*Paper presented at
Training Course on Handling
and Dissemination of Agricultural
Information 12-23 August, 1985
at the International Livestock Centre for Africa,
Addis Ababa, Ethiopia*

MICROCOMPUTERS FOR INFORMATION CENTRES

SOME PRACTICAL CONSIDERATIONS

13393

BY

K.P. BROADBENT

JULY, 1985

504

Microcomputers for Information Centres: Some Practical Considerations

General

The reduced prices of microcomputers and their widespread availability in most countries now means that small information services and libraries can make use of them for increased productivity.

Whilst there are many good reasons for purchasing a microcomputer, there are pitfalls. There are now many texts and magazines on the market to assist would be purchasers of these machines. However, people in developing countries often do not have access to a lot of the brochures, advertizing material or specialist data and it is difficult for most people to understand and recognize the problems.

Firstly, it is important to understand what is meant by microcomputer because the all-embracing term "microcomputer" includes a very broad range of equipment and includes "home computers", "personal computers", "video game" machines, "micro-main frames", "small-business" computer compatible calculators, etc. The kind of system that could be considered by most people depends on the use, but a typical research department would need to consider a machine with at least 64K bytes of memory and two disk drives, a video display unit, keyboard and printer capable of printing lines of at least 80 characters. This would be what we call a single-user system but it cannot be expanded. However, the brand and configuration of the microcomputer should be determined with the software in mind, to ensure a functioning system. The main consideration for libraries and information services that need larger storage capacity is to realize that machines such as this will not be suitable as a long term solution because document collections tend to grow at a rate that often outstrips the computer's memory. Retrieval time will be inordinately slow, making the service little better than semi-mechanized methods.

The first point to consider, therefore, is to be **aware of the end-use.** If you purchase a microcomputer with the vague idea of "automating" the collection of documents or "assisting" in retrieval of literature, it is quite likely that you will be disappointed. You may end up purchasing a piece of hardware whose software capabilities are of little practical use.

It may not be capable of doing what you want it to do, or if it does serve your initial requirements, you will be disappointed later when you want to do a lot more things with it.

The process of choosing a microcomputer, therefore, should not be a hasty one. **Talk to other users.** Talk to those who have experience. Most countries now have specialist user groups who get together to share common problems and ideas. These user groups can be valuable sources of local information, and can often help you establish your system, troubleshoot and generally help.

For exchange of information, **compatiblity** is a major concern. Information comes in different packages and formats. If you plan to obtain data from a variety of sources and run different data bases, it is important to understand the need for compatible systems for efficient exchange of information.

For efficient use of any system, you will need **trained personnel.** Very often the only form of instruction available locally will be via the distributor or the instruction manual. The best person to operate the computer will need to be someone in your institution with an interest in computers. Someone without the necessary curiosity will probably not wish to explore enough to discover the full range of possibilities presented by the machine. The person also should have a logical and methodological manner to problem solving. Ultimately, you will want to develop computer skills among several staff members to maintain continuity.

The Question of Input or Data Entry

Suitable software for a micro for data entry of, say, the AGRIS type would be a text editor e.g., WORD-STAR or WORD-MASTER or any word processing system. The main point to consider here is that the software should allow for free text data entry with upper and lower case characters with an automatic skip to the next line when one line is completed. There should be

no limitation on the total number of charcters per data element because this
can range between 5 and 1000 characters. There should be provision for
corrections e.g., character or line deletion/insert/replacement; and there
should also be provision for the replacement of one character string by
another. This will obviate retyping of recurring data, e.g. journal title
or corporate authors. You will need a good, hard copy printer, so that
proofing can be done.

Data Output

Most people want a microcomputer so they can organize their files and obtain
faster information from their files, i.e., performing queries and serving
users better. However, the storage capacity of most micros is a limiting
factor that has to be recognized. Large data CAB, FSTA, AGRIS, etc. cannot
be stored on a microcomputer system. However, micros can be used
to **complement** the work done on larger computers.

There are two ways in which microcomputers will be of use to small
information centres:

- as intelligent terminals; and
- as a means of retrieving a set of documents selected by a broad query
from a large data base.

