PN-AEE-226

64569

Working Paper No.

45

BIBLIOGRAPHIC INFORMATION MANAGEMENT USING MICROCOMPUTERS: THE EXPERIENCE OF THE INTERNATIONAL COUNCIL FOR RESEARCH IN AGROFORESTRY (ICRAF)

R. Labelle

January 1987

TABLE OF CONTENTS

	Page
ICRAF	
Institutional context	5
Documentation programme	5
Computerization	
Need for computers	6
Selection of computer hardware: micros vs larger computers.	6
Selection of microcomputer hardware	7
Selection of DBMS software	8
Human resources	9
ICRAF Library data base	9
Information retrieval	10
A typical query	11
Information retrieval with formatted output	12
Limitations of KMAN	12
Speed of retrieval	12
Access points	13
Indexed access	13
Conclusion	14

LIST OF TABLES

Table 1. Comparison of d-BASE II with KMAN.
Table 2. Structure of the ICRAF library data base.

LIST OF FIGURES

Figure 1. Preprogrammed mask used for data input and editing.
Figure 2. Legend explaining name and meaning of each field used in the library data base.
Figure 3. Formatted output on dot matrix printer of a search of the library data base.

APPENDICES

Appendix 1. Printout of programme to set input mask (LIBFORM.DEF).
Appendix 2. Programme to produce formatted output at the printer.

Summary

The Information & Documentation programme at ICRAF operates a information analysis centre on agroforestry that integrates the activities of the documentation service, the library and publications. This paper reviews the considerations leading to the decision to acquire microcomputers within the institutional context of this programme. The paper explains the approach taken, problems encountered and experience acquired in implementing a microcomputer based information management system for an international agricultural research centre located in a developing country. The author's views of the needed developments in microcomputer software for bibliographic information management are presented.

ICRAF : the institutional context

The International Council for Research in Agroforestry (ICRAF) promotes agroforestry, a term applied to all land use systems in which trees are deliberately combined with crops and/or animals Lundgren & Raintree (1983). Agroforestry has the potential for stabilizing and ameliorating agricultural production and for helping farmers and land managers in developing countries to become self sufficient. The Council is the acknowledged international centre of excellence in agroforestry.

In order to fulfill its mandate, the Council identifies,acquires,analyses and disseminates information and publishes research results on agroforestry. Within ICRAF, it is the Information and Documentation programme, known as Info/Doc, which is responsible for doing this.

Info/Doc is a specialized information analysis and disseminatior centre which provides a variety of information products and services. Info/Doc attempts to serve the information needs of a specified user community made up of researchers, development planners and others working or interested in agroforestry in developing countries. Priority in the provision of services is given to individuals or institutions working on field-related development/research projects in these countries.

Over the past four years, Info/Doc has directly answered over 500 separate queries for information on agroforestry.

Documentation programme.

The documentation activity provides a variety of services and products to the Council's target audience. At present, this activity includes a question and answer service, which involves locating and acquiring documents to meet user demands, including designing outside literature searches and corresponding with various literature services worldwide, evaluating documents received and undertaking subject analysis and input of bibliographic records to the computerized library catalogue, consulting with ICRAF's multidisciplinary team of specialists on specific questions posed by users and , finally, corresponding with these users. As part of this work, several abstract journals and external data bases are screened monthly for relevant literature . This screening is directly responsible for nearly all reprint acquisitions and results in over 100 titles received every month, mostly in reply to reprint request cards that are mailed to authors.

Other services include the preparation of specialized bibliographies, the annotation of bibliographies and the referral of users to other ICRAF specialized data bases to answer requests. Starting in 1986, a selective dissemination of information or SDI literature service will be offered to selected users and more bibliographies will be researched and produced. Reviews on selected topics in agroforestry are a natural outcome of bibliographic research and are being produced (Labelle,1986a).

Need for computers.

In acquiring bibliographic resources so that they could then be used to meet the current and future needs for literature on agroforestry, it became clear that an electronic data base management system (DBMS) could be useful. Using a DBMS, each

bibliographic unit or document could be described as one record in a file that could be updated as more bibliographic units were added to it. Each record would contain the normal bibliographic descriptors, i.e. author, title, publisher, keywords, etc., so that records containing specific data values for any of the descriptors could be readily retrieved.

Problems of management of a physical file or catalogue of documents would be decreased. There would be no problems of losing catalogue cards or of sorting the cards, etc. Producing and sorting indexes would be facilitated. Specialized bibliographies could also produced and even library catalogue cards.

Automation of library functions such as acquisitions, budgets, tracking book and periodical orders as well as circulation control could also be possible.

The next question concerned the computer environment that would support these tasks and wether they could be undertaken using microcomputers. In 1982 when this was being considered, microcomputers had begun to establish themselves in the marketplace.

Answers to these questions came with the realization that DBMS software existed for microcomputers that would be used to undertake these tasks, with a promise of rapidly evolving and increasingly available software and hardware innovations that would bring minicomputer sophistication and power to the microcomputer environment in the near future.

d-BASE II, a relational DBMS was tested initially and KMAN (Knowledge Man), a more sophisticated DBMS was acquired later.

However before considering the software, let us first examine the reasons for choosing microcomputers because these reasons may be relevant and useful to some institutions, especially those operating on a not for profit basis or those operating an information service in a developing country.

Selection of computer hardware: micros vs larger computers.

