

ARCHIVAL INFORMATICS NEWSLETTER

Part 1 of Archival Informatics Newsletter & Technical Reports

ISSN 0892-2179

WINTER, 1987

Volume 1, #4

ANOTHER YEAR

With this issue, the Archival Informatics Newsletter completes its first full year of publication. I'm pleased to report that next year it, and the Archival Informatics Technical Reports, will be available not only through Archives and Museum Informatics, but also through the Society of American Archivists in Chicago and the Museum Documentation Association in Cambridge, England. Individual members of SAA are now entitled to a 10% discount on subscription prices whether they renew through Archives and Museum Informatics or sign up through SAA. All European orders will be handled through the MDA from now on. I am extremely pleased to have the SAA and MDA signing up subscribers and distributing publications. Along with the mail I receive, this is further testimony that the Newsletter and Technical Reports are serving a useful purpose.

In 1988, the Archival Informatics Technical Reports will include:

- * A Directory of Software and Systems for Archives and Museums
- * Proceedings of a Seminar on Authority Control
- * A Case Study of Appraising Online Governmental Records Systems with discussion papers, and
- * Functional Requirements for Exhibit Management Systems.

The Newsletter will continue to have the same current awareness content it has had and hopes to include software reviews of new systems scheduled for release in 1988 by IME and the MDA (TINmus), Vernon Systems (Collection), and Cactus Software as well as others as they appear.

As an inducement to prompt renewal, an index to the 1987 Archival Informatics Newsletter will be sent to those who renew for 1988 or purchase back issues as part of a new subscription in 1988.

I want to thank those who have contributed to the publication this year, and especially thank Tom Brown whose "Machine-Readable Views" column has been consistently stimulating. Tom has agreed to continue his column in 1988.

TABLE OF CONTENTS

Articles:

MARCON: A Software Review	p.66
David Bearman	
Machine Readable Views	p.70
Thomas E. Brown	
Editorial Commentary	p.71
SAA-CGAP Planning Group Report	p.72

Regular Features:

Conferences	
Museum Computer Network	p.74
Am. Soc. for Info. Science	p.75
Authority Control Seminar	p.76
Conference Calendar: 1988	p.77
In-Box	p.78
Projects & Proposals	p.82
Software Forum	p.84
Software Products	p.86
Standards	p.87
Technical Report Exec. Summary	p.88

MARCON PLUS: TEXT RETRIEVAL SOFTWARE A REVIEW

MARCON and MARCON Plus are versions of a software package from AIRS, Inc., 335 Paint Branch Dr., College Park, MD 20742; It runs on an IBM PC-AT, COMPAQ 386 or 286 or compatibles with hard disk, under DOS 2.1 or later (3.1 for networked version). It requires a minimum of 640K RAM and at least a 10MB hard drive. MARCON operates with a color or monochrome monitor, outputs to most standard impact printers and accepts input from the DEST PC-Scan and other OCR devices. MARCON is available without security, cross database searching or saved search features for \$495. MARCON Plus, includes the above three features for \$695. A LAN version of MARCON Plus for five workstations costs \$3200 (\$6400/10; \$11,000/20). MARCON Plus is also used as a retrieval system for custom developed CD-ROM; prices negotiable. This review of MARCON Plus 3.01 was conducted on an IDR 386, with 1MB of RAM and 20MB of disk.

MARCON is text retrieval software par excellence. It does wonders with large databases of irregular, variable length records including significant text. It builds databases with sophisticated authority control and retrieves from any or all fields of records in one or more files using equally sophisticated online retrieval approaches. Archives and museums will be attracted to MARCON by the ease with which they can build and search complex databases but it is not without weaknesses. While MARCON is a powerful, generic text retrieval system, ideally suited to constructing and searching files comprising a research catalog, it is less well suited to public use as a catalog, or to collections management and/or project administration applications.

MARCON is a tool. The system is driven by a data dictionary which defines every field used in any record in any file (AIRS calls files "collections"). Because MARCON is a tool, and not a ready made application, everything in MARCON begins with the fields and the collections which they comprise. When we want to define a new database, we name a collection (a file), identify any previously defined fields which will be used, and define new fields in the collection. When we want to add data, we define a collection and are permitted to add data to fields according to the rules which we previously established in the data dictionary. When we want to search a collection, we identify the fields which will comprise the search argument, and then fill in the values which we want to search in each field. Because MARCON is a

tool, this review examines two major aspects of its use: how easily it can construct an application (and how complex that application can be) and how well it performs data entry, retrieval and reporting from the resulting application.

BUILDING AN APPLICATION

Getting Started: When we first encounter MARCON, it consists of five disks of highly compressed program data and a manual. Installation proceeds smoothly, although it may require modifications to your "config.sys" file and will consume about 4MB once finished "unchrunching", "unpacking" and "unsquashing" (messages which appear on your screen as decompression routines owned by PKWARE shareware, go to work). Initially, you will get directly to the main menu, but as systems administrator you will be free to introduce security control requiring others to provide a password at this stage to view, or update, any collections and you may direct them to a different initial menu.

Documentation and User Interface: At the main menu, if you select "Data Entry and Design" you will be asked to identify a collection in which you want to enter data, or that you want to design. About this point you will encounter one of MARCON's weaknesses: its online help often goes on for several densely packed screens, and the printed documentation, which now comes in a very professionally printed and bound package, is not as good as it was in earlier releases. Indeed, it often neglects to give essential instructions, or provides them in the most

unlikely places. Occasionally it conflicts with the online help; and it rarely builds from the same structure. The documentation no longer follows the way in which a user is likely to encounter MARCON, and I found the logic difficult to follow, though the index was good. Apparently, AIRS hired a technical writing firm to produce this mess; they advised me that they have now decided to return to their in-house writer and will soon republish the documentation, but beware.

Fortunately, one of MARCON's strengths is that it has a very consistent user interface. Although my documentation failed to explain that F7 would delete collections, records, fields, etc.; I eventually found it in a table explaining the function keys (below), and was fine from then on. Similarly, once I learned that F8 would permit me to change the item on which the cursor was located, and that F10 saved my changes, I could move through the design process with ease.

Summary of the Function Keys

F1	Help
ALT-F1	User comments
F2	Index list window
ALT-F2	Index list evaluation
F3	Page to previous record
ALT-F3	Copy from previous record
SHIFT-F3	Saved query - use
F4	Page to next record
ALT-F4	Import from DOS file
F5	View enlarged field
ALT-F5	View all fields
SHIFT-F5	View query
F6	Print
ALT-F6	Export Dos file
SHIFT-F6	Reports
F7	Delete
F8	Edit design
SHIFT-F8	Edit ANY/ALL query
F9	Define a new...
ALT-F9	Define new nested field
F-10	Save/Go
ALT-F10	Save query
SHIFT-F10	Save/Go and begin another...

Menus can always be navigated by the cursor, tab, or a unique first character, whichever the user preferred. Items can be transferred from one window to another by "point & shoot" (locating the cursor on the desired item, and hitting ENTER or RETURN) which is an especially nice feature when combined with the windows MARCON employs for indexes. Of course, consistency does not mean that the interface is intuitive; and the 23 commands created by combining the Alt and Shift keys with ten function keys, provide numerous features which users must learn, or they will fail to take advantage of MARCON's power. It would be helpful, especially for occasional users, if the most basic of these functions were defined on relevant screens in an options bar.

Design: The power of MARCON becomes evident as soon as we begin to define the first field of the first collection. We can define a highlighted screen box of any size, and a field of any size, so that the field may be required to fit in, or scroll through, the box. We can define the data as the usual alpha-numeric, numeric, and date, or as a date stamp from system time, and also have the option to define it as calculated from a defined equation. We may restrict the field to one occurrence or define it as multiply occurring, in which case it may have many values within a single record, or we may define it as inverted text, so that each word, (excepting those on an editable stop word list), will be indexed separately. We can define a "list" or index, which will hold values of the field, and restrict data entered in the field to values pre-existing in that list, or to unique values (not in the list), or permit the list to grow by addition of new values, thus constructing an index on the fly. Ultimately we will discover that many fields can be defined to the same list, that the list will be viewable in a window at data entry and searching, and that values may be entered, or used in searches, by point and shoot from that window, thus reducing errors. The list can be viewed in a sorted or

unsorted fashion. And we have a huge range of value checks that can be applied to the field, either as a warning or as an error condition, each with its own system administrator provided messages. Clearly, designing each field will take some time; we may want to change our choices frequently in early stages of the design. MARCON supports such changes easily, even building and unbuilding the necessary indexes as soon as the changes are implemented.

But this only begins to hint at the complexity of a potential MARCON database. First, the field we designed is part of a data dictionary; we need to invoke it in the first collection, even if that collection is still the only one. This means that the field is available for other collections, and cross database searching (supported by MARCON Plus) can exploit this feature. Also, the field we defined is a simple field, and MARCON supports "nested fields" or "sub-fields" which retain their relationships. Nested fields can be concatenated into a single index, serving as a meaningful key (author/date) or a hierarchical term (Bldg/room/range/shelf, etc.). And like values of other fields, these indexes can be displayed on data entry or searching.

