Marine Science Information: an International Commodity
(eds) Grundy, Ford & Beardsley
IAMSLIC. 1986.

THE COMING REVOLUTION IN LIBRARY SCIENCE

Thomas J. Benner
24314 Norchester Way
Spring, Texas 77811

ABSTRACT

The technology in the computer world is changing so rapidly nowadays, that innovations are viewed as antiquated systems in only a few years. There are even more changes coming up in the next few years to make computing more accessible and rapid than ever before.

INTRODUCTION

It wasn't so many years ago, that I, a young undergraduate student in Arizona found myself listening to an "Electronic Calculator" salesman who was a guest lecturer in an Introduction to Engineering class midway through one semester. The lecture was particularly amusing to me, since I had just invested more than $40.00 (a sum roughly equivalent to a king's ransom in those days) in a new K&E slide rule that must have had a jillion scales on it. Now, there stood before me a "salesman" with this bulky looking plastic box in his hand. The "miracle of modern electronics" as he called it, was only capable of adding, subtracting, multiplying and dividing! You couldn't get a sine nor even a square root from it and if you wanted to use logarithms, you still had to refer to the tables. Worst of all, that ugly little box with the red numbers cost more than $300.00! Imagine, I thought, anyone silly enough to trade in his trusted K&E Slide Rule (complete with real leather carrying case and built in belt loop) for such a frivolous and limited novelty item. To me and most of my classmates even the thought of such an act was absurd.

FUTURE COMPUTING

This all occurred in 1971. How was I to know that silly little box would within the next few years fall in price so dramatically, that today the modern pocket calculator (with more functions than my old K&E slide rule ever dreamed of) can be purchased nearly everywhere for less than $10.00? Even more astounding is that our more advanced

hand calculators have more power than computers which filled entire buildings only 25 short years ago! Isn't it amazing? Imagine, the computers that the majority of our children play with, the very machines that seem almost mundane to the children of the 80's would not have been thought possible barely 10 years ago! There truly has been a revolution going on all around us in computer technology. A real revolution, which has resulted in miraculous achievements. A revolution which has spawned a new generation in computer hardware just about every 3 to 4 years! Isn't that truly amazing?

Would you like to hear something just as amazing as the revolution in computer technology? There is a revolution beginning to brew that will effect every Librarian or Records Manager, from the most hallowed largest universities right down to the small local neighborhood library next door. Its coming, in fact, it's already started. It will affect every aspect of what a librarian does. Some of you will love it, other will hate it, but come it will whatever your personal desire.[1] Words like **DataBase, CRT, CPU, RAM, ROM, Batch Process, Multi-User, Interactive** and many more will become a permanent part of your vocabulary. Best of all, there is no need for any of you to be afraid. The revolution will be painless and will open up vast new applications for your skills.[2] For with forethought and planning on all of our parts, this revolution will enable a fantastic amount of knowledge to be accessed and shared as never before. It is for this reason that all of us must understand what has and is transpiring to make this revolution possible; and hopefully can help those of us who are bridging the gap between computerized data manipulation and manual techniques to do so in a way that will give you the best possible combination of advanced computer hardware and software.

NEW APPROACHES

The computerization of many library functions certainly is not a totally new field. There are several programs available to "computerize" a select group of library functions utilizing what we in the computer industry commonly refer to as "Main Frame" computers. There are also a fewer number of programs available for another class of computers which many of you probably have more direct experience with, the Personal Computer or "PC" as it is commonly called. Not too many years ago there was a very distinct difference between computer classes and in order to fully appreciate what has happened to the computer industry, it is important that you understand the differences between these traditional categories of computers.

In the beginning there were:

Main Frames. These physically large machines (with corresponding large price tags) dominated the computer industry for decades. They were until recently the most powerful of all computers, having the ability to process millions of instructions per second (MIPS). They were and are characterized as requiring special environments (cooling

& humidity), large staffs of computer professionals (to keep them running), operate primarily in the batch mode (that is jobs are submitted to the processor in groups via card decks or CRT's where they are processed without human intervention or interaction__ and use a Time Shared architecture (more about that later).

Mini Computers. Increasingly referred to as Super Mini's these have some of the attributes of the Main Frames, in that they require special environments and utilize Time Shared architectures, but unlike the mainframes, these machines operate in both the Interactive (that is a human interacts with the applications program being run by the machine as well as batch modes. Further, these machines while not as powerful as the Mainframe, cost considerably less and require substantially fewer computer professionals in order to keep them running. Many users believe that this class of computer is easier to use as well (being more "user friendly").