The choice of microcomputers over larger computer systems was based on the following considerations:

a- ease of use, microcomputers are easier to use than larger computers and user friendliness is one of the main considerations in their design.

b- cost, which is much less for microcomputers, both from the point of view of initial purchase cost and subsequent maintenance and staffing costs.

c- scale, ICRAF is a small institution, with modest administrative and scientific requirements for electronic data processing.By its nature, the Council does not undertake primary scientific research and thus has a limited need for data acquisition and statistical analysis usually undertaken on mainframe or minicomputers.

d- backlog of data, when the Council began its current programme of work in 1981, no significant backlog of data existed.

e- value of investment, the continuing and remarkable technical innovations that brought the microcomputer to individual users in the first place have established microcomputers as a business or work place productivity tool. This has created a demand that has brought down prices and which is likely to continue as more innovations are made. Micros are becoming standard office features, the micro business is booming, prices per unit of computing power are falling. The net result of these anticipated

innovations will be more powerful and user friendly systems that can meet the increasing demands for data base processing at ICRAF.All these reasons contribute to making the purchase of the more popular computers, the so called industry standards, a sound investment.

f- local service, microcomputers are serviced locally in Nairobi and there is no need to hire a service engineer to set up and operate the systems. Most microcomputers at ICRAF have been user installed and error checking is done by ICRAF technical staff, i.e. microcomputer programmers and operators. There are local sales and service representatives for the microcomputers installed by ICRAF.

The following example demonstrates that the Council is already benefiting from this choice.

Innovations in hardware and software are worthy of note. From 8 bit microcomputers running at 4.77 MHz and capable of addressing 64 Kb of RAM, the Council has migrated to 16 bit microcomputers capable of addressing up to 16 Mb of RAM while running at 8 MHz or more. Faster RAM memory for micros is also becoming available.

Software innovations are showing even greater promise with the arrival of sophisticated DBMS packages and multiuser multitasking operating systems from mini and mainframe computer environments. Operating systems capable of addressing more than the current limit of 32 megabytes of mass storage imposed by PC DOS 2.0 or 3.0 while permitting multiuser and multitasking facilities are now available, i.e. UNIX/XENIX, although applications software written for these systems in an IBM compatible microcomputer environment are not yet plentiful.Some of the advantages of these innovations will be discussed below. Understanding them is necessary in order to begin planning for DBMS implementations that are required to meet the growing needs of small to medium size data base applications, that is data bases ranging in size from a few hundred to over 64,000 records and possibly more, i.e. from less than 1 Mb to more than 64 Mb in size. This is within the anticipated range of capacities of DBMS applications envisaged at ICRAF.

These innovations also apply to the field of publishing. Typesetting and page layout can now be done using an Apple MacIntosh and Apple laser printer and associated software for under USD 12,000, CIF Nairobi. ICRAF has established the feasibility of producing documents using a word processor in the IBM PC environment (Multimate or Wordstar for example) and of subsequently transferring these to the Apple MacIntosh (Labelle, 1986b). The MacIntosh is superior for graphic intensive operations required for page layout and typesetting. Near typeset quality printing, which is suitable in many cases, can be produced using the Apple Laser Writer in connection to the MacIntosh. The output can optionally be set using a high resolution typesetting machine.

Selection of microcomputer hardware.

ICRAF uses Wang Professional Computers for word processing and IBM Personal Computers for data base management, although both machines can and do adequately perform these tasks, along with many others.

The IBM PC was selected for the Information & Documentation Programme because it is an industry standard, it conforms to an open architecture, it is very well documented, more software exists for the IBM PC than for any any other microcomputer.

A variety of high quality peripherals are available from third party manufacturers to improve the performance of the IBM product as it is. IBM is also locally represented. The IBM PC can largely be maintained in house on the basis of the technical information provided by IBM and using the Council's own computer programmer/ operators.

Inexpensive peripherals, including extra memory, can be added to a PC without any need to call for expensive servicing.Similarly, much information in the form of independent product reviews, commentary about the computer industry and advice for programmers and system operators is available for the IBM PC's and 100% compatibles in journals such as BYTE, Personal Computer Magazine and other periodicals. This advice can readily be acted upon by ICRAF staff without unduly relying on or being tied to local service agents and burdened with sometimes excessive service contracts. Of course , using this information implies that there are staff motivated to do so and that they are technically competent to evaluate it. To obtains such competence, ICRAF relies on locally employed computer professionals who have obtained their training overseas.

The biggest drawback of the IBM PC is its slow clock speed of 4.77 MHz compared to 8 MHz for the Wang PC or Olivetti M24 microcomputer and the 8 bit bus. The IBM PC AT has now overcome this problem.

Selection of DBMS software.

ICRAF uses DBMS software for several tasks, many of which are filing operations dealing with textual material and involving no mathematical manipulations. Bibliographic listings, results of surveys or questionnaires and mailing lists are some of the filing operations needed to be performed at ICRAF, and for which a microcomputer was acquired. Specifically, here are some of the tasks for which a DBMS can be considered:

* data input with error checking,
* data storage,data security and data manipulation, transformation,editing,
* selective retrieval and display of information,
* printed output in a user defined format.

