respect to the group.

(a) With reference to figure 4.1, the consistency values of indexers with respect to the group ranged from 28.65% for indexer RM to 47.5% for indexer BO.

(b) Indexer BO had the highest consistency value of 47.5% with respect to the group, followed by indexer HN with 42.68%, then followed by indexer LA with 32.28%, indexer TK with 29.25% and indexer RM with the least score of 28.65%.

(c) The consistency values obtained in this study is within the range of values reported by earlier studies on consistency. Earlier studies on Consistency reported values ranging from 10% to a maximum of 80%.

(d) There is a remarkable difference between the consistency values of indexers BO and HN on one hand, and indexers LA, TK and RM on the other hand. The difference between the scores of indexer BO with 47.5% (the highest score) and indexer RM with 28.65% (the least score) is about 18.85%.

(e) The disparity within the consistency scores of the various indexers with respect to the group is wide. The variability within the group is found to be 8.5. That is:

x	$(x - \bar{x})$	$(x - \bar{x})^2$
47.50	11.50	132.25
42.68	6.68	44.62
32.28	- 3.72	13.84
29.25	- 6.75	45.56
28.65	- 7.35	54.00
180.36		290.27

$5 \overline{)\ 180.36} = 36$

Using the formula for standard deviation $s = \sqrt{\frac{(x - \bar{x})^2}{n - 1}}$

$$\therefore \quad s = \sqrt{\frac{290}{4}} = \sqrt{72.5} = \underline{\underline{8.5}}$$

The variability among consistency values of each indexer with respect to the group is wide and much wider than what is obtained for the pairs which was 5.54.

(f). The group average consistency of the indexers is about 36%. Indexers BO and HN with scores of 47.5% and 42.68% respectively have scores above the average while indexers LA, TK and RM with scores of 32.28%, 29.25% and 28.65% respectively are below the group average of 36%.

(g) The difference between the consistency scores of indexers BO and HN that fall within P_1 is small. The differences among those in P_2 is much smaller than those of P_1.

(h) Consistency in the system has fallen drastically from as high as 47.5% to 28.65% within a span of 12 years. That is, it is decreasing with an average of 1.6% per year.

4.4 Speculations:

The outcome of this study led to speculations on pair-wise consistency of indexers and one-indexer consistency with respect to the group of indexers.

(See Appendix II for full names, qualifications and years of Service of indexers.)

4.4.1 Speculations on Pair-Wise Consistency Values:

The following speculations are made on the basis of the pair-wise consistency values obtained from the research.

(a) The pair with high consistency values tell us that the pair has been indexing materials in the system for long. High consistency values of a pair may also mean similarity in subject background and/or experience of indexers.

(b) The larfer the number of identical documents indexed, implies the higher the average consistency values of the pair. Indexer HN and LA indexed 20 pairs of documents and scored average consistency value of 63.7%, while the pair of indexer BO with LA indexed 17 sets and scored 65.4% consistency value.

(c) The high consistency values scored by each pair with indexer LA implies that, indexer LA has been long with the system; see Appendix II. In addition to that the additional qualification of Associate degree in science she has obtained has helped improve her consistency in the system.

(d) The small number of identical documents indexed by the pairs of indexers TK and RM with others in the system may be because indexers TK and RM are new to the system as such they have not indexed much to Warrant high consistency values. Another reason could be because

the two indexers have library Science background, it made them at a disadvantage compared to others in the system. Indexer TK had BLS while RM had DLS.

(e) The pairs of indexer HN and LA, and that of indexer BO with LA Scored the highest consistency values because indexers BO and HN were the founding members of the system so they have indexed much and later joined by LA. This high consistency values could be due to the science background which they have and is relevant to IAR.

4.4.2 Speculations on One-indexer Consistency with respect to the group of indexers in the System.

(a) Consistency values decrease with the change of indexers. That is, with more indexers consistency values tend to decrease and this decrease in consistency values with the increase in the number of indexers over time is due to human differences. Most of the indexers had different background and that contributed in no small way to the differences.

(b) Indexer BO had the highest group consistency value of 47.5% because of the indexer's background in science (relevant area to IAR System) which enabled him analyze document contents and assign the terms that adequately decribe the documents.

(c) Being one of the founding members indexer BO might be using the thesaurus strictly in indexing.

(d) Indexer HN Scored the next highest value to indexer BO. That is, indexer HN scored 42.60%. This high score may be due to the fact that the indexer is a specialist in information handling and experienced in indexing.