Each subsequent file within the database can be designed to link in a number of ways. It can literally have a redundant data element (including one that is unique in the first "collection" and likewise unique in the linked "collection"). Or it can point its field values to indexes shared between files. Or it can be related by any word in the text of the record, regardless of field (MARCON supports a modifiable stop word list to exclude local terms or words which are meaningless in a particular context).

USING AN APPLICATION

Searching: The linkage across words in the text brings us to the second part of the evaluation of MARCON. How does it perform as an application, once the database is built? For a sophisticated searcher, MARCON is a stupendous information retrieval product. Any field can be search, and combined in a

Boolean AND or OR with any other hits. Any number of values may be searched in a field, and each may be combined as an AND, OR, or NOT. Values may be specified as exact matches, as wild cards of one or more letters, as range searches within dates or numbers, (or even letters) and for their absence. Any word in text fields may be specified and it may be requested in proximity to another word by number of words, by sharing a sentence or a paragraph. Queries may be saved (in MARCON Plus) and re-executed. Queries can be viewed during the search. Number of hits are displayed for each field and for the overall statement. The number of matches is shown when the first match is displayed.

For assistance in formulating a query, record indexes may be displayed by any indexed term and views may be shifted from one to another. Terms may be selected from lists and appended to any Boolean expression with a single keystroke. Query statements may be extensively edited prior to execution. Execution is quick, since the indexes have been built in batch. It would be difficult to find a product anywhere with all these features. Unfortunately, most users won't experience it because all these delicious options are not displayed on the screen in a guide bar, and are inherently complex, but experienced searchers will find MARCON a delight.

Data Entry: First the data has to get in. MARCON is again a pleasure for text heavy databases in this respect, since data can be keyed in or read in by scanner or from an ASCII file. Keying data is made easier by a number of facilities in addition to the "point & shoot" capability alluded to earlier. Among these are full record copy from prior record and field copy from prior records. Some help can be provided at the data dictionary design stage (or later) by providing user messages for errors (but note that my copy displayed only one of three lines of error message). Global changes can be made to records from an index list, making it easy to correct input errors and to change headings across an entire database. Several odd

decisions in the data entry module keep it from being among the best. For example, after you enter a record, the screen defaults to a "choose a record to modify" screen, rather than directly to the next blank template. Fields can't be defined as incremented nor can default values be defined in the data dictionary - both of which would make data entry go faster for some fields. In my copy, warning messages acted like error messages and could not be over-ridden as documented.

Reporting: For two years, reporting has been the crippling weakness of MARCON. This release, 3.0, was designed to address this handicap with a new report writer. If the report writer performed as documented, the module would be weak, but acceptable. However it doesn't, and isn't.

Two exceptionally useful kinds of "default" reports are provided by MARCON. The first is a printout, or write-to-disk, of each record in one collection. The user selects which fields will appear, up to thirty (!) sort fields, and whether each record will print on a separate page or three lines after the previous record. Field labels will print, each field will appear on a new line, and data from only one collection will be reported. The second default report is a printout of any of the index lists, complete with the number of occurrences per term if desired.

MARCON now allows a third kind of report, one that can be defined using the new report writer. The only reporting capability that works flawlessly here will print a table of fields and values for selected records or all records, again from one collection. All three of these reports are easy to use (tables are more complex to define, but not much) and are useful.

Unfortunately, for the present, we must stop here. The report writer documents the method for writing a report in which the page location of each field selected is defined by coordinates and fields may be displayed with or without labels anywhere on a page of any size (effectively permitting formats such as mail labels, card catalog cards etc.). In my copy, these reports don't work. AIRS

acknowledges that there have been serious bugs and is working to repair them, apparently with success, but buyers should evaluate this feature fully.

MARCON files may be unloaded, and reloaded, in a MARCON specific format. Further improvements in file output are planned, including implementation of MARC record input and output, but apparently it is hard for a company which makes what might be the premier PC-based text retrieval package to invest enough in reporting. For the present, the inability to report across collections is a serious limitation for anyone designing a collections management system, where multiple files are linked in ways which cannot entirely be satisfied even by nested fields and repeat occurrences.

UTILITIES:

Database management: MARCON provides some useful utilities to assure that the databases being constructed are consistent, up-to-date, and error free. One function updates all the inverted text indexes for newly entered, or altered, records. One feature alluded to earlier permits global changes in records from indexes, automatically replacing values in records with new values in the list, and allows indexes to be deleted. A "database manager" function reorganizes the storage of a database which has been heavily edited, saving space and increasing searching efficiency. A batch loading facility provides a method to reload MARCON data or load an ASCII file into a field or collection. These functions, which take a reasonable amount of batch system time, have useful on-screen countdowns indicating the number of records left to process and summary screen reports when the process is completed.

Security administration: In MARCON Plus, the systems administrator is empowered to establish user accounts with passwords for the application and each of its major modules, for a collection, for a field, and for a function (add, edit). This level of security is adequate for most applications and works well. Oddly, within this module

the administrator also sets the colors for each users' screens. While I don't imagine this will be customized by most systems administrators, it is worth noting that MARCON has done an excellent job using color, which contributes to ease of entry and searching, but that its highlighting is also fine for a monochrome monitor.

Thesaurus: MARCON developed out of a system that AIRS president, Ted Durr, designed for management of documents at BRISC, an urban studies research center in Baltimore known to many for the Urban Information Thesaurus which Durr authored there. The system was designed to handle a thesaurus, and is still marketed with a module for thesaurus development, but my copy did not contain this module. It has been used by the Art & Architecture Thesaurus and other archives and museum projects, but is not being pushed by AIRS at present because of the strain it places on the system. AIRS expects to reissue the thesaurus as part of a new architecture for the overall system, in 1988.

SUMMARY

MARCON is an exceptionally full featured information retrieval system, well suited to the construction of online research catalogs, and databases containing large quantities of text. It is less well suited to structured database applications, especially those requiring substantial printed reporting, for which it will be found wanting.

David Bearman

NOTE: After announcing that I would write this review, I was contacted by AIRS to assist them in designing a MARC I/O module for their system, and have accepted a contract for that work. This review has not been seen by anyone at AIRS and I will not begin work on the MARC I/O module until after it is published, so I feel there is no influence from them, but I do want to clarify what could be seen as a conflict of interest.

MACHINE-READABLE VIEWS

by **THOMAS E. BROWN**

To conclude the first full year of these ruminations on archives and machine-readable records, this column will harken back to earlier comments. These remarks will probably tie up some loose ends, and, at the same time, unravel some tied ones.

The first issue of the Archival Informatics Newsletter (vol. 1#1) contained a technical information leaflet on commercially available database management systems (DBMS's) which Bill Reader and I wrote. I have since had the opportunity to discuss the leaflet with a technical representative of a vendor whose product the leaflet discussed. Happily, the leaflet accurately outlined the technique to extract information from the DBMS and put it into a software and hardware independent format. Further, the vendor representative estimated that someone familiar with the DBMS would need only about 15 minutes to write the program to unload the data in the desired format. However, he admitted that processing time would be inordinate. He estimated that it would take about 12 hours of run time (3 hours of CPU time) to unload a database of about 5 million characters. Thus, for this DBMS, it is not a question of whether off-loading the data is technically possible, but whether it would be feasible.

In the second issue (vol. 1#2), John MacDonald of the National Archives of Canada urged that archivists should view themselves as corporate users who need to include archival concerns in functional requirements during the system design phase. I note that about a year ago, the National Bureau of Standards issued guidelines to U.S. Federal Agencies on how to write functional specifications for database management systems. The obvious question is whether this guideline, issued as a Federal Information Processing Standard (FIPS),

addresses the archival concern over transportability.

The guideline alludes to transportability twice. First, it defines conversion capability as one of the "global data factors". Unfortunately, the sample specification here is for inputting fixed length or comma delimited data into the DBMS, not for outputting similar formats. The second reference appears under "other specifications" as a requirement for "database transporting". Here the specification does include requirements for offloading data. It says:

"The DBMS must have a facility that can selectively dump portions or the entire database (i.e., the defined schemas and associated data occurrences). The dumped file must be loadable by the DBMS. The DBMS must allow copying the database record occurrences to standard files for information interchange."

Interestingly, the vendor whose DBMS needs 12 hours to unload a 5MB database into a transportable format would meet this specification.

My discussion of the evolution of appraisal of electronic records (vol. 1 #3) elicited considerable comment. The review alluded to a meeting on May 27, 1983, of a National Archives Appraisal and Disposition Task Force. At the conclusion of lengthy discussions, a consensus developed around four positions. Their articulation was not formally incorporated into the final report of the Task Force, but I think these statements may still be basically valid and deserve to be disseminated and discussed. I welcome your reactions or comments.