KnowledgeMan or KMAN , a generalized DBMS produced by Micro Data Bases Systems, Inc. in the USA (1), has been used at ICRAF for data base management. It was selected because it supports up to 255 fields and over 64,000 characters per field or per record, total. It also fulfills the tasks mentioned above. One independent reviewer of note considers KMAN one of the best DBMS' for microcomputers (Pournelle, 1986).

These limits are beyond the needs of all present ICRAF applications. KMAN has the added advantage of permitting several data bases to be used at the same time , although the speed of multi file operations can be slowed down considerably when larger files are used simultaneously.

d-BASE II, a very popular commercially available DBMS was rejected because it did not support the large number of fields per record that are needed for the various ICRAF data bases.

d-BASE III, the successor to d-BASE III adapted to the 16 bit environment has not been evaluated.

A comparison of KMAN with d-BASE II and d-Base III is given in Table 1.

KMAN works on both the Wang and IBM PC microcomputers.

KMAN offers several features that permit a user friendly interface. Several variables can be user defined to configure and set up particular environments for given tasks, including menus. Macro expansion and a high level structured language are facilities for application development and user learning. The manual is written clearly and has been a useful learning tool for non programmers, such as the author , who can routinely undertake simple programming tasks using the KMAN command language. KMAN also has a built in spreadsheet which allows up to 255 rows by 255 columns (version 1.07 to 2.10) and can exchange data with data base records. A data base can also be treated as a spreadsheet under KMAN and analyzed as such, each field corresponding to a cell and vice versa.

Simple statistical analyses, calculations, functions and procedures for operating on numeric and string data also exist. Virtual fields and file encryption are also supported. Ad hoc enquiries, screen input/output, forms and printed forms management are also supported.

Very powerful data retrieval based on relational and Boolean logic operators can be performed using sequential searching techniques. A simple but fast indexing technique is implemented but is limited to retrieving character strings located at the beginning of indexed fields. (For more information on indexing, see below). Relational and Boolean operators cannot be efficiently used for indexed retrieval. Versatile sorting and other manipulations are also supported. Separate companion modules permit forms painting (K-Paint), text processing (K-Text), business graphics (K-Graph),mouse processing (K-Mouse) and a run time version (K-Run), (MDBS, 1985).

Human Resources.

Immaterial of what the advertisers of software packages or microcomputers have to say, user friendliness has not developed to the stage where complete novices can sit in front of a computer and begin to install,set up and master the system for immediate beneficial use. A learning period is be required. It is ICRAF's experience that the time and effort involved can be greatly reduced by hiring a computer programmer with at least a B.Sc. level of training in computer science. Even at this, learning is helpful in order to grasp the potential of the system and in order to better communicate with the programmer concerning the tasks that are to be performed.The programmer will need this feedback in order to properly design a system. This applies especially for software packages such as KMAN or d-BASE III because application development requires high level programming in the languages appropriate to these packages.

The ICRAF Library data base.

The ICRAF library data base is an electronic form of the library catalogue. Each record in the data base is equivalent to a catalogue card. In fact, the manual library catalogue can be produced in printed format from the library data base. The structure of this data base is shown in Table 2. The library data base was structured to store information according to general bibliographic rules that Info/Doc adapted to suit its own requirements along with the main categories of information that

the Council retrieves the stored information by. It is made up of 20 fields that contain the usual bibliographic categories along with subject content identifiers relevant to the interests of ICRAF. All together,the most important of these are the English and other language title fields of 200 characters in length each, the subject and species descriptor fields each 200 characters in length and the geographic region and country code fields. The latter two are but a few characters in length owing to the use of codes to represent the ecological zones and countries.

In Figure 1, the preprogrammed form or mask used for data input and editing shows these and the other fields that together make up an individual record in the library data base. The hard copy of this mask is known as an input sheet. A printout of the command file used to define the mask and to call it up on demand is given in Appendix 1. A menu driven system for data entry and editing can also be used. Figure 2 shows the name of each field and explains the meaning of each field and of each code used. Field 12 contains the two letter ISO (International Standards Organization) codes for geographical region and for countries. Field 11 is for the ecological zone codes. These are adapted from a modified Koppen classification as described by Young (1983). Fields 15, 18 and 19 are bibliographic descriptors and are used according to Morin-Labatut and Sly (1982). Each subject and species descriptor in fields 16 and 17 are delimited with "/". The reasons for this are described below.

All together each record occupies a little over 1 kilobyte of storage space. This large space is required immaterial of wether character strings are entered into the various fields because KMAN does not support variable length fields. Most records in the data base do not need all of the space allocated to them. Variable length fields only store what is entered in them. KMAN and other popular DBMS', including both d-BASE II and d-BASE III (2), do not support variable length fields. At present, the library data base contains over 7000 records and occupies 7.7 megabytes of hard disk storage. Of this, over 5 megabytes are used for empty parts of fields. Obviously, an improvement here would call for variable length fields or other storage conservation or data compression techniques.

Information retrieval.

One of the most important tasks a DBMS can perform is the selective retrieval of information from electronic files. Also important is the ability to output the information in a user defined format. Selective retrieval is performed using a combination of logical operators such as 'and', 'or', 'not', to link previously identified subsets of data obtained from files. There are therefore two distinct steps in selective retrieval; the creation of subsets from the data base and the linking of these subsets using logical operators sometimes known as Boolean operators.

Examples of selective retrieval under KMAN are given by the following simplified command:

CONVERT FROM LIBRARY FOR (conditions)

CONVERT FROM LIBRARY is the command to create a separate data base that will contain all of the records from the Library data base that are TRUE for a given set of conditions.