(e) Indexer KM got the least score of 28.65% because the indexer is not as experienced as the others in the system. The indexer was the last to join the system and has Diploma in Library Science background which is weak in relation to the IAR collection.

(f) The least figure of 28.65% consistency value means the indexer did the least indexing in the system.

(g) The high consistency values of 47.5% for indexer BO and 42.68% for indexer HN, means they did most or a lot of the indexing into the system.

(h) Indexer TK had the second lowest consistency value of 29.25% because of her library science background and experience. She had BLS which is weak in relation to the IAR collection.

References:

1. Painter, A. F. (1963). An Analysis of Duplication and Consistency of Subject Indexing Involved in Report Handling at the Office of Technical Services, U.S. In: Zunde P. and Dexter, M. E. (1969)."Indexing Consistency and Quality," American Documentation, 20(3) pp. 259 - 267.

2. Korotkin, A. L. and L. H. Oliver (1964). The Effect of Subject Matter Familiarity and the Use of an Indexing Aid upon Inter-Indexer Consistency. In: Zunde P. and Dexter, M. E. (1969)."Indexin Consistency and Quality." American Documentatio 20(3) pp. 259 - 267.

3. Hooper, R. S. (1963). "Facet Analysis System, Automation and Scientific Communication Short Paper." In: Zunde P. and Dexter, M. E. (1969)."Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

4. Stevens, M. E. (1965). Automatic Indexing: A State - of - the - Art Report. In: Zunde P. and Dexter, M. E.,(1969)."Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

5. Bryant, E. C. et al. (1963). Some Technical Notes on Coding Errors. In: Zunde P. and Dexter, M. E. (1969)."Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

6. Zunde, Pranas and Exter, Margaret E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3): pp. 259 - 267.

CHAPTER FIVE

SUMMARY AND CONCLUSION

5.1 Summary:

Alot of time is used-up in the search for specific information by researchers and the need for quick access to information is on the increase daily. In trying to find solution to the problems of handling information people have deviced many ways and means of storing and retrieving information over the years. Among the means deviced are Pre-coordinate systems, such as Dewey Decimal Classification, Universal Decimal Classification, Library of Congress, Colon Classification etc. The Others are Post - Coordinate systems such as Uniterm, peek-a-boo and edge-notched cards. Of all the invention of third generation computers enhanced capabili .es of processing information greatly.

Most, if not all of these information handling devices relied on human indexers for input into the system, but human indexer is not consistent. This inconsistency affects retrieval effectiveness of a system. Earlier studies in indexing and the consistency of the indexing have shown that in-consistency is a major problem in manual indexing. Based on this background this study examined the evolution of consistency values within an optical coincidence system of the Institute of Agricultural Research Library, at Ahmadu Bello University, Zaria in order to evaluate the present level of performance and where necessary improve on the retrieval performance of the system.

The study intended to find answers to the following questions:

a. Is there indexing consistency in the system over the period 1974 to 1986? That is measuring consistency to establish continuity.

b. If there is inconsistency, then identify some of the factors responsible for the inconsistency.

In the study variables such as document length, document type/form, change in research direction and revision of thesaurus are assumed to have no effect on consistency level in the system.

In conducting the research index entries for similar or identical documents indexed at different times by various indexers were examined for inter-indexer consistency. The degree of consistency was determined by totalling the number of identical terms assigned to the documents at each indexing, as compared to the total number of terms assigned to the two documents. Index entries from the inception of the system in 1974 to date were examined.

Consistency pair for all possible pairs in the system was calculated and using Hooper's Model the Consistency of each indexer with respect to the group was also calculated.

Pair-Wise Consistency values obtained from the research ranged from 51% to 65.4% for different pairs. Table 5.1 is a summary of the values.

Pair of Indexers	% Consistency Values
HN and LA	63.7
HN and TK	56
HN and RM	51
BO and LA	65.4
BO and TK	63.6
BO and RM	61

Table 5.1 Percent Average Consistency Values of all Pairs.

One-indexer Consistency with respect to the group of indexers yeilded values ranging from 28.65% for indexer RM to 47.5% for indexer BO. One-indexer Consistency in the system is on the decrease. Table 5.2 below is a summary of the values.

Indexers	% Consistency Values
HN	42.68
BO	47.50
LA	32.28
TK	29.25
RM	28.65

Table 5.2 Percent Consistency Values of One-indexer with respect to the group.