1) The disposal provisions for machine-readable administrative or housekeeping records may not always be the same as those which have been developed for textual records. The long-range strategy is to incorporate appropriate disposal provisions into the General Records Schedules which now concern administrative records in only textual form. Until this is achieved, the short term solution may be to state within those General Records Schedules that the

disposition do not apply to machine-readable records with similar information.

2) While machine-readable records have been appraised primarily for their informational value, they may have evidential value. Thus machine-readable records may be appraised as having evidential value worthy of continued preservation although textual records with comparable contents exist.

3) The overall goal of the National Archives is to accession only records in a software and hardware independent format. However, in unusual situations, the information may have such value that destruction is not justified and may be so software and hardware dependent that the resource commitment for reformatting is not justified. In such a situation, the National Archives can accession a data file in a hardware and software dependent format and leave it to a future researcher to bear the costs of reformatting.

4) During the scheduling process, the appraisal should include a technical analysis of any software or hardware dependency. However, the full extent of these technical problems may not be apparent until the time of scheduled transfer. If significant technical problems become evident at the time of accessioning, these problems can warrant the destruction of information which was earlier appraised as permanent.

T.E.B.

EDITORIAL COMMENTARY

I recently encountered a case of the consequences of the result of following Tom's recommendation #3. University of Massachusetts project archivist Judy Hensche is looking for firms which can reconvert a massive collection of microform representing one million captured Vietnam war documents whose index entries were (once) in machine readable format, and are now recorded as microfilm images of "Flexowriter" punched cards.

Documentation of the Intelligence Data Handling Systems (IDHS) of the U.S. Military Assistance Command in Vietnam, obtained from the National Archives, is helpful as background to understand the system which produced a huge collection of microforms at the University of Massachusetts Boston Harbor Campus, but it does little to help Judy and her researchers to access the documents on the films. What can be gleaned from the documentation is that the Department of Defense used an "Automated" Document Storage and Retrieval System manufactured by FMA Inc. and sold commercially during the 1960's as the FileSearch (Flexowriter) system, to index captured documents.

Index terms were highly coded to save space, and the codes were keypunched on the Flexowriter. The document, with its associated index worksheet and machine-punched card, were filmed together. The Filesearch system used a unique, seven bit, binary code, which complicates matters somewhat, but the system as a whole is reasonably well documented.

Unfortunately, the cards were not read onto magnetic media and into standard data representation codes, as these new and non-proprietary approaches superseded the Flexowriter system. As a consequence, the costs of re-indexing the entire collection, or of converting the data from the micrographic image of the Flexowriter cards will fall to future researchers.

It is intriguing to consider whether this represents good social policy, is cost effective and is fair. I find myself thinking that even though it is obviously more costly to convert the data now than it would have been in the early 1970's, it seems on first examination to be reasonable social policy since it places the burden on those who want to use the materials. But I would not be willing to accept the consequences if the data were U.S. voting returns or information "essential" to a knowledgeable citizenry. I would welcome your comments and other examples. Judy would appreciate offers (or bids) to convert the data. D.B.

SAA-CGAP PLANNING GROUP REPORT

In February 1987, the SAA Committee on Goals and Priorities created a planning group to develop a detailed action agenda to implement goals and objectives relating to automated records and techniques outlined in Planning for the Archival Profession (commonly known as the GAP Report). The planning group met in Ann Arbor on July 28, 1987. Its report consists of two sections - Automated Techniques and Automated Records. Reprinted below are "Proposed Activities" statements from the planning group report. The full report includes additional text on Purpose and Significance, Potential Actors and Possible Resources, and a Timetable for each proposed activity.

AUTOMATED RECORDS: A PROPOSED ACTION AGENDA

- I. "Establish a clearinghouse or means of exchanging information about activities in the area of electronic recordskeeping practices and techniques. The clearinghouse could contain information about existing or planned activities in this field, examples of general records schedules, examples of legislation, information about access and privacy issues, and technical material about preservation, data transfer, and new technologies."
- II. "A regular series of articles should be written and published in the SAA Newsletter on various aspects of automated records and techniques. The articles would summarize activities and projects that are being undertaken in these two areas by the Society itself, by information professionals in other associations and organizations, as well as within archival institutions throughout Canada and the United States. These articles should be made available to other publications, such as the newsletters of regional archival associations and specialized newsletters, e.g. Archival Informatics."
- III. "A research agenda should be developed for the archival profession in the area of automated records and techniques. The

research agenda should identify activities and projects that are currently being undertaken in these two areas by the Society and its members, by other associations and organizations, as well as by archival institutions throughout Canada and the United States. The research agenda should also identify the outstanding issues that have yet to be addressed by the archival profession. The agenda should describe in considerable detail the various actors or players who should be involved in these activities, the resource implications, and propose a number of options for coordinating the work of the various actors, thereby ensuring that long-term objectives are attained."

IV. "Develop a workshop on the life-cycle management of automated records."

V. "Workshops for archival educators to introduce or improve knowledge of concepts and methods of teaching about automated records and techniques in basic archival education courses."

VI. "Influence the interim certification board to ensure that there will be adequate coverage of machine-readable records and techniques included in the certification process."

VII. "Identify college and university organizations and associations whose members create and maintain records. Develop guidelines with these organizations for preserving archival documentation, especially machine-readable records of the colleges and universities."

AUTOMATED TECHNIQUES: A PROPOSED ACTION AGENDA

I. "Continue to offer workshops developed by the Automated Records and Techniques Task Force and develop additional advanced topic workshops."

II. "Develop a workshop for planning for automated techniques for archival administrators."

III. "Develop mechanisms to ensure that the appropriate SAA groups (committees, task forces, etc.) are consulted about official SAA publications that are to be developed or

revised in their particular area of interest and expertise. Such groups could then integrate such publications into their own plans and activities, suggest additional materials (such as workbooks) that might supplement the publications, and suggest distribution mechanisms and avenues for maximum impact."

IV. "Develop teaching techniques to disseminate basic information to beginning archivists."

V. "Propose a mechanism to influence NARA and LC to provide leadership on automated techniques."

VI. "Develop functional specifications for local archives and manuscript control systems."

VII. "A. Develop a mechanism to evaluate archival descriptive standards.

B. Undertake retrieval studies in automated systems using the MARC AMC format to develop better descriptive standards.

C. Work towards clarifying definitions for MARC AMC fields 655 (Genre/form terms) and 755 (Physical characteristics access). Develop guidelines for applying terms.

D. Appoint a Library of Congress Subject Headings Task Force to suggest changes and modifications of LCSH to the Library of Congress. Develop guidelines for archival use.

E. Produce a frequency distribution list of terms used in the various access points in the MARC AMC 6xx fields in the RLIN AMC file.

F. Revise and expand Archives, Personal Papers, and Manuscripts.

Planning group participants were: Larry Dowler (Harvard), Fynette Eaton (NARA), Tim Ericson (SAA), Carolyn Geda (ICPSR), Margaret Hedstrom (N.Y. State Archives), David de Lorenzo (Gallaudet), Harold Naugler (Nat. Archives of Canada), Donn Neal (SAA), Lee Stout (Penn. State), Richard Szary (Smithsonian), Bill Wallach (Michigan), and Lisa Weber (SAA)

CONFERENCES

MUSEUM COMPUTER NETWORK

The Annual Meeting of the Museum Computer Network held October 13-14, in Cambridge, Massachusetts, witnessed a smorgasbord of papers in what seemed like a single unending session, distinguished by the numerous important projects discussed but hampered by lack of formal discussion periods. Exhibitors and project demonstrations competed with the formal session and with numerous *ad hoc* discussion and special interest groups, and were, as a consequence, much less successful. Abstracts of papers presented at the meeting will be published in the next issue of SPECTRA, and some of the full papers will appear in subsequent issues (available as a membership benefit for \$25 from MCN, Box 111, East Winthrop ME, 04343).

Participants heard reports on many major international projects, including papers by D. Andrew Roberts on the Museum Documentation Association of the U.K., Toni Peterson on the Art & Architecture Thesaurus, Jim Blackaby on the revision of Nomenclature and on the "Common Databases for History Museums" component of the AASLH Common Agenda project, Marilyn Schmitt on the Getty/ADAG and other information sharing projects, Rachel Allen on the Inventory of American Sculpture, Margrethe Floryan on the Art Index of Denmark, Joan Bacharach on the U.S. National Park Service National Catalog, and Mary Sullivan on the Virginia Museum of Fine Arts/Canadian Heritage Information Network cooperative systems analysis project. Also, a panel of members who participated in the fall ICOM/CIDOC meeting reported on that organizations' working group projects.