'FOR (conditions)' causes the command to act on a given record for which the conditions are TRUE. The conditions are the sequence of relational expressions that are evaluated for each record in the table as either TRUE or FALSE. An example of a relational expression can be all records for which the year of publication is 1978, or YEAR="1978" in KMAN language. Every record in the data base for which the year of publication is "1978" will be evaluated as TRUE and a subset of the main data base containing all records for YEAR="1978" will be created. Several other operators can be used here. They are equal to (EQ or =) as above, not equal to (NE or <>), greater than (GT or >), greater or equal to (GE or ≥), less than (LT or <), lesser or equal to (LE or ≤). A special operator, IN . is also used when searching for a particular character string in one or more record fields. An example would be a search of all records for those having the word 'tree ' in the English title or TITLEENG IN ["*TREE*"]. The asterisk '*' is a wildcard string match symbol that specifies that the character string 'tree' can be located anywhere in the field, either as a lone character string , or as part of a word, for example 'trees' or 'treed' or 'raintree' (a type of tree) would all result in a TRUE condition in this example.

The result of using these operators is the creation of a subset of records from the data base that are TRUE for the conditions specified, i.e. all those records that have the character string 'tree' somewhere in the English title field or those that were published in 1978.

In order to be even more specific in a search, the conditions can include the linking of subsets using the logical operators (also known as Boolean operators) AND, OR, NOT, XOR (and/or).

Combining the above two examples or subsets provides a clear illustration of this; that is searching for all records with 'tree' in the English title field and published in 1978. Each record satisfying these two conditions form another subset.

A typical query.

Under KMAN, the search criteria can be very specific. Consider the following query. An ICRAF library user seeks references in the literature to the use of Acacia trees (a genus widespread in the dry tropics) for fuelwood or charcoal in Kenya, Tanzania and Ethiopia. This researcher is interested in literature published after 1979 and written in English. A quick listing can be obtained using the SELECT command. Only a few of the fields specified are called for in the printout because a complete list would clutter up the printed output. The fields are chosen to provide the user with the most information required to decide wether he or she will consult the document itself. Using SELECT is therefore a quick way of providing the information requested for a researcher who has easy physical access to the ICRAF library.

The listing for this command is:

```
SELECT CURREC(LIBRARY), ACCESSNO, UDC, YEAR, AUTHORED, TITLEENG,
GEOGCOUN, ECOLZONE FROM LIBRARY FOR ( SPECIES IN ["*ACACIA*" ] AND
SUBJDESC IN ["*FUELW*","*CHARCOAL*" ] ) AND ( YEAR > "1979" AND
LANGUAGE = "EN" ) ORDER BY ASCENDING YEAR,AUTHORED
```

Brackets are used to indicate the precedence with which the operations are undertaken. Those in brackets are undertaken

first. Note that, as indicated above, the output is ordered by ascending year and then alphabetically in ascending order by author.

Information retrieval with formatted output.

The CONVERT command is an intermediate step in selective retrieval when the user wants to print the output in a predefined format. Figure 3 shows a printer output formatted with the KMAN program listed in Appendix 2.
Note that under KMAN/2, CONVERT produces another data base which is then printed with a user programmed command file that produces the output shown in Figure 3.

Limitations of KMAN.

Speed of retrieval.

Although KMAN allows powerful searching, it is relatively slow to identify and retrieve records meeting selective conditions, although KMAN/2 is much faster than the previous versions of the software. KMAN/2.1 is supposed to be even faster because all of the available memory, up to 640 Kb, can be used by the programme and not just the first 320 Kb.
Under KMAN/2, a simple search using the library data base, which contains over 7800 records, each a little over 1 kilobyte in size, takes about 10 minutes on the IBM PC AT before even producing formatted output at the printer.

Indexed vs sequential retrieval.

As mentioned previously, sequential searching proceeds one record at a time. It is a procedure analogous to searching for concepts or information in a book by starting from page 1 and reading to the end of the book. Each occurrence of the data sought is noted and the page(s) on which it is found is (are) kept in memory. Once the end is reached, all of the occurrences (page numbers) are listed. A data base index uses the same principle as a book index. It points to the record containing the information required or, as in the above analogy, each entry in the index lists the pages on which that entry is to be found. Like an index in a book, a data base index is a list of the searchable terms with the record number in which the term is to be found. In fact, a computerized index is another data base that is created for fast retrieval. The file created by indexing a data base is called an inverted file. Each searchable data element is listed alphabetically and the corresponding occurrences of the term, that is the record numbers, are listed alongside.
KMAN does have an index feature to speed up retrieval, but it does not support relational and logical manipulations. These operators can only be used for the sequential searching of each specified field (or fields) in each record of the data base. This is the case above with the SELECT and CONVERT commands. This limits the use of the index to predetermined key fields where the access points are located at the beginning of the field. Searching for character strings imbedded in the text of the field is not possible. This is a limitation of KMAN. As we shall see below, there are programmes that not only support these operations using indexes, but that also permit multiple access points to indexed terms that are imbedded in a field. The results

are discussed below and include dramatically improved access times.