5.2 Inferences from Analysis:

The following can be infered from the results and analysis of the results of this study.

(a) Consistency values decreases with more indexers in a system. That is, the more hands we have in indexing, the lower the inter-indexer consistency values will be.

(b) Indexing Consistency mostly depends on individual differences. That is, human differences and differences in understanding the subject content of a document, largely determines consistency of the index. This can be spported by the inter-indexer consistency values obtained in the study.

(c) Indexing Consistency in the System is on the decrease due to the change in indexers. That is, every indexer comes to the system with his human differences and that affects the indexing and subsequently the effectiveness of the retrieval system.

(d) It can be infered that, the more materials one indexes, the higher will be his consistency value. This can be seen from table 3.1, where indexers that have indexed a large number of identical materials have higher consistency values for the pair.

(e) The use of thesaurus does not guarantee high inter-indexer consistency values within an information retrieval system.

(f) It is difficult to score consistency values more than the 80% reported in earlier studies by Hooper[1].

(g) One can also infer that indexing in-consistency is human and inevitable but can be minimized as pointed out by Huruitz[2].

(h) The inconsistency evidenced by the results of this study has led us not to reject our hypothesis.

(i) Subject background and experience determine indexing Consistency to a large extent.

5.3 Conclusions:

The results of the findings led us to the following Conclusions.

There is indexing inconsistency in the information retrieval system of IAR. This has made us not to reject our hypothesis which states that; consistency in the system varies with the change of indexers over a period.

We concluded that consistency values decrease with the change of indexers in the system.

The inconsistencies can be attributed partly to human differences, relevant subject background and experience of indexers.

The out come of this study corroborated the findings of earlier studies, that is, inconsistency is human and inevitable.

5.4 Recommendations:

In the course of this investigation some few problem areas have been identified. These problems are only indirectly related to the parameter measured in the study. They are however mentioned here so as to direct attension on to them and if possible future researchers can investigate into these areas.

1. Lack of strict adherance to indexing policy has been causing problem of inconsistency. Every indexer applies the policy to his convinience. There are many instances of over indexing and under indexing in the system. There is need for strict adherance to the laid down indexing policy in order to improve consistency.
2. The thesaurus should be revised. It was last revised in 1981 and since then many developments might have taken place and the thesaurus is suppose to be a dynamic tool.

3. Researchers should be allowed to use the information retrieval system without necessarily having the information specialists to assist where they can handle the search comfortablly.
4. The system should be open to users for as long as the library is open so that the practice whereby the library is opened but the information retrieval unit is not working will cease.
5. Enough hands should be employed to handle indexing of both current and backlog of materials so that the information stored in the system will be up to date and comprehensive. As a matter of policy in future employments, indexers with subject background relevant to IAR should be employed so that indexing consistency would be improved or be better than what is obtained now.
6. IAR should check its indexers and thesaurus for that decrease in inter-indexer consistency over the years.
7. Enough hands should be assigned to handle the patching of wrong holes, holes for withdrawn or lost documents etc. And there should be full time typists to handle the typing of document cards to speed up the process of input into the system.
8. IAR Library needs a generator and as a matter of urgency, a standby generator should be installed (provided) for the library so that the systems over-dependance on NEPA for electricity supply is minimized. The provision of a standby generator will enable the information retrieval unit to be in operation all the time when the library is open.

9. During the course of this study it has been discovered that holes for withdrawn documents or those punched by mistake are normally patched with a light piece of paper and many such pieces, used for patching the holes have been torn of due to constant use. To avoid this problem of patching materials being torn off, I would recommend that rubber solution or some other material that will be permanent be used in patching the holes.

5.5 Suggestions for Further Studies:

The following areas seem to be relevant to the subject of this study only that they have not been given adequate treatment or completely ignored in this study, therefore I would like to suggest that future studies should look into them.

Future studies should look into the quality of the indexing in the system and relate it to retrieval effectiveness of the system.

The thesaurus and how it affects indexing consistency also should be looked into. There is need to look into the relationship between consistency of the index and retrieval effectiveness of the system.

It would be adviseable if future studies would ascertain whether inter-indexer consistency is indicative of the quality of the indexing.

References:

1. Hooper, R. S. (1965). Indexer Consistency Tests - Origin, Measurements Results and Utilization. In: Zunde P. and M. E. Dexter (1969). "Indexing Consistency and Ouality." American Documentation, 20(3): pp. 259 - 267.