In addition to updates on these major international ventures, participants were treated to a series of papers and panels on implementation issues, ranging from database design (at the Museum of Games, University of Waterloo; Mystic Seaport, and Princeton University Art Museum) to more general issues such as forming computer

users groups, attending to the legal issues in software acquisition, managing risks in implementations.

The keynote address by David Bearman of Archives & Museum Informatics argued for increased attention to development of standards for information exchange and introduced the concept of "open" or permissive standards which codify practices but do not require uniformity between in-house systems which use different subsets of the standard at their option.

Two research reports stand out for me. Dr. Jane Stone (Dept. of Computer Science, Montana State University) reported on a simple technique she developed to measure items depicted in an archeological imagebase and, potentially, to scale digitized images so that they could be displayed together. She locates two points on each image for which the distance is known (a pencil or shoe or person or object of known length) and the machine then computes the scale and calculates the dimensions of other objects in the image. To date the technique has been applied only to videodisc images, but Dr. Stone intends to extend it to digitized data.

Dr. D.D. Hilke reported on audience research she conducted in conjunction with the SITES, "Laser at 25" exhibit. Concern has often been expressed that interactive components in exhibits might compete with other means of presentation, so Dr. Hilke set out to assess whether the computer was useful in its own right and whether it detracted from other parts of the exhibit in a controlled experiment involving the exhibit at the Maryland Science Center in Baltimore and the Discovery Center in Fort Lauderdale when the 15-minute interactive computer program was operating and when it was not. Researchers observed visitor behavior (visual information seeking, verbal exchange, hands-on interaction) and recorded some demographic data (adult/child and gender). The findings are extremely encouraging to software developers: all kinds of interaction and retention were higher when the computer was on than when it was off, even among visitors who did not themselves use the computer! 15% more people visited the exhibits when the

computer was on, they spent more time in the exhibit, but no less time on the rest of the exhibit just because the computer was available, and they engaged in 24% more "information seeking" behaviors when the computer was on.

Two developments within the MCN itself should also be noted. The MCN extended voting privileges to all individual members of the organization (voting was previously limited to organizations) and it elected to hold a full day "vendor/user" forum in conjunction with its meeting in Los Angeles, October 26-28, 1987. The forum will provide an opportunity for vendors to discuss strategic decisions in their software and contrast their products and for users to compare systems. A call for papers will be issued in January; preliminary proposals for papers or sessions should be submitted by March 1, to David Bearman, 1988 MCN Program Coordinator, Archives & Museum Informatics, 5600 Northumberland St., Pittsburgh, PA 15217.

AMERICAN SOCIETY FOR INFORMATION SCIENCE

ASIS held its 50th annual meeting in Boston, the week before MCN. The event was staged in a grand fashion, with plenary addresses daily from a stellar cast including Robert McCormick Adams (Secretary of the Smithsonian), Joseph Duncan (Chief Statistician, Dun & Bradstreet), Congressman George Brown, Arthur Miller (Harvard Law School) and Jacques Vallee (EUROLINK International). Seventy other sessions were held in concurrent slots, so this report will be impressionistic and reflect my choice of sessions and the competing demands of exhibits and my committee activity as in-coming Chairman of the Arts & Humanities Special Interest Group (SIG-AH). I left the meeting impressed by ideas I heard articulated around two broad themes: the societal and organizational impact of information systems and empirical studies of retrieval methods and indexing theories.

Dr. Adams tried to place the "information society" in an evolutionary, rather than revolutionary, relation to prior eras in his philosophical opening remarks. Calling on historical evidence, he downplayed the ability of technology to increase knowledge, noting that knowledge, unlike information, is a personal acquisition, requiring a human intermediary, but his broad perspective did not dampen the enthusiasm for technological fixes and razmataz shown by others.

Evidence of recent information management evolution was reported by James McKinney and his colleagues at Harvard Business School. McKinney et al. have been studying internal corporate communications in great detail for the past six years. Their intensive observations of all communications between employees of several firms on at least two occasions of at least five months duration each, provided a basis for an exceptionally rich analysis of the impact of electronic mail systems on businesses. Their findings that E-mail decreases the percentage of time managers spend in meetings and increases the amount of face to face contact are less interesting in themselves than the fact that use of electronic mail works. It works by improving preparation for meetings (the largest reason for the decrease in time spent in meetings), improving service response time, increasing the depth of employee understanding of management priorities, and shortening development time in projects (in part by reducing the number of unanticipated bugs). For electronic mail to succeed, however, McKinney and his colleagues found that organizations must change the way in which they work - building informal networks in place of formal reviews and budget cycles and by managing the communications process itself. In the question period, McKinney acknowledged that one area of corporate structure that has not accommodated itself to the electronic environment is records management and that mechanisms for ensuring corporate accountability have yet to emerge in the firms he has studied.

Jamie Grodsky of the U.S. Congress, Office of Technology Assessment, reported on another aspect of corporate behavior which

has yet to adjust: dissemination of press information and information to the public through the Freedom of Information Act. Grodsky pointed to the plethora of information outlets maintained by agency press offices, noted the resulting reduction of access to information, and hinted at the likelihood of a federal wire service in the near future. She then turned to the FOIA and examined the problems which the law (enacted when there were fewer than 3000 computers in the entire Federal government) and the case law which has grown up around FOIA litigation, have created for access to information in electronic formats. The case law for records in traditional formats has clearly established that agencies are required to release "records in being", not to create new records, but requests for electronic information pertain to records which are created in response to the request and are not "in being". The law has established that records must be retrievable through a "reasonable" search, which case law has defined in a fairly constrained way; with electronic systems, it may not be unreasonable to demand more extensive searches. Finally, because the law limits release of proprietary materials, the software in which data resides is not releasable, and this may effectively make many data sets unavailable through FOIA. OTA is studying these and other information dissemination issues and invites comment.

A large number of papers grappled with empirical analysis of the relationship between indexing methods, retrieval success, and other factors. Fortunately, many of these papers are published in full in the Conference Proceedings, available from Learned Information (members \$32, others \$40).

In two valuable papers, Marcia Bates reintroduced the content of the databases as a too often ignored variable in users studies and Carolyn Eastman examined the role negation plays in queries. In a challenging session on automatic indexing, Susan Humphreys reported further on the frame-based indexing methods in use at NLM and Gerald Salton expanded on his arguments for

weighted terms and vectorized searching over Boolean query.

Two papers of special interest to archives and museums, both dealing with new technologies are printed in full in the Proceedings. Howard Besser presented the experiments in digital imaging at the University Museum in Berkeley in an exceptionally rich paper, and George Thoma discussed design issues involved in the implementation of cost-effective image based document management systems at the National Library of Medicine's Lister Hill Research Center. Thoma documents second by second that optical scanning throughput, the human labor component of imaging, will continue to be the critical variable in capturing texts and images using digital technologies, but that many minor engineering adjustments can be made which, in the aggregate, significantly affect costs.

ARCHIVES AND AUTHORITY CONTROL

A seminar on Archives & Authority Control was sponsored by the SIBIS-Archives Users Group and the Office of Information Resource Management (OIRM) at the Smithsonian on October 27. Full proceedings of the conference, including discussions, will be published as an Archival Informatics Technical Report in the summer of 1988.

Over a hundred Smithsonian archivists, curators, information systems support staff and guests attended the seminar to hear Jackie Dooley (U.C. San Diego) introduce the concept of authority control, Tom Garnett (SI Libraries) discuss the experience of a library implementing authority control, Marion Matters (Minnesota Hist. Soc.) contrast it with the experience of a state archive, Lisa Weber (SAA) detail the state of authority control as an issue for the archival profession and Rich Szary (OIRM) assess the relationship between technical capabilities of the software system and needs of the archives users community. David Bearman (Archives & Museum Informatics, led a discussion of the speakers and the audience.