Micro's. (also called PC's)They have been traditionally small (physically) machines capable of single user access only with limited amounts of memory and speed available. However, these machines usually require no special environmental consideration, are easy to use and relatively inexpensive. Where a mainframe would typically operate at 2 to 50 MIPS, a Mini would operate at .5 to 4 MIPS.[3] (data resources)

All of the above used to be a reasonably accurate description of the various categories of computers, that is **used** to be! Today, Mini's have MIPS rates which far surpass some Main Frames, and some Micro's outperform both! Size is of little meaning anymore as is cost, since for some applications the less expensive Micro does a better job than the more expensive mainframe. As if that isn't enough, it is now necessary to add two new categories of computers:

Super Computers. These machines are the fastest computers available, They have MIPS rates of 90 or better and use a combination of the traditional time shared architecture and vectored processing. Because of their brute force speed, Supers have rigid environmental requirements, required scores of technicians and of course if you have to ask, you can't afford one!

Multicomputers. Lastly, but certainly not least, enter the Multicomputers. Unlike all of the computers discussed thus far, the MultiComputer does not have a Time Shared architecture. Rather, it literally ties together a multitude of processors into one tightly knit group or network, such that each task or user has his/her own processor. In order to fully appreciate the significant benefits available in a Multicomputer, it is necessary to understand the differences between a Multicomputer and traditional multiuser computers. Until now, multiuser computers have utilized a "time-shared" design in which there is one processor which is shared by all users on the system. To illustrate the time-shared concept, imagine a clock with a ticking second hand. This second hand represents the computer's attention as it is divided among the users. When the hand stops at each second, the computer devotes 100% of its attention to the user associated with that second's position, and the other users

associated with the other 59 seconds are idle, waiting their turn or share of the computer's time. Using this time sharing concept or architecture, as more users are added to the system, it becomes slower because its capacity is divided or "shared" by more users.

A Multicomputer does not use the time shared architecture, instead it uses a "Master/User" architecture. A multicomputer is not one computer but a system of tightly coupled processors sharing certain peripherals such as printer, disk drives, plotters, etc. Each user has a dedicated processor ("User" or user processor) while a Master processor manages the User processors and their access to the shared resources. The net result of the Multicomputer architecture is that each user gets his own computer. Unlike traditional time-shred computers, a Multicomputer can be expanded almost indefinitely with little or no degradation in performance.

Time shared computers are limited in the number of users they can support due to lack of the processor's ability to handle a greater load and lack of physical attachment points for users. For example, a time-shared computer which can physically attach 32 users can rarely accommodate more than 25 users without suffering serious performance degradation.

A Multicomputer however, is almost unlimited in expansion capacity. Since each user on a Multicomputer has their own processor, additional users can be added to the system by simply inserting a processor board (User processor) and adding a terminal. In situations where the number of users (User processors) needed exceeds the capacity of the cabinet, additional cabinets can be "linked" together without limiting a user's access to shared resources or communicating with other users. In this way the Multicomputer System can be expanded to fit the needs of the users, rather than forcing the user to purchase newer, bigger, more expensive equipment each time capabilities need to be increased.[4]

Now that we understand the various differences between computer hardware types, let us explore the differences in Software. We shall define "software" as the instruction set which is utilized by the computer hardware in order to accomplish a specific task or function. Because software is rather dependent upon the available computer hardware, it typically lags behind hardware development, sometimes by years. Also since software is to some degree hardware dependent, there is usually a direct relationship between the cost of software and the cost of the hardware that it requires to run on. That is,a program which runs only on multi=million dollar equipment usually sells (or leases) for many thousands of dollars (sometimes hundreds of thousands). Correspondingly, programs which run on inexpensive Micro's rarely cost more than a few hundred dollars. Although, increasingly, programs which could only be run on mainframes are now finding their way down to Micro's and Mini's; usually with corresponding reductions in price.

It is not possible for software companies to convert all mainframe software to run on PC's. Frankly, some of the software that is being

run on main frames is so antiquated and hard to use, that no PC owner would buy it anyway. However, as PC's become ever more powerful I'm sure that we will see more and more main frame type software being converted to take advantage of the more powerful PC;s and the much larger PC market. But more importantly, as the PC becomes more powerful, and as the Multicomputer becomes more commonly available, software companies will increasingly take advantage of these factors to create increasingly more sophisticated programs. Also, since there is a far greater percentage of software being developed today for PC's and Multicomputers, you the end user will begin to see software that is not only very user friendly, but more powerful than any you have been accustomed to even on the main frames of just a few years ago.