2. Huruitz, F. I. (1969). "A study of Indexer Consistency." American Documentation. (20(1): pp. 92 - 94.

BIBLIOGRAPHY

Afolabi, Michael (1985). Information Science: Its definition, evolution and place in Nigerian Departments of Library Science. Paper presented at the Department of Library Science Seminar, Jan. 23.

Berko, H. (1968). "Information Science: What is it?" American Documentation, 19(1): pp. 3 - 5.

Bradford, S. C. (1948). Documentation. Crosby Lockwood, London, p. 11.

Bryant, E. C. et al. (1963). Some Technical Notes on Coding Errors. In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

Case Western Reserve University Centre for Documentation and Communication Research (1968). An inquiry into testing of information Retrieval Systems Part I: In: Caudra C. A. (Ed.) Annual Review of Information Science and Technology - Encyclopedia Britanica Inc., Chicago, 4:pp. 81-82.

Cleverdon, C. W. (1962). Interim Report, Association of Special Libraries and Information Bureaux, Cranfield, England.

Cleverdon, C. W. (1962). Report on Testing and Analysis of an Investigation into the Comparative Efficiency of Indexing Systems. College of Aeronautics, Cranfield, England.

Conference on Training Information Science Specialists, Oct. 12 - 13, 1961 and April, 12 - 13, 1962. Proceedings. Institute of Technology, Atlanta, Georgia.

Cooper, Williams S. (1966). "Is Interindexer Consistency A Hobgoblin?" American Documentation, 20(3): pp. 268 - 278.

Costello, John C. Jr. (1966). Coordinate Indexing Volume VII. Rutgers University Press, New Jersey, 219p.

Fairtherne, R. A. (1958). "Automatic Retrieval of Recorded Information." Computer Journal, 1(1): pp. 36 - 42.

Fairtherne, Robert A. (1955). "Essentials for Document Retrieval." Special Libraries, 46:pp. 340 - 353.

Foskett, A. C. (1967). A guide to Personal Indexers. Clive Bingley, London, 80p.

Foskett, A. C. (1977). The Subject Approach to Information 3rd ed. Clive Bingley, London, P. 1.

Hooper, R. S. (1963). A Facet Analysis System, Automation and Scientific Communication. Short Papers, Annual Meeting of the American Documentation, Institute, 2:p. 253.

Hooper, R. S. (1963). "Facet Analysis System, Automation and Scientific Communication Short Paper." In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

Howerton, Paul W. (1965). "Organic and Functional Concepts of Authority Files." In: Simon M. Nemman (Ed.) Information Systems Compatibility. Spartan Books, Washington.

Huruitz, F. I. (1969). "A study of Indexer Consistency." American Documentation, 20(1): pp. 92 - 94.

Jacoby, J. (1962). Methodology for Indexer Reliability Tests. In: Zunde P. and Dexter, M. E. (1969) "Indexing Consistency and Quality." American Documentation, 20(3): pp. 259 - 267.

Jacoby, J. and V. Slamecka (1962). Indexer Consistency under Minimal Conditions. In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3): pp. 259 - 267.

Jahoda, G. (1963). "Indexing with Edge-notched and Internally punched Cards." In: Simonton, W. (Ed.) Information Retrieval Today, University of Minnesota, Minneapolis, pp. 37 - 47.

Kapple, Frederic J. (1939). "Looking Foward-A Fantasy." In: Danton, Emily Miller (Ed.) The Library Tommorrow, American Library Association, Chicago, p. 5.

Kent, Allen (1966). Textbook on Mechanized Information Retrieval, 2nd ed. Interscience Publishers, New York, 371p.

King, Donald W. (1968). "Desing and Evaluation of Information Systems." In:Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, Britanica Incorporated, Chicago, 3: p. 119.

Kondakov, I. P. (1966). "The Library and Documentation." Libri, 16(3): pp. 216 - 220.

Korotkin, A. L. and L. H. Oliver (1964). The Effect of Subject Matter Familiarity and the Use of an Indexing Aid upon Inte-Indexer Consistency. In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

Lancaster, F. Wilfred (1968). Evaluation of the MEDLARS demand Search Service. U.S. Department of Health, Education and Welfare. Public Health Service, National Library of Medicine, Bethesda. Md., 276p.