Indexed access under KMAN is nevertheless useful for quickly obtaining all references produced by an author, as long as the author's name appears first in the field, i.e. first authors only. Thus the command to use an index in KMAN, 'PLUCK', can be used to quickly list all titles authored by "Lundgren" for example. The equivalent command using SELECT takes over 8 minutes to produce an output compared to less than five seconds to retrieve a record using an index. The other indexed fields in the library data base are the publisher (PUBLISHR) field and the accession number (ACCESSNO) field. The latter is a unique number assigned to each book and is a useful access point to an individual record.

Indexing is also used with another large data base, the mailing list. This data base contains over 4800 addresses. The index is used to retrieve records on the basis of an alphabetized field, called 'ALPHA' which contains the first 8 digits of a last name or institutional acronym. The index in this case is very useful for editing entries, although it cannot be used directly for this purpose, only to locate a record for subsequent editing. This is another drawback of KMAN.

Access points.

The KMAN indexing technique is fast, which is what an index is all about. Indexes are sometimes called pointers because they point the user to the required information. Unfortunately, under KMAN, it does not operate on imbedded text. This means it has limited access to data elements. Only the first value entered into a field can be retrieved using the KMAN index, as mentioned above.

Take for example the following string, which is a printout of the species field from a typical record in the library data base.

"ACACIA SEYAL/ACACIA SENEGAL/LEUCAENA LEUCOCEPHALA/ACACIA ALBIDA/
EUCALYPTUS CAMALDULENSIS/GMELINA ARBOREA/AZADIRACHTA INDICA/
CASSIA SIAMEA/PROSOPIS JULIFLORA/"

It is impossible to retrieve the character string ' ACACIA ALBIDA' using a KMAN index with the 'PLUCK' command. Access is limited to the leading characters of a field. The inherent advantage of an index, its speed of retrieval , cannot be used to advantage under KMAN because powerful relational and logical search capabilities are not supported with the index. The net result is that for all searches, on line methods are not permissible because they take too long for feedback to the operator sitting at the terminal. Batch processing is the only possibility. This limits user interaction and the possibility of quickly modifying search designs to either limit the number of records retrieved or to widen the breadth of the search, that is to either refine or broaden a search.

Indexed access.

Multiple access through the use of indexes that support Boolean logic operators is a very effective way of improving performance. Multiple access indexing is a technique whereby inverted files are created by marking with a delimiter (such as a /) or otherwise identifying those terms that are to be inverted or indexed (the same thing). An inverted list is a data base where

each entry corresponds to a data value , usually a character string, and to which is associated a list of all occurrences of that data value in the main data base. The following excerpt from the user manual of the 'Mini Micro version of CDC/ISIS' explains this well.

' The logical relationship between the major files of a CDS/ISIS data base is best perceived by examining the way in which retrieval is performed. Retrieval from a data base is done by specifying a set of access terms which are looked up in the inverted file to locate the list of master record numbers associated with each term. These lists are then manipulated by the program according to the Boolean logic operators you have specified in your search formulation until, at the end of the search, a single list , called the hit list, is obtained corresponding to the MFNs (Master file numbers or record numbers) satisfying your specifications. If at this point you request a display of the records answering the question CDS/ISIS will retrieve from the Master file each record in the hit list, format it, and display it on the screen.' (Unesco, 1985).

It is noteworthy that the increase in speed of retrieval is not all that dependent on the hardware. Even the slow IBM PC affords rapid retrieval using indexed access.

Furthermore,as the data base grows, and assuming that the list of descriptors used is finite , then beyond a certain data base size, the same descriptors will be used over and over again. Any increase in the size of the data base will only result in more entries under already existing descriptors. The inverted list will not grow much more in size. Thus access times, which are somewhat dependent on the size of the inverted list will not increase in direct proportion to increases in the size of the main file, as it would under conditions of sequential access. Of course this is only possible with conceptually related data sets. It is to this level of performance that Info/Doc hopes to arrive at.

Conclusion.

In conclusion, ICRAF has implemented a DBMS that.satisfies many of the present needs of the Council. However, more powerful and efficient software and hardware combinations are being sought. On the hardware side, Info/Doc has acquired the IBM PC AT microcomputer which offers much improved performance over the IBM PC (it is 2.5-3 times faster) while maintaining an open architecture and offering many peripherals and competitive prices. Info/Doc has acquired Micro CDS/ISIS from Unesco in order to benefit from the advantages of a sophisticated, proven, powerful , fast and user friendly software package for the manipulation of textual data and for the eventual standardization of data base formats.

The next improvements will come with new operating systems and DBMS software that will allow multiuser and multitasking facilities so that more than one user can have access to a file at the same time and more than one operation can be undertaken at once, i.e. sorting a file while searching on a second one with the same CPU. Once this happens, the distinction between microcomputers and larger computers may disappear and powerful computer based applications will be available to development information systems.

(1) Micro Data Base Systems, Inc (MDBS),
P.O.Box 248
Lafayette
Indiana 47902
USA ... are the producers of KMAN.

(2) d-BASE III supports a variable length comments field only. Data fields are of fixed length. When enquiring with MDBS, the producers of KMAN, about introducing variable length fields in updates of KMAN, the reply was that these would slow down file operations.

Bibliography.

Labelle (1982) A preliminary agroforestry word list with definitions. ICRAF Working Paper No. 8 .Nairobi, ICRAF., 30pp

Labelle, R. (1986a) Agroforestry: general concepts, early work and current initiatives. A review of the literature. ICRAF Working Paper Series. (In press)

Labelle, R. (1986b) Desk top publishing with microcomputers. Which system do we go for? Technical note. ICRAF, Nairobi. 6p.