Since the software companies who are developing programs for PC's are depending upon selling their products in this very large market place, they are more willing to amortize the cost of the development of these programs over more installations. As opposed to software developed primarily for mainframes which because of the much smaller number of installations, must be amortized over a very few systems (not uncommonly, only one). The net result to you the end user is that you will have available to you programs which are not only very sophisticated, but cost less than what you would have expected to pay for the same program which was developed to run on a mainframe only a few years ago.

My good friend Ruth Grundy, Senior Librarian at the University of Texas Marine Science Institute Library in Port Aransas, Texas believes that computer programs should and could now be developed to perform the following functions in most Libraries:

Journal Control
Journal Exchange Control
Circulation management
On Line Card Catalog (including search capabilities)
Interlibrary loan generation and accounting
Acquisition Automation and Control
Inventory Control
General Accounting
Electronic Mail
Reprint & Technical Report/Government Document data base
Building off line BRS & DIALOG search strategies
Records Control

There is no doubt that Ruth is correct, in that for the first time in history we now have both the hardware and software technologies necessary to accomplish all of the above tasks in both a cost effective and user friendly manner. In my mind the dilemma is which of the tasks on this long list do we begin with? Keeping in mind that the only realistic way to effectively satisfy all of the above areas will be through the creation of a large relational data base, such that as new

tasks (or problem solvers) are added to the program the end result will be a coordinated and effective solution. In other words, we need to treat all of the individual problems as integral portions of a greater whole (which a fully relational data base allows us to do). In my opinion this is the only way that the computerization of a library can be accomplished in a meaningful manner.

What does all of this mean to you, the Librarian? How will these advances in technology effect your profession? What will it cost to get started with a state of the art library computerization project? What new set of problems will arise out of these new advances? These are just some of the many questions which are being asked about the revolution which is going on about us. The remainder of this paper will be devoted to answering them as best I can.

The Library of 1995 will be very different from what most of you are accustomed to today. The card catalog, and technical reports section will have pretty much been replaced by CRT's and possibly even voice actuated information stations. Your job as librarian will be in part to oversee the input of the various documents for which you are entrusted. In the larger libraries, you will undoubtedly be well into a program of inputting old technical reports and other here-to-fore uncataloged information. Some of you will become specialists in search programs, which have been written to access the phenomenal amounts of data being entered into the system. Others of you will become communications coordinators who will be responsible for maintaining and enhancing the intralibrary communications network, where nearly all major libraries computer systems are interconnected into one big information network. Using this network, you will be able to access virtually any document contained in any library on the system from your desk in less time than you can now walk out to the stacks and retrieve a book!

One of the greatest problems facing your profession over the next few years will be the coordination of data input and information sharing between libraries. There now exists the hardware and there will soon exist the software necessary to maintain online the title, author, key words and the abstract of every document in every library. Once this wealth of information is stored in the system,it could be accessible via boolean logic searches by any library user who had been given access to the system! The physical location of the document or even the computer will be immaterial! A user at Woods Hole could be given access to the entire collection of Scripps, without ever leaving his office in Massachusetts! Further, he will even be able to get paper copies of this information printed at his location!

However, in order for this kind of information access to take place, there must be a coordinated effort on the part of the Library Profession for the rational input of the millions of documents currently being held in your libraries. It makes little sense for those documents which are being held in more than one library to be entered into the system more than once. There will have to be a coordinated effort combined with the willingness to share entered data in order for your profession

to truly realize the benefits of these advances in technology. I am told that some of you have literally millions of documents in storage at your libraries which have never been cataloged, and which some of you have virtually locked away in closets out of the frustration of being unable to handle such a massive volume of information with your limited staffs and budgets. Isn't it a shame that such a wealth of information has been lost for all practical purposes. Who knows what discoveries lay hidden away on your shelves? Someday, it will be available to anyone who wants it. But, we all have to work together toward this common goal, if we are to save the information with which you have been entrusted.

Those of you who like I, have received their indoctrination to computer via the mainframes have probably by now concluded that the cost to automate your libraries can only be measured by the millions of dollars. Your estimate would be quite correct if this modernization were to take place with Main Frames or even Mini computers. Obviously, even the smaller libraries will need very powerful machines if they are to fully realize the many benefits outlined above. If the older time shared technology was used to accomplish this task, each library would be forced to buy a very powerful processor which would be capable of handling this Herculean task. Such a machine, even one which was capable of taking advantage of some of the newer advances in computer technology could easily cost more than a million dollars and in larger institutions as much as ten times that number. My personal opinion is that a good average for a modest 100 user system would easily be two to three million dollars. Even more discouraging than the numbers themselves is that each Library would be forced to fund the bulk of this amount at one time, since for the most part time shared machines are not easily expandable as I have explained earlier. It is true that CRT's can be added to these systems with relative ease and for only a few thousand dollars per CRT; however, if the computer itself is not sufficiently large to handle the additional burden of ever increasing users, its performance will degrade so rapidly so as to make it functionally unusable in a real time environment. Of course, when this happens, the entire system must be replaced with yet a larger system which will cost substantially more than the original system. Worst of all, the now inadequate smaller system probably won't be usable or even compatible with the new system. It is not uncommon for used computer systems to have lost 80 to 90 percent of their original value after only a year or two of service. In other words, that $1,000,000 system which you bought two years ago, might only fetch $200,000 or less today!