The intelligent terminal concept means using the microcomputer to store
log-on procedures, standard queries, SDI functions and pre-formulated
queries. The second usage, i.e., to process broad queries at a control
point enables you to reformat the result and store it on a floppy disk.
This type of data can either be received via telecommunication links or
through the mail on a floppy disk. Refined retrieval is then done on the
microcomputer by an information retrieval system, such as SEARCH and FIND
(SAF) developed by the Swedish Company, Program Markana. This is an
interactive retrieval system for bibliographic data. Input is done by
WORD-STAR text processing which SAF interfaces internally.

The system supports inverted file generation which is an integral part of the interactive dialogue. Retrieval is performed with a standard set of Booleau logic operators (AND; NOT; OR). The system inverts every word in the text excluding the stop words. It supports left, right, or middle word transactions in the search query formulation.

Important Terminology

Before embarking on a microcomputer system. It is important to understand some of the terminology.

Bits & Bytes

Computers are information processors. They differ from each other in the amount of information they can store, so we must have some way of quantifying information in order to make useful comparisons between them.

The simplest unit of information is a binary digit or bit. It is called binary because it can have one of two possible values - zero or one. Computers group them together into strings of 8 bits which are known as bytes. If bits are expressed as 0's and 1's, a byte might look like 10100010. There are 2^8 or 256 possible combinations of 8 bits, so a byte can have 256 possible values.

When a computer processes textual information, each 8-bit combination of 0's and 1's represents a character (01000001 is an A, for example). Bytes are equated with characters when storage capacities are expressed. For most purposes, an integer (or whole) number will require two bytes of storage and a real number will require four bytes.

Since a byte is a small unit of information, a shorter form of notation is commonly used. Microcomputer capacities are measured in kilobytes (Kbytes) and megabytes (Mbytes). From standard metric terminology you would expect 1K byte to be 1000 bytes. On computers, quantities are measured in powers of 2 and kilo was adopted to represent 1024 units. 64K bytes is actually 64*1024 or 65,536 bytes.

Operating Systems (OS)

All operating systems are written in machine language - the processor in
your computer will determine what operating system you can use. The three
most common operating systems that run on different computers are CP/M-80
(formerly known simply as CP/M), MSDOS (also known as PCDOS or IBM DOS), and
CP/M-86.

Memory

Most microcomputers contain two kinds of memory - ROM (Read-Only Memory) and
RAM (Random-Access Memory). RAM is the real working storage of your
computer. The amount of RAM a computer has is a very significant indication
of how powerful it is, how flexible it is, how easy it is to use, and how
large a task it can handle. It is more and more common these days to see
software packages that require 64K of RAM, and for machines like the Apple
III or IBM PC it is not unusual to find programs that require as much as
256K of RAM.

The amount of ROM a computer has is not as significant as the amount of RAM,
although there are some cases where a manufacturer's intelligent use of ROM
can be an advantage to you. A good example of this is the IBM PC, which has
most of its BASIC language sorted in ROM. This means that when you run a
BASIC program on this computer, some of the RAM that would otherwise have to
be used for the BASIC language is available to your program. An IBM PC with
64K of RAM can run a BASIC program over 41K in size, while a typical CP/M-80
computer with 64K of RAM is limited to about 30K for BASIC programs. Note
that because the ROM in the IBM is used for BASIC it does not give you any
advantage when running non-BASIC programs (like VisiCalc, Wordstar, etc.).

Floppy Disks

Floppy Disks are so-named because they are small circles of very flexible plastic. A floppy disk is often described as a cross between a cassette tape and a phonograph record. The disk is coated with a magnetic material. You can buy disks containing "prerecorded" programs or you can buy blank disks to store your own programs and data. The piece of hardware that reads and writes the disk is called a disk drive. There are two common sizes of floppy disks, 8 inches in diameter and 5 1/4" in diameter. Typically, a 5 1/4" diskette can hold about 350K bytes of data.

The hard-disk drive is a higher performance mass storage device known as a Winchester drive. The hard disk gets its name from the fact that the disk is made of polished aluminum. Hard disks cost much more ($1,000 to $15,000) but can hold from 10M bytes to 50M bytes of data.

There is another kind of hard-disk drive that is just appearing on the market. It is known by several names, including removable hard disk, exchangeable hard disk, or cartridge hard disk. In this design the disk itself is enclosed in a small plastic box which fits into a slot in the disk drive in much the same way that a floppy disk does. The disk drive prices are comparable to those for regular hard disks and the cartidges cost approximately $100.

These offer several advantages over fixed disks, by buying several cartridges, you can store tremendous amounts of data, without buying another drive. A cartridge gives you the same kind of flexibility as a floppy-disk drive does.