Lundgren,B. & J.B.Raintree (1983) Sustained agroforestry. Reprinted from:Agricultural Research for Development:Potentials and Challenges in Asia. The Hague, ISNAR; Nairobi,ICRAF, p 1-26

Micro Data Base Systems Inc. (1984) Knowledge Man Reference Manual. The Knowledge Manager Software from MDBS. Sept. 1984. Version 1. 250 pp + Appendices & Glossary.

Micro Data Base Systems Inc. (1985) KnowledgeMan Reference Manual. The Universal Knowledge Management System. July 1985 version 2. Lafayette, Indiana.

Micro Data Base Systems Inc. (1986) KnowledgeMan Reference Manual. The Universal Knowledge Management System. KnowledgeMan version 2.01 Release Notes. July 1986. Lafayette, Indiana.

Morin-Labatut,G. & M.Sly (1982) Recommended Methods for Development -Information Systems; Volume 1.,Ottawa,Canada, IDRC.-TS40e, 272 pp

Pournelle, G. (1986) Computing at Chaos Manor. Byte vol 7 July 1986

Unesco (1985) CDS/ISIS. Mini-Micro Version. Paris, Unesco. 198p

Young,A. (1983) An Environmental Data Base for Agroforestry. Nairobi,ICRAF. 60 pp

Table 1

Comparison of d-BASE II with KMAN.

	d-BASE II	KMAN 2.0
Records per file or table	65535	2,147,483,647
Characters per record	1000	65,535
Fields per record	32	255
Characters per field	254	65,534

Sources: Micro Data Base Systems Inc. (1986)
Ashton Tate Inc. (1982)

Table 2

STRUCTURE OF THE ICRAF LIBRARY DATA BASE UNDER KMAN/2

Field Name	Type	Length	Comments
LOCATION	C	10	Physical location of document.
ACCESSNO	C	6	Alphanumeric accession number.
UDC	C	15	Decimal numeric of the Universal Decimal Classification. Used for books only at this time.
AUTHORED	C	35	Author(s) editor(s) of document.
TITLEENG	C	200	English title of document. If original is not in English, it is translated into English.
TITLEOTH	C	200	Title of document in original language.
PUBTITLE	C	100	Title of the publication, if journal or serial article.
YR	C	4	Year of publication.
VOLNO	C	16	Volume and number, if journal article.
PAGINAT	C	9	Pagination.
LANGUAGE	C	2	Language of publication.
NOREFS	C	8	Number of references in bibliography.
DATATYPE	C	9	Based on AGRIS bibliographic level.
RECOTYPE	C	2	Based on AGRIS.
LITINDI	C	7	Based on AGRIS.
GEOGCOUN	C	14	Two letter ISO region and country codes to identify document content based on geographic and country atributes.
ECOZONE	C	9	When indicated, codes of ecological zone based on modified koppen classification.
SPECIES	C	200	Descriptors of species used.
SUBJEDESC	C	200	Descriptors of subjects covered.
PUBLISHR	C	35	Publisher of documents.

Figure 1

Record Number:

LOCATION UDC CLASS ACCESSION NO.
ENGLISH TITLE

OTHER TITLE

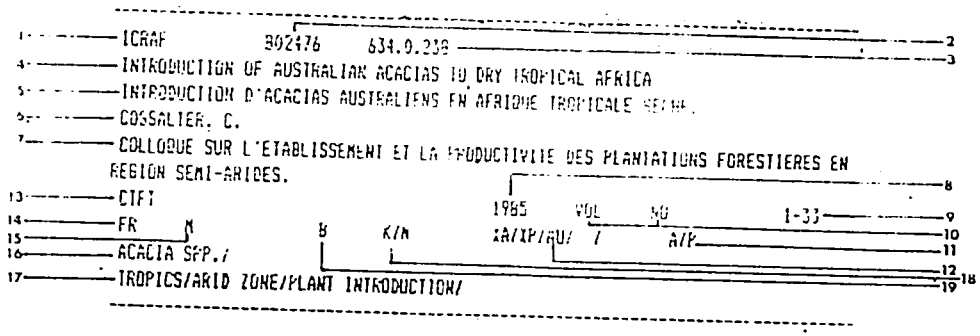
AUTHOR
AFFILIATION
PUBLICATION TITLE

PUBLISHER
YEAR 19 VOLUME NO VOL NO PAGINATION - NO OF REFS ref
sLANGUAGE DATA TYPE RECORD TYPE LITERARY INDICATOR
SPECIES

SUBJECT DESCRIPTOR

GEOGRAPHICAL COUNTRY / / / / ECOLOGICAL ZONE _____

Figure 2



1 - Location
2 - Accession No. (B ... for books)
3 - UDC Class No.
4 - English title
5 - Title in other language (original language)
6 - Author
7 - Publication title
8 - Year of publication
9 - Pagination
10 - Volume & No. in case of journal articles
11 - Ecological zone (see attached list of codes)
12 - Geographical location (either by country or region, see attached list of codes)
13 - Publisher
14 - Main language of the document
15 - Data type (using the AGRIS bibliographic level) e.g. A - analytical
M - monograph
S - serial