Lancaster, Wilfred F. (1968). Information Retrieval Systems: Characteristics, testing and evaluation, John Wiley, London, 222p.

Lancaster, F. W. (1969). "MEDLARS: Report on the Evaluation of its Operating Efficiency." American Documentation, 20(2): pp. 119 - 142.

Lancaster, F. W. and Jenkins, Grace (1970). "Quality Control Applied to the Operations of Large Information System." JASIS, 21(5): pp. 370 - 371.

Landau, Robert M. (1974). "Impact of Information Technology. In: Anthony, Debons (Ed.) Information Science: Search for Identity: Proceedings of the NATO Advanced Study Institute in Information Science: 1972, August 12 - 20; Champion, R. A. Marcel Dekker Inc., New York, pp. 191 - 196.

Lawani, S. M. (1976). "Indexing and Indexes." Nigerian Libraries, 12(2 and 3): pp. 111 - 128.

Leonard, Lawrence Edwards (1975). Inter-indexer Consistency and Retrieval Effectiveness: Measurement of relationships (Ph.D. Thesis). In: Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, Britanica Inc., Chicago, 12:pp.97-99.

Loukopoulos, Loukas (1966). "Indexing Problems and Some of their Solutions." American Documentation, 17(1): pp. 17 - 25.

Lufkin, Richard C. Determination and Analysis of some parameters affecting the subject indexing (Thesis B.S) In: Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, Britanica Inc., Chicago. 7:pp.

McColvin, (1978). Quoted in Borko, H. and Bernies (1978). Indexing Concepts and Methods. Academic Press, London, pp. 1 - 6.

Meadus, A. J. (1978). "Documentation and the Scientific Producer/User." Journal of Documentation, 34(4): pp. 324 - 332.

Mohrhardt, Foster E. (1966). "Librarianship and Documentation: Relationships in the United States." Libri. 16(3): pp. 211 - 215.

Momoh, Sidiku (1979). A Comparative Study of Readers Service in the Libraries of National Veterinary Research Institute, Vom and Institute of Agricultural Research, Unpublished BLS Project. Ahmadu Bello University Zaria, Nigeria.

Oruma, Ben Oviss (1978), The Information Retrieval System in Nigerian Agriculture (An Evaluative Study of the Feature Card System of the IAR, Samaru.) MLS Thesis. Ahmadu Bello University, Zaria, Nigeria.

Painter, A. F. (1963). An Analysis of Duplication and Consistency of Subject Indexing Involved in Report Handling at the Office of Technical Services, U.S. In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality," American Documentation, 20(3) pp. 259 - 267.

Ranaganathan, S. R. (1975). Five Laws of Library Science. Blunk, London.

Rayward, W. Boyd and Svenonius, Elaine (1967). Consistency Concensus sets and Random deletion. In: Caudra, C. A. (Ed). Annual Review of Information Science and Technology, Encyclopedia Britanica Inc., Chicago, 3: pp. 61 - 103.

Rodgers, D. J. (1961). A study of Intra?Indexer Consistency. In: Zunde P. and Dexter, M. E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3) pp. 259 - 267.

St. Lawrent, M. C. (1967). A Review of the literature of indexer Consistency. In: Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, Britanica Inc., Chicago. 3: p. 82.

Salton, G. and Young, C. S. (1973). "On the Specification of Term Values in Automatic Indexing." Journal of Documentation, 29(4): pp. 351 -372.

Saracevic, T. (1979). "An Essay on the Past and Future of Information Science Education:1. Historical Overview." Information Processing and Management, 15: pp. 1 - 15.

Savage, Terry R. (1965). "A Boyle's Law of Indexer Consistency." American Documentation, 16: p. 33.

Schultz, Claire K. (1976). "ASIS notes on its Founding i and Development." Bulletin of the American Society for Information Science, 2(8): pp. 49 - 51.

Shera, Jesse H. and Egan, M. E. (1950). "Documentation in the United State." American Documentation, 1(1): pp. 8 - 12.

Shera, Jesse H. (1951). "Documentation: Its Scope and Limitations." Library Quaterly, 21(1): pp. 13 - 26.

Shera, Jesse H. (1968). "The Propaedeutic of the New Librarianship." In: Simonton, W. (Ed.) Information Retrieval Today. University of Minneseta, Minneapolis, pp. 5 - 20.

Shera, Jesse H. (1968). "Of Librarianship, Documentation and Information Science." Unesco Bulletin for Libraries 22(2): pp. 58 - 65.