Computer-To-Computer Communications

When it is necessary to transfer files between two microcomputers or between a microcomputer and a mainframe, the preferred transfer method is floppy disk. Unfortunately, this requires that the two computers have a common disk format. If they do not, you must make an electrical communication link between the two machines and transfer the file over that link.

There are two common methods of doing this. Both methods require that the two machines have an RS232 serial interface as well as communications software. The simpler method can be used if the two computers are physically close to each other (in the same room, for example). It involves connecting a cable between the serial interfaces of the two machines and then sending the files via that cable. The second method can be used if the two machines are some distance from each other but each of them is close to a telephone. It requires two additional pieces of hardware called modems (an abbreviation for modulator/demodulator) and cables to attach each modem to the serial inteface in one of the computers. This method is almost alway used if one of the computers is a mainframe.

There are dozens of factors involved in communications, including switch settings, cable wiring, software settings, file formats, and operating system, all of which must be correct for the transfer to work properly.

Voltage Regulators and Uninterruptable Power Supplies

A voltage regulator will keep the voltage supplied to the computer within a specified range, even if the wall voltage varies considerably. Voltage regulators typically cost $300 to $500. Uninterruptable Power Supplies (UPS) will provide power for a few additional minutes after a power failure. This gives the user time to finish what he or she is doing, and allow for the computer to be shut down in an orderly fashion. A UPS, by design, usually also acts as a voltage regulator. Prices start from $500.

Bibliographic and Data Base Systems

There are many data base software packages available on the market today. As well, there are a few which are specifically designed or enhanced to handle textual or bibliograhic data. Common packages available are:

A Bibliographic Systems

Micro CAIRS: Leatherhead Foods R.A.
                            Randalls Rd.
                            Leatherhead, Surrey
                            KTZZ 7RY
                            U.K.

Inmajic   Warner-Eddison Assoc. Ltd.
                            186 Alewife Brook Parkway
                            Cambridge, MA 02138
                            U.S.A.

B Simplified File Management Systems

CARDBOX - Plus   Business Simulations Ltd.
                            Scriventon House
                            Speldhurst
                            Tunbridge Wells, Kent
                            TN3 OTU
                            U.K.

C Generalized Data Base Systems

dBASE II and dBASE III   Ashton-Tate
                          10150 West Jefferson Blvd..
                          Culver City, CA 90230
                          U.S.A.

C <u>Generalized Data Base Systems</u> (cont'd)

SIRE   KNM Inc.
                        6118 Swausea Street
                        Bethesda, MD 20817
                        U.S.A.

Knowledgeman   Micro Data Base Systems, Inc.
                        P. O. Box 248
                        Lafayette, IN
                        47902
                        U.S.A.

Revelation   COSMOS Inc.
                        19530 Pacific Highway S.
                        Seattle, WA
                        U.S.A.

Inverted File

For library systems to be at all useful, the data that is likely to be referenced directly (titles, authors, key descriptor words, etc.) need to be stored in a computer in such a way that they can be found quickly. To accomplish this, the data to be referenced is extracted from the primary data base, and stored in a separate file called an **inverted file.** In the inverted file, the data is organized differently, being sequenced alphabetically for rapid access, and containing the appropriate identifiers to the data records in the primary data base.

Of course, data not inverted in this way can still be searched. However, the complete file would have to be browsed to find the records in question.

Typical Prices

Examples of typical small (and very small) business computer systems and their prices

CSI group	Maker and model	Typical Price
Group I	Apple II	$1,750-6,500
	Commodore 8000 Series	$3,100-5,300
	Commodore 4000 Series	$2,800-3,700
	Tandy TRS-80 Model II	$4,100
	Olympia Boss	$5,500
Group IIa	Sirius I	$5,500+
	Olivetti M20	$5,000
	Apple III	$5,300+
Group IIb	DDC Profi 203	$8,500+
	Nixdorf 8820/2	$12,000+
	Kontron PSI8000	$15,500+
Group III	Nixdorf 8870/1-Model 10	$22,000
	IBM System 34	$20,500+
	Digital PDP 11/23	$33,000+
	Kienzie 9066-10	$23,500+
	Olivetti S6000	$40,000
Group IV	IBM System 34 (F36)	$91,000
	Digital PDP 11/24	$46,000+
	Philips P4530	$54,000
	Nixdorf 8870/8	$100,000