CONFERENCE CALENDAR 1988

February 4-11, Art Libraries Society of North America, Annual Conference, Dallas TX [ARLIS/NA, 3900 Timrod St., Tucson, AZ 85711]

March 21-24, RIAO-88 "User-Oriented Content-Based Text and Image Handling", Cambridge, MA [M.I.T., Conference Services, Bldg 7 Rm. 111, Cambridge, MA, 02139]

April 13-17, Society of Architectural Historians Annual meeting Chicago [SAH, 1232 Pine St, Philadelphia PA 19107]

April 21-22, Nottingham, England
"Computer in Museums seminar, sponsored jointly by the Museum Documentation Assoc. and the Museums Assoc., [MDA, 347 Cherry Hinton Rd., Cambridge, CB1-4DH, UK]

May 15-18, American Society for Information Science-Mid-Year meeting on Artificial Intelligence, Ann Arbor, MI. [ASIS, 1424 16th St. NW, Washington, DC 20036]

May 26-29, International Association of Social Science Information Systems & Technology (IASSIST-88), Washington, DC [Pat Doyle, Mathematica Policy Research Inc., 600 Maryland Ave., SW, Suite 550, Washington, DC 20024]

June 3-7, American Association of Museums Annual Conference, Pittsburgh, PA [AAM, 1225 Eye St., NW, Suite 200, Washington DC, 20005]

June 6-10, Association of Canadian Archivists Annual meeting, Windsor, ON [ACA, P.O.Box 2596, Station D, Ottawa, K1P-5W6, CANADA]

June 12-16, Nebraska Videodisc Design Workshop. Basic workshop dates for 1988 are March 20-24; May 22-26; July 17-21. [1800 North 33rd. St., Lincoln, NE 68583]

July 20-23, National Association of Government Archives and Records Managers, Annapolis MD [NAGARA, New York State Archives, Room 10A75, Cultural Education Bldg, Albany, NY 12230]

August 22-26, "Museum Archives", Washington, DC, [Smithsonian Institution, Office of Museum Programs, Washington DC 20560]

September 14-17, American Association for State and Local History, Annual meeting [AASLH, 172 Second Ave. North, Suite 102, Nashville, TN 37201]

September 21-24, Second Annual Museum Documentation Association Conference, "Terminology for Museum Documentation"; Cambridge (UK); preceded by a study tour for the interested [MDA, 347 Cherry Hinton Rd., Cambridge CB1-4DH, UK]

September 29-October 2, Society of American Archivists Annual Conference, Atlanta Georgia [SAA, 600 S. Federal St., Suite 504, Chicago, IL 60605]

October 2-6, 2nd Annual Library & Information Technology Association Conference, Boston MA [LITA, 50 East Huron St., Chicago, IL 60611]

October 23-27, American Society for Information Science, Annual Conference, Atlanta, GA [ASIS, 1424 16th St., NW, Washington, DC 20036]

October 26-28, Museum Computer Network, Annual Conference, Los Angeles [MCN, P.O.Box 111, East Winthrop, ME 04343] Conference includes Vendor/User forum on October 26. To participate as a vendor or propose sessions or speakers, write David Bearman, 5600 Northumberland St., Pittsburgh, PA 15217

IN-BOX

REPORTS

A Common Agenda for History Museums: Conference Proceedings, February 19-20, 1987; edited by Lonn Taylor, Nashville, TN, American Association for State and Local History, 1987, 53p.

This volume contains several thoughtful articles which are important in their own right, not just as a call to action. James Blackaby's categorization of the types of information found in history museums is of particular interest. The Common Agenda project has recently received additional funding to enable it to pursue the development of common data standards.

Government Information: An Endangered Resource of the Electronic Age, papers presented at the first Special Libraries Association, State-of-the-Art Institute, October 19-22, 1986, Washington DC, SLA, 1986, 277pp. \$21.75

Papers in the on-going debate about the appropriate role of government and the private sector in dissemination of public information.

Program Reporting Guidelines for Government Records Programs, National Association of Government Archives & Records Administrators and the Council of State Governments, Sept. 1987, 26pp. [CSG, Iron Works Pike, P.O. Box 11910, Lexington, KY 40578]

This protocol for statistical data collection represents the culmination of years of efforts by NAGARA to arrive at measures of activity which could represent cumulatively the impact of public records programs. One hopes it will be adopted and tested widely and evolve over time so that reliable data can be compared and programs can be assessed. Very useful definitions and measurements are proposed which other repositories will want to study.

Report of the First National Conference on Issues Concerning Computerized Public Records, Public Records Division, Massachusetts Secretary of State, 1987, 2 vols.

The conference reports on the conclusions of four working groups consisting of state freedom of information act officials, commercial and public advocacy representatives and legal and political figures, which addressed a common set of issues. The consensus report seems to open criteria for access by concluding that release of automated records under FOIA constitutes release of "pre-existing" records, not creation of new records, but to restrict criteria by agreeing that access to certain kinds of records, such as property assessments or professional licenses, which involved individual citizens could be denied. The details of the working group discussions are informative for anyone concerned with access issues. A follow-on conference is scheduled for 1988.

Virginia Museum of Fine Arts Information Systems Framework, Richmond VA, VMFA and Ottawa, Ontario, CHIN, 1987, unpaginated.

This gigantic report on the results of a collaborative effort by the Canadian Heritage Information Network and the staff of the Virginia Museum of Fine Arts to apply CHIN's "Corporate Information Systems Framework" (a methodology akin to IBM's Business Systems Planning) is worth careful study. Its scope is all the information systems of a moderate size museum (106 separate data entities, 23 major activity groupings each comprised of numerous sub-activities), and its objective is total intellectual integration. It is the best museum systems foundation document I know of, including those on which I worked while at the Smithsonian.

NEWSLETTERS

CD Data Report (ISSN-8755-5727), Langley Publications Inc. 1350 Beverly Rd., Suite 115-324, McLean, Va 22101, \$225 p.a.

Editor Linda Helgerson has done a superb job of following this emerging market and reporting on it monthly from all angles; hardware, software, users, products. Issues are lengthy (40+p.), well-written and timely.

Information Retrieval & Library Automation (ISSN-0020-0220), Lomond Publications Inc., P.O.Box 88, Mt. Airy, MD 21771, \$66.p.a.. Commencing vol.24 in 1988. Its edited press releases and publication notices include some items which are news to me every month.

Registrar, Newsletter of the Registrar's Committee of the American Association of Museums (\$10 plus AAM membership).

The Fall 1987 issue of Registrar is devoted to automation, but previous issues on aspects of the registrarial function such as the use of couriers (Spring 1987) are equally substantive and no less useful as guidelines for standard practice.

Research Libraries Group News (ISSN-0196-173X) is a free, occasional, house organ which often has articles of interest to the archives and museum community. The September 1987 issue (#14) contains a report on the King Papers Project and its use of MARC-AMC in a primary publication, historical editing, project.

SCOPE: Humanities Computing Update (ISSN-0735-8296), Paradigm Press Inc., Osprey, FL, 34229-1057, bi-monthly, \$52.p.a.

SCOPE contains a wealth of bibliographic, references, references to hardware and software reviews, and software notices in addition to an occasional short article. Selected past items from SCOPE software, courseware reviews and software reviews are issued on floppy, as DiscSCOPE (\$5).

ARTICLES & BOOKS

Barnett, Patricia & Lucker, Amy E., Procedural Guide to Automating an Art Library, Occasional Papers #7, Tucson, AZ, Art Libraries of North America, 1987 40p. \$15.00

The title is misleading, but the papers in this volume, on retrospective conversion, online catalogs, authority control and the relationship between object and bibliographic information systems, are useful additions to the literature and the "Directory of Projects and Systems" compiled by the Clearinghouse on Art Documentation and Computerization will provide at least one new, and essential, contact for anyone.

Bearman, David, Towards National Information Systems for Archives and Manuscript Repositories: The National Information Systems Task Force (NISTF) Papers, Chicago, SAA, 1987, \$10, 119pp.

The NISTF planning papers and two conference papers on NISTF from 1983 and 1984 are combined here with a new introduction discussing the significance of the issues raised by NISTF for archivists today.

Crawford, Walt; Patron Access: Issues for Online Catalogs, Boston, G.K.Hall, 1987 250p. + index

This small book contains so much uncommon sense about online catalogs that it should be required reading for anyone designing a reference database. From the first dictum, that "online catalogs should serve patrons at least as well as card catalogs do" a series of assertions are made (and amplified by examples) which are, collectively, the best guidance yet published. I am particularly impressed by the implications Crawford draws from, his assertions that the online catalog is "not a thing, but a perspective" (ie., a view of a database), and that it "should not be a computerized card catalog" but rather a source of dynamic "status" information.

Directory of Information Management Software for Libraries, Information Centers, Record Centers, 1987-88 Edition, Ed. by John Kazlauskas and Pamela Cibbarelli, Studio City, CA Pacific Information, 1987, 300pp., \$49. pbk.

Two page, systematic descriptions of 118 packages, a list of non-respondents, and indexes by vendor, hardware, modules and product name. The information should be adequate to decide whether to evaluate a product further, but more comparative tables would have been useful. While the scope is appropriate for the repositories in the title, archives and museums will find some products of interest here as well.

Dryden, Jean, "Subject Headings: The PAASH experience", Archivaria, 24, Summer 1987, p.173-180

The development of the Provincial Archives of Alberta Subject Headings is the most complete account we have of the challenges to archivists of modifying LCSH and the pitfalls and potential benefits of such an undertaking. The article, and the headings themselves, must be required reading for archival descriptive catalogers.

Dunlap, Ellen S. & Reed, Kathleen, "Borrowing of Special Collections Material for Exhibition: A Draft", Rare Book and Manuscript Librarianship, v.2#1 Spring 1987 p.27-37

This draft guideline, prepared for the Rare Book and Manuscript Division of ALA has uses for all archives and museums, and is being proposed as the basis for a standards development process.