A document can be a combination of any two or three codes
e.g. AM - Analytical monograph - chapter of a book
AS - Analytical serial - journal article
AMS - Analytical monograph in series

16 - Main species in the document
17 - Key words - or subject descriptors
18 - Literary indicator (see attached list of codes)
19 - Record type - (AGRIS codes)

Literary indicator codes	Record type codes
K - Conference	B - Monograph
L - Dictionary	D - Drawing
N - Numeric data	F - Film
U - Thesis or dissertation	G - Map or atlas
W - Legislation	J - Serial article
Z - Bibliography	R - Report
Y - Map included	T - Computer medium or or output
E - Summary	
R - Review article	

Figure 3

```
!*****************************************************************************
!        LIB.ITB                          PAGE NO.   1            01/10/87  14:18:17*
!*****************************************************************************
-----------------------------------------------------------------------------
ICRAF          05400
ROLE OF SAMPLING FOR CROP-LOSS ASSESSMENT AND NEMATODE MANAGEMENT
HARRER, E.R. ET AL.
AGRIC. ECOSYSTEM ENVIRON.
                                        1985     VOL 12  NO          355-369
EN     AS            J      B/N        IN/US/  /  /       0
NEMATODES/PEST MANAGEMENT/CROP LOSS/HOST PLANT/PEST POPULATION/
-----------------------------------------------------------------------------
ICRAF          05401
ROLE OF THE HOST-ARBUSCULE, INTERFACE IN THE VA MYCORRHIZAL SYMBIOSIS: ULTRACYT
OLOGICAL STUDIES OF PROCESSES IN PHOSPHATE AND CARBOHYDRATE EXCHANGE
GIANINAZZI, S. ET AL.
PLANT AND SOIL
                                        1983     VOL 71  NO          211-215
EN     AS            J      E          XZ/  /  /  /       0
ALLIUM CEPA/ACER PSEUDOPLATANUS/GLOMUS MOSSEAE/
MYCORRHIZAE/SYMBIOSIS/FUNGUS/PHOSPHATE/POLYSACCHARIDES/PROTEINS/
-----------------------------------------------------------------------------
ICRAF          05402
THE EFFECT OF RAINDROP IMPACT AND SLOPE LENGTH ON SEDIMENT CONCENTRATION OF RUN
OFF
MANNERING, J.V. & MOLDENHAUER, W.C.
INDIANA ACADEMY OF SCIENCE: SOIL AND ATMOSPHERIC SCIENCES
                                        1983     VOL     NO          464-467
EN     AS            J                  IN/US/  /  /       0
RAIN/RUNOFF/SLOPE/EROSION/WATERSHEDS/
-----------------------------------------------------------------------------
ICRAF          05403
ASSESSING SOIL LOSS IN ZIAMBU AND MURANGA DISTRICTS, KENYA
LEWIS, L.A.
GEOGR. ANN.
                                        1985     VOL 67  NO 3-4      273-284
EN     AS            J      E/N        XA/KE/  /  /       0
SOIL LOSS/EROSION/POPULATION CHANGE/DATA COLLECTION/SOIL CONSERVATION/FOOD CROP
S/CASH CROPS/
-----------------------------------------------------------------------------
ICRAF          05404
THE ROLE OF THE TERMITE ODONTOTERMES GURDASPURENSIS (ISOPTERA: TERMITIDAE) IN P
LANT DECOMPOSITION IN A TROPICAL GRASSLAND
GUPTA, S.R. ET AL.
PEDOBIOLOGIA
                                        1981     VOL 22  NO          254-261
EN     AS            J                  XP/IN/  /  /       P
CHENOPODIUM ALBUM/DESMOSTACHYA BIPINNATA/
TERMITES/SOIL PHYSICOCHEMICAL PROPERTIES/TROPICS/GRAZING LANDS/DECOMPOSITION/
-----------------------------------------------------------------------------
ICRAF          05405
INFLUENCE OF FLORISTIC COMPOSITION ON THE NET PRIMARY PRODUCTION AND DRY MATTER
TURNOVER IN A TROPICAL GRASSLAND
GUPTA, S.R. & SINGH, J.S.
AUST. J. ECOL.
                                        1982     VOL 7   NO          363-374
EN     AS            J      E/N        XP/IN/  /  /       P
SESBANIA BISPINOSA/DESMOSTACHYA BIPINNATA/
PRIMARY PRODUCTIVITY/GRAZING LANDS/BIOMASS/SEASONS/TROPICS/
-----------------------------------------------------------------------------
        LIB.ITB
                                                                  201/10/8

                                                                14:18:30*
```