Shera, J. H. (1972). Foundation of Education for Librarianship. Becker and Hayes, New York, 293 p.

Shera, Jesse H. and Cleveland, D. B. (1977). "History and Foundations of Information Science." In: Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, Encyclopedia Britanica Inc., Chicago, 12: pp. 249 - 275.

Slamecka, V. and J. Jacoby (1963). Effect of Indexing Aids on the Reliability of Indexers, In: Zunde P. and Dexter, M. E. (1969). Indexing Consistency and Quality. American Documentation, 20(3) pp. 259 - 267.

Slamecka, V. and Jacoby, J. E. (1965). The Consistency of Human Indexing, in the coming of Age of Informatio: Technology, Documentation Incorperated, Washington D. C., pp. 32 - 56.

Stevens, M. E. (1965). Automatic Indexing: A State - of - the Art Report. In: Zunde P. and Dexter, M. E. (1969). Indexing Consistency and Quality. American Documentation, 20(3) pp. 259 - 267.

Swets, J. A. (1963). "Information Retrieval Systems." Science, 141: pp. 245 - 250.

Swets, John A. (1969). "Effectiveness of Information Retrieval Methods." American Documentation, 20(1): pp. 72 - 89.

Tague, Jean (1979). "Information Science in Graduate Library Programs." Canadian Library Journal, 36(3): pp. 89-99.

Tarr, D. and Borko, H. (1974). Factors Influencing Inter -Indexer Consistency. In: Zunde, Pranas (Ed.) Information Utilities: Proceedings of the 37th Annual Meeting of the ASIS, Volume II. Oct. 13 - 17, Atlanta G. A. Washington, D. C.

Taylor, Robert S. (1967). "Interface Between Librarianship and Information Science and Engineering." Special Libraries, 58: pp. 45 - 48.

Tinker, John F. (1966). "Imprecision in Meaning Measured by Inconsistency of Indexing." American Documentation, 17(2): pp. 96 - 102.

Van Rijsbergen, C. J. and Jones, Sparck, K. (1973). "A Test for Separation of Relevant and Non-Relevant Documents in Experimental Retrieval Collections." Journal of Documentation, 29(3): pp. 251 - 257.

Vickery, B. C. (1968). On Retrieval Systems Theory, 2nd ed. Butterworths, London, 191p.

Vickery, B. C. (1973). Information Systems. Butterworth, London, 350p.

Vickery, B. C. (1978). "Concepts of Documentation." Journal of Documentation, 34(4): pp. 279 -287.

Zull, Carolyn G. and Baumanis George J. (1967). Textual Correlates of relevance: Preliminary Investigation. In: Caudra, C. A. (Ed.) Annual Review of Information Science and Technology, Britanica Inc., Chicago. 3: pp. 84 - 85.

Zunde, Pranas and Dexter, Margaret E. (1969). "Indexing Consistency and Quality." American Documentation, 20(3): pp. 259 - 267.

APPENDIX I

TABLE FOR DATA COLLECTION

Author	
Title	
Ti	

APPENDIX II

FULL NAMES, QUALIFICATIONS AND YEARS SPENT BY VARIOUS INDEXERS IN THE SYSTEM

1. HN - Miss Helen Nation.
   She holds B.A. and BLS degrees and had working experience with information retrieval system before she was sent in by the Canadian University Services Overseas 1974 and left in 1978. The founder of the system.
2. BO - Mr. Ben Oviss Oruma.
   The first indigineous staff to join the system. He had B.Sc degree in Biological Sciences an area relevant to IAR. He worked there from 1974 to 1978.
3. LA - Mrs. Lamatu Adeoye.
   This indexer joined the Library as a Senior Library Assistant in 1967, and later obtained Diploma in Library Science. She started actual indexing in the system in 1979 after the two founding members of the system (Ben Oruma and Miss Helen Nation) had left, by the end of 1978. She indexed till 1981 and went for further studies. She came back with Associate degree in science in 1983 and continued work until 1984. She finally left at the end of 1984.
4. TK - Mrs. Tinu Komelafe.
   She had BLS degree and indexed in the system from October 1980 to August, 1983. She was the fourth indexer to join the system.

5. RM - Mrs. Rakiya Musa.

The fifth indexer to join the system. She had Diploma in Library Science and has been with the retrieval Unit Since 1982 up to date.