LC Thesaurus for Graphic Materials: Topical Terms for Subject Access, Washington DC, Prints & Photographs Division, Library of Congress, 1987, 591pp. \$30. from CDS

Subject access to image material has long been a problem for libraries, archives and museums. This structured thesaurus, containing a huge hierarchy of terminology, will be of great assistance to anyone providing such access. Jackie Dooley's introduction to the volume stands alone as a

valuable statement of the ways in which such subject term authorities can relate to traditional library practice.

Overmire, Rozell, Records Management in Museum Travelling Exhibits: Planning a Computerized System, Masters Thesis, July 1987, John F. Kennedy University, San Francisco, CA, 119p. & appendices

A survey of computerization projects in the travelling exhibits field; rich in detail from which to develop requirements but short on advice, design considerations or analysis of the sources of successes and failures in the field.

Parker, J.J.N., "The Hull Domesday Project Database", Humanistiske Data, 2-87, p.4-22

An introduction to the work surrounding the analysis of the Domesday book as a machine-readable data file.

Rush, James E. & Tannehill, Robert S. Jr., "Bibliographic Data Elements for Computer Programs", Library Hi Tech, #14, p.79-93

I failed to cite this essential article in a previous discussion of cataloging of software, and hasten to correct my oversight.

Shillingsburg, Peter L.; Scholarly Editing in the Computer Age: Theory and Practice, Univ. of Georgia Press, 1986, \$11.95

The real of building archives and museum data and image bases is to use them to interpret our past. Shillingsburg combines practical advice with discussion of the premises of documentary editing and potentials for transforming texts by use of electronic techniques.

Sturges, Paul, "Policies and criteria for the archiving of electronic publishing", Journal of Librarianship, vol.19#3, July 1987 p.152-172

The author takes a broad view of the practices of librarians, archivists, data archives, media centers and others and makes recommendations, previously advanced in his report to the British National Bibliography Research Fund.

EPHEMERA

Annual Review of OCLC Research, July 1986-June 1987, Dublin, OH, OCLC, 1987, 53pp.

OCLC continues to conduct exciting basic and applied studies of networking, interfaces, automatic indexing and cataloging, image and information capture and retrieval, speech recognition, new media and a host of other issues of concern to cultural repositories. Their annual report is clear, informative and includes references to more detailed publications.

The Beckman Center for the History of Chemistry: Five Year Report, 1982-87, Philadelphia, PA, The Beckman Center, 1987, 36pp.

Neither museum nor archive, this disciplinary study center sponsored by the American Chemical Society and the American Institute of Chemical Engineers at the University of Pennsylvania is engaged in both enterprises, and in information exchange about the history of chemistry.

Guide to Fund Master Fund Raising Software, 1987 Edition, Master Software Corporation, 8604 Allisonville Rd., Suite 309, Indianapolis, IN 46250, 85pp.

Although this booklet specifically sells Fund Master software, its discussion of the issues involved in fund raising is an excellent analysis of functional requirements for this activity and a good place to begin defining local needs and assessing local systems.

Information Packet on Library Technical Standards, ALA, Library & Information Technology Association, 1987, \$3.50

LITA has collected into a pocket folder copies of ANSI, NISO, SISAC, BISAC, ALA and LITA publications brochures on standards. Whether this is useful depends on whether you have picked them up yourself at meetings or want to save the effort of writing these organizations.

"Lasers in Graphics Arts: White Paper" available free from Compugraphics Corp., Marketing Communications, 200 Ballardvale St., Wilmington, MA 01887, is a technical introduction to lasers in image setting (type setting), scanning and printing. It is straightforward and non-commercial.

Optical Discs for Storage and Access in ARL Libraries, SPEC Kit 133, April 1987, [Systems Procedures Exchange Center, Office of Management Studies, Association of Research Libraries, 1527 New Hampshire Ave., NW, Washington, DC 20036] contains optical systems planning reports from MIT and the Smithsonian, proposals for projects at Kent State, MIT, U. Michigan, Syracuse and Vanderbilt, progress reports from LC and MIT, and evaluations from Brown, U. Illinois and Vanderbilt, mostly from late 1985 and 1986. Useful for anyone pursuing similar ventures.

Resources for the History of Computing: A Guide to U.S. and Canadian Records, is available from the Charles Babbage Institute, 103 Walter Library, University of Minnesota, Minneapolis, MN 55455, for \$9. It contains approximately 350 entries.

Trinet and the 1987 SIC Revision is free from Trinet Inc., 9 Campus Dr., Parsippany, NJ 07054. It contains an overview of the Standard Industrial Classification system, its history and structure, and the 1987 revisions to the SIC, with only a modest plug for Trinet.

Visual Image Transmission: An examination of electronic delivery of visual images and text from the library to the academic community. Final Report to the Council on Library Resources, Merrill W. Smith, Patrick A. Purcell and Christopher P. Thorman, MIT, Rotch Visual Collection, Sept. 1986 26pp.

Report on experiments in image transmission and a variety of real and hypothetical configurations to support such efforts.

PROJECTS & PROPOSALS

Getty Conservation Information Network

The J. Paul Getty Trust has announced the advent of the Conservation Information Network, the result of more than two years of planning. Its three online databases feature access to technical literature, conservation materials and products & suppliers, as well as an electronic mail system enabling conservators to consult each other. The bibliographic database will provide access to almost 100,000 citations including ephemeral technical reports. The materials database provides access to 1,000 properties of materials relevant to conservation practice. Network access (including one account number, one E-mail box, and a full set of documentation) is \$50, and includes access to a help-line and the "Network News", a bi-monthly report on its services. Users pay \$40 per hour, prorated, for access including communications charges. E-mail users pay \$70 per 1K bytes sent or received. For subscriptions, contact, User Services, Conservation Information Network, 450-3 Glencoe Ave., Marina Del Rey, CA 90292

U.S. Government Information Inventory

The Office of Management and Budget (OMB) Bulletin 87-14 requires Federal Agencies to inventory their information dissemination products and services. NTIS is to collate these submissions into a government wide database available for sale. OMB intends to revise Circular A-130 to incorporate this as a permanent requirement.

Image Management Needs Assessment

New Jersey is conducting a Statewide image processing survey to identify requirements of administrators for traditional microform, unitized microform, CAR systems, electronic imaging systems, and COM. Contact Cesar Iacovone, Director, Div. of Archives & Records Mgmt., CN 307, 2300 Stuyvesant Ave., Trenton, NJ 08625

Scholarly Communications Office

The ACLS Newsletter, Summer 1987, announced that American Council of Learned Societies is closing its office of Scholarly Communication in Washington DC. The Office published Scholarly Communications and was a useful switching circuit. It, and Director Herbert Morton, will be missed. Meanwhile, ACLS has just published its first "Occasional paper", "A Life of Learning" by Carl E. Schorske, the delightful and fascinating text of his Charles Homer Haskins Lecture on April 23, 1987. Contact the ACLS at 228 East 45th St., New York, NY 10017-3398.

Canadian Strategic Planning

The Association of Canadian Archivists (ACA) has established a Strategic Planning Committee charged with submitting a report for consideration at the 1988 meeting in Windsor, Ontario. The Committee has circulated a questionnaire to ACA membership. A second questionnaire, on the potential role of the Bureau of Canadian Archivists, is being distributed by the ACA and by the Association des Archivistes du Quebec. Contact Colleen Dempsey, President and Committee Chair, ACA, P.O. Box 2596, Station D, Ottawa, Ontario, K1P-5W6, Canada

Canadian Archival Certification

The Bulletin of the ACA (Sept. 1987) contains the text of a discussion paper on certification drafted for the ACA. U.S. archivists will be interested in its similarity to (and difference from) their own recently adopted plan.

British Archival Description

The Archival Description Project at the University of Liverpool, which produced the Manual of Archival Description (MAD) in 1986 has been awarded funds by the British Library R&D Department to revise the rules to include in them description of such special documentary formats as title deeds, cartographic records, photographs, architectural and planning archives, sound archives, and moving image and machine-readable materials. It is also charged with

testing the suitability of MARC AMC in conjunction with "MAD2" and with exploring the implications of optical digital media.

The project team is arranging experiments, demonstrations and discussions with as many UK colleagues as possible and is desirous of communicating with others who have examples of cataloging instructions, in-house rules for description, or proposed standards. Contact Margaret Proctor, Fieldwork Coordinator, Archival Description Project, Archives Unit, University of Liverpool, Liverpool, L69-3BX, England

International Museum Database Survey

The International Council of Museums, International Documentation Committee, is conducting a survey of museum databases. To be included or see results, Contact Mary Case, Director, Office of the Registrar, Smithsonian Institution, Washington, DC 20560

Smithsonian Institution Information Architecture

An executive summary of the Information architecture project at the Smithsonian was issued in October 1987 by the Office of Information Resource Management. While the project has a distance to travel, its success to date is encouraging. For information about the museum collections architecture, Contact, Jane Sledge, Office of Information Resource Management, Smithsonian Institution, Washington DC 20560

U.S. Government Databases on CD-ROM

In addition to databases such as Medlars, formally published by Government Agencies, many agencies are now using CD-ROM for "internal" dissemination of information. The U.S. Patent and Trademark Office is distributing its "Classification and Search Support Information System" (CASSIS) in this fashion and using it in its reference rooms in Crystal City, Va. The Air Force is compiling a database of acquisitions related materials to publish on disc. NAVSEA is looking at CD-ROM for parts lists and servicing data for the fleet. The U.S.