Appendix 1. The printout programme to set up the input mask (LIBFORM.DEF).

```
/* form for library.itb database */
form libform1
@ 3,2 put "LOCATION" WITH "R"
@ 3,11 get location
@ 3,22 put "UDC CLASS" WITH "R"
@ 3,32 get udc
@ 3,49 put "ACCESSION NO." WITH "R"
@ 3,62 GET ACCESSNO
@ 4,2 PUT "ENGLISH TITLE" WITH "R"
@ 4,16 GET TITLEENG
@ 7,2 PUT "OTHER TITLE" WITH "R"
@ 7,14 GET TITLEOTH
@ 10,2 PUT "AUTHOR" WITH "R"
@ 10,9 GET AUTHORED
@ 11,2 PUT "AFFILIATION" WITH "R"
@ 11,15 GET AFFILIAT
@ 12,2 PUT "PUBLICATION TITLE" WITH "R"
@ 12,20 GET PUBTITLE
@ 14,2 PUT "PUBLISHER" WITH "R"
@ 14,12 GET PUBLISHR
@ 15,2 PUT "YEAR" WITH "R"
@ 15,7 GET YEAR
@ 15,12 PUT "VOLUME NO" WITH "R"
@ 15,22 GET VOLNO
@ 15,40 PUT "PAGINATION" WITH "R"
@ 15,52 GET PAGINAT
@ 15,61 PUT "NO OF REFS" WITH "R"
@ 15,74 GET NOREFS
@ 16,2 PUT "LANGUAGE" WITH "R"
@ 16,11 GET LANGUAGE
@ 16,14 PUT "DATA TYPE" WITH "R"
@ 16,24 GET DATATYPE
@ 16,34 PUT "RECORD TYPE" WITH "R"
@ 16,46 GET RECORTYP
@ 16,50 PUT "LITERARY INDICATOR" WITH "R"
@ 16,70 GET LITINDI
@ 17,2 PUT "SPECIES" WITH "R"
@ 17,10 GET SPECIES
@ 20,2 PUT "SUBJECT DESCRIPTOR" WITH "R"
@ 20,21 GET SUBJDESC
@ 23,2 PUT "GEOGRAPHICAL COUNTRY" WITH "R"
@ 23,23 GET GEOGCOUN
@ 23,38 PUT "ECOLOGICAL ZONE" WITH "R"
@ 23,54 GET ECOLZONE
ENDFORM
```

Appendix 2. Programme to produce formatted output at the printer.

```
/* LIBFORM3.PRN KMAN FORM DEFINITION OF ACCESSION LIST FORM PRINTOUT*/
E.STEP=TRUE;E.OPRN=TRUE;E.PEJE="\\L";EJECT
E.PEJE="";E.DECI=0;MACRO EX EJECT
PAGENO=1
£A
FNAME=£DEFAULT
N=1
E.OPRN=FALSE
FORM TITL
        @ N,2 PUT "|*****************************************************
        @ N+1,2 PUT "|          "+FNAME+".ITB
        @ N+1,45 PUT "PAGE NO."
        @ N+1,54 PUT PAGENO USING "ddd"
        @ N+1,70 PUT £DATE
        @ N+1,80 PUT TIME()+"*"
        @ N+2,2 PUT "|*****************************************************
ENDFORM
FORM LIBLINE1

        @ N,2 GET LOCATION
        @ N,17 GET ACCESSNO
        @ N,28 GET UDC

ENDFORM
FORM LIBLINE2

        @ N,2 GET TITLEENG

ENDFORM
FORM LIBLINE0

        @ N,2 GET TITLEOTH

ENDFORM
FORM LIBLINE3

        @ N,2 GET AUTHORED

ENDFORM
FORM LIBLINE4

        @ N,2 GET PUBTITLE

ENDFORM
FORM LIBLINE5

        @ N,2 GET PUBLISHR
        @ N,42 GET YEAR
        @ N,51 GET VOLNO
        @ N,70 GET PAGINAT

ENDFORM
FORM LIBLINE6

        @ N,2 GET LANGUAGE
        @ N,9 GET DATATYPE
```

```
            @ N,23 GET RECORTYP
            @ N,30 GET LITINDI
            @ N,42 GET GEOGCOUN
            @ N,61 GET ECOLZONE

   ENDFORM
   FORM LIBLINE7

            @ N,2 GET SPECIES

   ENDFORM
   FORM LIBLINE8
            @ N,2 GET SUBJDESC

   ENDFORM
   FORM LIBLINE9
            @ N,1 PUT "--
   ENDFORM
   EX
   E.SUPD=TRUE
   CONTROL=1
   NOPRNTED=1
   PERFORM "INFODOC" ;
   E.PWID=100
   PRINT TITL WITH GET
   E.PWID=80
   PRINT LIBLINE9
   N=N+2
   LASTR=LASTREC(£B)
   WHILE CONTROL LE LASTR
            OBTAIN CONTROL
            PRINT LIBLINE1 WITH GET
            IF TITLEENG > 2 THEN PRINT LIBLINE2 WITH GET;ENDIF
            IF TITLEOTH > 2 THEN PRINT LIBLINE0 WITH GET;ENDIF
            IF AUTHORED > 2 THEN PRINT LIBLINE3 WITH GET;ENDIF
            IF PUBTITLE > 2 THEN PRINT LIBLINE4 WITH GET;ENDIF
            PRINT LIBLINE5 WITH GET
            PRINT LIBLINE6 WITH GET
            IF SPECIES >2 THEN PRINT LIBLINE7 WITH GET;ENDIF
            IF SUBJDESC > 2 THEN PRINT LIBLINE8 WITH GET;ENDIF
            PRINT LIBLINE9
            CONTROL=CONTROL+1
            NOPRNTED=NOPRNTED+1
            IF NOPRNTED=7 THEN EX;NOPRNTED=1;PAGENO=PAGENO+1;TALLY TITL;E.PWID=100:
              PRINT TITL;E.PWID=80;
              PRINT LIBLINE9;ENDIF
ENDWHILE
E.PEJE="\\L";EJECT;E.PEJE=""
RETURN
```