On the other hand if a multicomputer was used to modernize your library it would be possible to start with a small basic system which met your existing budget. Then as you were able to secure additional funding for the system it could be made to grow to accommodate both your needs and your budget. Let's say that a large University Library determines that it needs to accommodate its user community which is spread between a main library, a science library and a marine library

as well as several auxiliary libraries spread over the main campus and two extension campuses containing a total of 2 million titles. All together, the computer system would need at least 300 CRT's with the potential to add at least 100 as money became available. A mainframe system to accommodate such a configuration would need at least 10 megabytes of RAM memory (although 50 MB would not be too much). Further, the system should have at least 3 giga (a giga is equal to 1,000,000,000) bytes of disk storage capabilities. Using a mainframe system (a mini or super mini couldn't handle the load) with probably less than 20 MB of RAM would require a budget of at least $6,000,000. The Multicomputer system would include 122.5 MB of RAM memory, 1.66 giga bytes of disk space 300 CRT's and 24 printers (3 Laser, 3 daisy wheel, 12 high speed dot matrix and 6 letter quality dot matrix).

If the university in our example were unable to secure the $6,000,000 funding for the mainframe they would have no viable alternatives. However, were they attempting to secure the $2,000,000 funding for a Multicomputer and were turned down, they could start with whatever funding they could secure and add the initial system as funding became available until they had achieved the complete 300 user system that they needed. Further, as other departments on the campus desired to access the libraries system, the modularity of the multicomputer would allow each department to add to the multicomputer out of their own budget, yet in such a way as to benefit the entire library system as well.

Whether the university in question chose to begin its Multicomputer system with the full $2,000,000 or $100,000 would only effect the number of CRT's, printers and disk drives available to the user community, not the performance of the system, or its capabilites. Again, as the University was able to add to the system they could do so without losing any investment in either their hardware or software, since their initial purchases would become an integral part of subsequent additions to the system.

Now that you understand how nearly every library will be able to afford its own multicomputer system, it is easy to see how it is inevitable that over the next few years libraries will join the age of computerization, which leads us to yet another set of problems to be addressed. As more and more libraries adopt the multicomputer type architecture ever increasing amounts of data will become available to larger and larger user communities. Who (if anyone) should decide which users should have what access to what information? Should any user of any library be empowered to perform virtually any search of any topic he/she wishes? If access is to be denied, who shall decide who has access and who shall not? With the computers built-in ability to track usage, shall users be charged for library access based upon their usage of library facilities? Should an effort be made to charge for the systems use (somewhat like we are now charged to access DIALOG, BRS, etc.)? By charging for usage can we expect the library user community to fully fund all library services? As you can see the advent of the com-

puterization of libraries will open an entirely new perspective and related questions regarding the library of tomorrow.

CONCLUSION

In closing, I'd like to propose a solution to one last problem here-to-fore unmentioned; namely, what is to be done with all of those wonderful solid oak card catalog cabinets once they have been replaced by CRT's. I've pondered this problem for some time, and believe that I have found a solution, in part due to my K&E slide rule (mentioned earlier). Being somewhat frugal in nature, I vowed many years ago to somehow recoup my investment in the K&E rule. In order to accomplish this seemingly impossible task, I've locked that old K&E away in my safety deposit box along with my other invaluable possessions. There it will stay, year after year, until someday there are no K&E's around. Then, I'll sell it through an antique dealer as a novel curiosity of the pre-calculator era. If I hold onto it long enough I'll bet that I get my $40 back with luck, maybe even more! If I can do it with my K&E, you can do it with your old card catalog files. In fact, you might not have to wait as long as I will, since there have to be lots of uses for those wonderful old masterpieces. They would make great storage places for small nuts and bolts that are always cluttering my garage. What about as a file cabinet for small sewing supplies, or recipes? (Scratch recipes, there'll be a computer in every kitchen for that sort of thing soon, but that's a whole other subject worthy of its own paper.)