Geological Survey is disseminating earth sciences data, and is readying a distribution of Global Digital Seismic Network data from the National Earthquake Information Center. The Census Bureau is exploring distribution of demographic data on cd's, while NASA's Voyager Planetary Data System and the National Oceanographic and Atmospheric Administration's database of solar flare activity are already available in this format.

Consortium to Develop Nineteenth Century Painting Image Base

Six major American art museums have received funding from the Pew Charitable trusts to develop a videodisc of paintings and works on paper created between 1860 and 1914 by European and American artists which they collectively hold. The Museum of Modern Art (NY), the Metropolitan Museum of Art (NY), the Brooklyn Museum, the Art Institute of Chicago, the Museum of Fine Arts (Boston) and the Philadelphia Museum of Art are cooperating in a project aimed at evaluating the potential of videodiscs for public education, database searching and authoring, and at building a foundation for future collaborative efforts.

Automation Surveys

Lisa Weber reports on the results of the SAA automation survey in the SAA Newsletter, November 1987. Rebecca Buck reports on the results of a survey of collections management software vendors in the fall issue of Registrar.

Smithsonian CD-I Publication Plans

The Smithsonian Institution Press has announced plans to publish a Compact Disc - Interactive (CD-I) version of Edwards Park's Treasures of the Smithsonian. The disc, created by American Interactive Media Inc. in Los Angeles, is to be available early in 1988. Narratives about 300 of the Institution's most famous artifacts are to be viewed under interactive user control, with branching provided for those desiring greater detail. Future titles are planned in what is to be a library of CD-I publications according to SI Press Director Felix Lowe.

SOFTWARE FORUM

Jon Reynolds, University Archivist at Georgetown University (37th & O St., NW, Washington DC 20057), writes:

"I started using *dbase II* in 1983, and have made quite a bit of progress. We use *dbase III Plus* and *Clipper* now, and I may get the chance to try out *Foxbase*. Naturally, I was interested in your statement about *dbase* in the last newsletter. Depending on the way you define powerful and generalizable archival application, you may be right, but I hope to prove you wrong. I am NOT making bets, or promising delivery dates, but I am hopeful.

My present system was designed to meet my most pressing need, which was (is) file folder access to current accessions. It works nicely, but is not space efficient, and is not a complete system by any means. I am enclosing a ROUGH PRELIMINARY diagram which indicates what I think is possible using *dbase* programs compiled with *Clipper*. Each box represents a *dbase* file. As you know, each file may have up to 128 fields (1024 in *Clipper*), 254 characters per field, 4000 total per record, exclusive of memo fields. The "iffiest" part of this is using memo fields for things like scope notes. In straight, unmodified *dbase*, the memo field is designed quite poorly. It is better if you substitute for the *dbase* editor, and better still in the *Clipper* compiler. How much better, and how well I will be able to implement it, remains to be seen.

The arrows on the diagram are a first attempt to illustrate how I expect to get things out. They are probably not as clear as I would like. The logic is: you enter at the left. Before entering you will have told the computer whether you wish collection level or folder level data, and whether you wish to qualify the term by some sort of MARC tag, and by form, and whether you want to see the results on the screen or in print. For example, you want a printout of folder level data containing letters written by Evelyn Waugh. So you are looking for "Waugh, Evelyn" as a creator, and you want folder

level data. The system finds Waugh in the lookup file, skips it if it isn't tagged as a creator, and passes to the folder level file if it is, and prints out, or displays, the appropriate data if found. Since you are using indexed fields, this would be found almost instantly, and would start printing or screen writing in about a second and a half.....If you wanted a photo of Waugh, the system would take you to a photo file, and so on. If you had asked for collection level data, it would take you to that file. I will make the collection level file as MARC-like as I can.

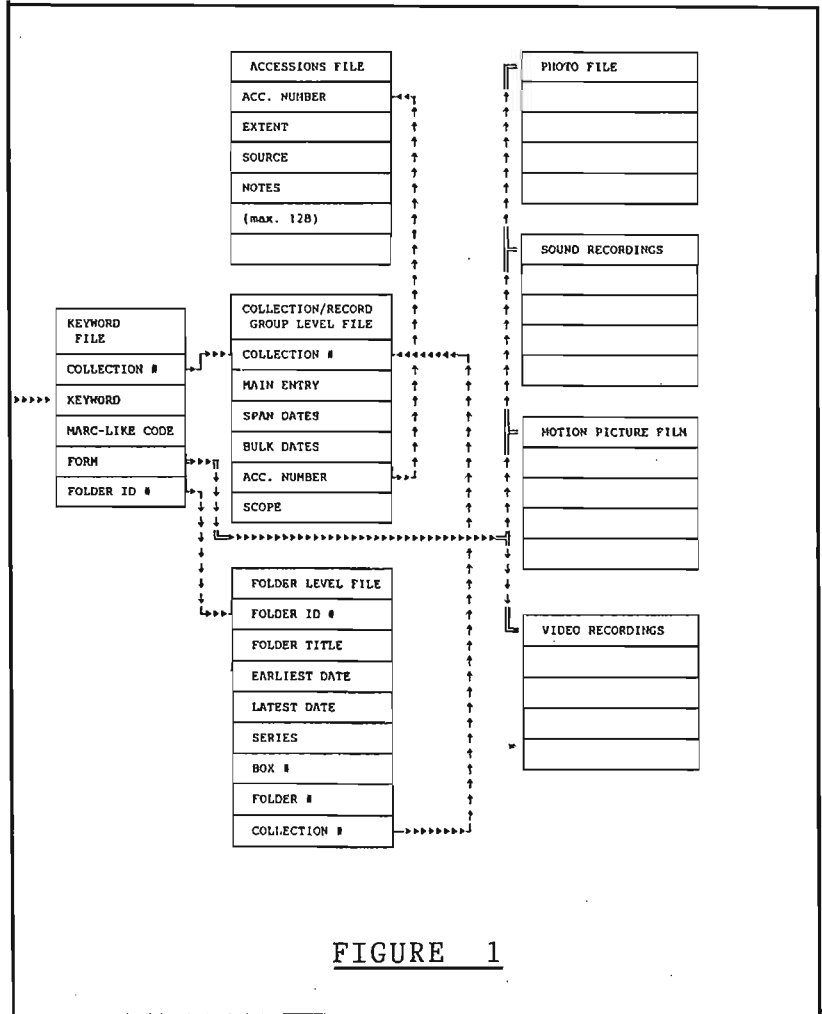


FIGURE 1

Now I think this will be clearly powerful; the generalizable part is more problematic. I am confident of the searching part. Idiot proof data entry will be a little harder, but not hopeless. I think it will prove possible to look terms up in a thesaurus during data entry, but we shall see.

If I do get it all to work, I'm not sure how to go about sharing it. I'm as poor as the next archivist, but I don't have time to set up a real business. I will probably continue to give away printouts of the code for the cost of copying and mailing, but will try to make some money on advising people how to go about putting it to use. If it ends up more or less fool-proof, I may try to figure out how to sell it for a modest sum.

P.S. Your readers may be interested to know that the SABA handscanner actually works! I was able to try one out for several weeks, and found that it reads Courier and several other typewriter fonts quite well, as long as the page is clean and the impression is good. It is not perfect, but it reads much better than I type. We didn't buy it because all our typed finding aids are in Bookface Academic; but for less than \$600 it is amazing."

Raimund Goerler, University Archivist, Ohio State University, (169 Converse Hall, 2121 Tuttle Park Place, Columbus, OH 43210-1169) writes that he has ordered release 4.0 of PROGRESS, but that using his current version, 3.2:

"... we continue to make, I think, significant advances in our application. As of this date, the data dictionary has been created and programs to add records at the collection level, the box level and the item level have been written and successfully tested. Most recently I wrote programs to search this database of accessions by means of keywords, thesaurus terms (from SAA's Thesaurus for College and University Archives) keywords and thesaurus, classification, and record i.d. Printed reports consist of an alphabetical listing of all keywords actually used in the database (with provisions for **see** and **see also**) and a listing of thesaurus terms

actually used. I decided to base the searches on the "BEGINS" command which uses indexes to search the database; rather than read each record. This will prove particularly useful in searches by hierarchical classification, in which I can gather all lower offices (e.g., a search by record group "31" would yield all departments of the College, e.g. Art, History, Dance, etc.). By the same token, we enter the thesaurus codes, rather than the words, of the SAA C&U thesaurus and can do inclusive searches by simply entering the first portion of code to gather narrow terms with broad ones.

As a test of the database and of my programs, I have entered some 150 records and searched them successfully. Data entry is as rapid as I could hope, particularly because in the case of file folder level entries we use a data file to read redundant but necessary information into the record, e.g., box number, folder number, collection dates, location.

At this point, I think I have only one more major task: the creation of menus.

After all of this work, I do think that we will no longer process in the traditional sense. What we are doing is making use of the inventories that offices provide, and entering the data directly into the database. I envision that processing would no longer mean consolidating accessions from the same office, weeding, and creating a new inventory. Instead, what is likely is that the weeding will take place and that the records in the computer will be adjusted."

Readers are encouraged to report on their own development efforts, comment on the software designs proposed by others, or suggest generalized requirements that these and other systems should meet. I will continue to run a forum, so that design issues can be aired. The forum is open to software vendors who wish to explain their system architectures as well. I may delete parts of discussions that are too lengthy or too promotional. Editor

SOFTWARE PRODUCTS

Advanced Projects International, Inc., [8000 Towers Crescent Dr., Vienna, VA 22180], a start-up firm headed by Hernan Otano who lead experimental projects in videodisc and digital imaging at the National Air and Space Museum, has announced four models of its *Digital Archival Retrieval System* (DARS). The basic components of these systems are image capture devices, storage systems, and information processing software. The image capture devices range from flat-bed scanners for uniform paper and photographic documents and over-head camera scanners for exceptionally large and small items, to automatic microfilm readers and video camera's for color documents, paintings and three dimensional objects. Storage is accommodated on magnetic erasable disks and optical WORM disks. The retrieval system is proprietary, but promises to extract text from images and permit output to high resolution laser printers.

Chadwyck-Healey Inc., [1101 King St., Alexandria, VA 22314], has been demonstrating a CD-ROM product that heralds a new way of presenting a wide variety of statistical information. *SUPERMAP* combines 1980 U.S. census data with maps of the U.S. to the county level to enable users to display relationships between pre-aggregated datasets cartographically. Susan Severson, President, expressed interest at a recent meeting in extending the product to allow local data sets to be displayed together with the census data.

Chadwyck-Healey has also announced February 1988 availability of "The Architectural Landscape of North America: An Image Bank on Videodisc" by Alan Gowans, consisting of 28,000 color images, indexed and captioned in two separate volumes.

Tri-Star Publishing [475 Virginia Dr., Fort Washington, PA 19034] is offering the Oxford English Dictionary (without supplements) on CD-ROM. Interactive,

random, access to the OED is an entirely new experience for dictionary-philes.

Vernon Systems Ltd. [P.O.Box 6909, Auckland, New Zealand] has announced *COLLECTION* a multi-user, PC-based application for collections management built on the *Revelation* DBMS by Cosmos Inc. that is consistent with MDA data standards. The first complete release was not available when I saw the system operating at the MCN meeting, but what I saw was extremely impressive.

COLLECTION supports a rich data structure with numerous data groupings (composed of "multi-valued" fields) that serve as authority data. Data entry is vastly simplified by these segments which can permit exception only data entry for much of a record. A true thesaurus capability, with both the use/use for and broader/narrower structures supports as many thesauri as the user wants. Indexes are constructed for 28 access paths. Navigation between any screens is a one step process and the user can return to the place of departure with a toggle. Maximum record length is about 65K per record and all fields can be variable in length and occurrence. Some unfinished aspects, such as security control, were not demonstrated, but other features, such as reporting and query, are handled by *Revelation* which has an impressive set of resident features. I look forward to the Spring 1988 release.

ABACUS Data Group Inc. [812 S.W. Washington St., Suite 600, Portland, OR 97205] recently sent me documentation of their Photograph Management System and Historical Society Membership System, software products aimed at the WANG-VS market. ABACUS also makes mailing list and accounting software aimed at historical societies, museums, and archives. The Photograph Management System features item or lot level description and multi-layer hierarchical subject assignment. The Membership System features member history, gift and affiliate society membership (third party billing) and interface to word processing.

STANDARDS

X.400/X.500

One of the silent events of 1987 was the acceptance by a large number of computer and telecommunications information vendors of the X.400 standard which is now several years old. X.400 is a standard for electronic messages which makes possible exchange of data between any systems. Acceptance was marked by two staged events - a vendor demonstration at the Hannover Fair in March 1987 involving British Telecom and the Deutsche Bundespost along with many of the world's largest computer companies and a demonstration at Telecom '87 in October involving the Swiss PTT, the French PTT and AT&T and many others.

X.400 would provide a world-wide electronic mail service by defining just how messages would be switched and relayed on the world's public telecomm services (providing an organizational infrastructure to carry out the potential transfers and implement the host of local, e.g., national, tariffs) and how the contents of an envelope would be described. X.400 can support all the existing telecommunications methods - telex, facsimile, teletext, videotex and straight data exchange and therefore promises to make electronic document interchange a feature of the 1990s landscape. Missing from the standard is a means of handling directories for remote addressees - an oversight being rectified in x.500, a U.S. standard being proposed to the CCITT early in 1988.

This one is very real, and almost upon us. Virtually all the major vendors have announced support and it is unlikely that any will be able to resist and survive into the 1990s. Archivists need to consider the implications uninhibited, universal, instantaneous electronic communications will have for work within their organizations and plan accordingly.

Using Standard Codes

I frequently find myself dismayed to discover yet another project making up codes for languages, countries, or industries

rather than using standard codes.

Presumably they wouldn't make up codes for states instead of using postal codes or for chemical elements? Perhaps they don't know of standards. If so, allow me to introduce the Standard Industrial Classifications (SIC) codes, a system managed by the U.S. Federal Government, which relates to U.S. economic and social data the way that census tracts relate to demographic data and electoral districts relate to political statistics...which is to say, you can't correlate anything with data which has not been so coded.

SIC is a hierarchical schema in which 11 one digit "Divisions" (Agriculture, Mining, Retail, Services etc.) are divided into 2 digit "Major Groups", three digit "Industry Groups" and 4 digit "Industries. The system is simple, comprehensive and powerful. A free guide to the 1987 update is available from TRINET Inc., 9 Campus Drive, Parsippany, NJ 07054.

In England, a Standard Industrial Classification used by the Central Statistical Office serves the same function. It has been incorporated into Social History & Industrial Classification published by the University of Sheffield in 1983 and available from the Museum Documentation Association which has incorporated its use into the Museum Documentation System.

Background & Projections

"Telecommunications and Information Systems Standardization -- is America Ready?" May 21, 1987 (87-458) written by David Hack, Analyst in Information Science and Technology at the Congressional Research Service contains a useful discussion of the background to international standards efforts and their potential economic and political implications.

CD-ROM Logical File Format

On October 5, the International Standards Organization passed the Volume and File Structure of CD-ROM for Information Interchange standard (ISO/DIS 9660) which grew out of the "High Sierra" group discussions and had previously been adopted as ANSI X3.B11. While access remains to be standardized, the format now seems stable.

AUTOMATED SYSTEMS FOR ARCHIVES AND MUSEUMS: ACQUISITION AND IMPLEMENTATION ISSUES

ARCHIVAL INFORMATICS TECHNICAL REPORT

Vol.1, #4, Winter 1987

Executive Summary:

Selecting and implementing an automated system is a complex and expensive process which few archives and museum directors have experienced. This report examines the technical requirements of each of the major types of automated systems installed in archives and museums. Specific technical features of membership and development, exhibits management, collections management, accounting and personnel (including volunteer) systems are reviewed. The emphasis in all cases is on analysis of options.

The report aims to educate the professional staff of an archive or museum who possesses the expertise to define functional requirements for their information systems, but who have little or no experience defining technical and performance requirements or selecting hardware and software systems. The report is intended to be used as a workbook. It presents a step-by-step guide to writing a request-for-proposal and developing evaluation criteria with which to select the system that comes closest to meeting institutional needs. It includes checklists of features, forms which can be used in evaluating systems and vendors, and project planning guidelines for the entire system life-cycle. And it discusses the issues which a cultural repository will face in implementing automated systems.

In addition, the report identifies important implementation issues and suggests methods for staff to assess implementation impacts and to plan realistically for installation, training, operations, maintenance and replacement. While this report will not obviate the need for additional professional help, it will prepare staff to provide on-going support for the automation of archives and museum information systems.

Available in January, 1988. Archives & Museum Informatics, 5600 Northumberland St., Pittsburgh, PA 15217. Separate copies \$45 pre-paid, includes handling, or by subscription to Archival Informatics Newsletter & Technical Report (\$160 p.a. includes all reports for the year plus the Newsletter; \$180 p.a. foreign subscriptions, includes airmail.).