ARCHIVAL INFORMATICS NEWSLETTER

Part 1 of Archival Informatics Newsletter & Technical Reports

ISSN 0892-2179 Summer, 1987 Volume 1, * 2

THIS ISSUE

Actually, Canada has been just north of us for a long time, even though it may seem from this issue of the Newsletter that we just discovered it!

As reported throughout this issue, archival automation activity in Canada has reached a crescendo recently. Tom Brown's column (p.5-7) reports on the exciting I-ASSIST and ACA meetings. The article by John McDonald (p. 14-15) explores the challenges being faced by the new National Archives of Canada and declares the rights of archivists to be enfranchised in the requirements definition processes for information systems serving all activities within their organizations. On pages 15-16, I introduce the Planning Committee on Descriptive Standards and their plans (what my Smithsonian colleagues were fond of chiding me by calling "plan-plans").

But Canada doesn't have a monopoly on debates about descriptive standards. On pages 9-13, Lisa Weber, the automation officer of the SAA, presents a discussion of the confusion and conflict surrounding description of microforms (and, by extension, any "copies" in another medium) within the framework of the MARC Formats for Bibliographic Description. If the details tend at times to be hair splitting, it reflects the problem of trying to define a specific descriptive cataloging rule to meet the needs of a variety of user communities. No single, "logically consistent" view seems to do for all.

This, of course, is the challenge faced by museum information professionals whose situation is the subject of my reflections on the AAM meeting which was just held in San Francisco (p.2-4). Amidst some promising developments in information standards and some tentative offerings of computer software, the chaos of Babble, each with his or her own language, reigned. And yet, there was a sense of promise in the air, of a new beginning

I am delighted to be running letters to the editor. The interest, controversy, and information exchange potential of the first issue is immensely gratifying. I hope that the requests for dialog issued by each of the authors of pieces in this issue will find an equally responsive audience.

The contributions I received to this issue are doubly welcome. First because they are challenging in themselves, and are the kinds of opinion pieces which would not have found their way to print except through a vehicle such as this, and second because they demonstrate a need for the Archival Informatics Newsletter more than anything I can say.

The first issue of Archival Informatics Technical Reports, on Optical Media in archives and museums, was published in May. Issue 2, on Collecting Software in archives and museums will be available in August. The fall issue will be devoted to the requirements for collections management. It will discuss not only the information needs of collections managers and define software systems to meet them, but also examine the relationship between collections management and other institutional missions.

TABLE OF CONTENTS

Museum Automation at the AAM
by David Bearman

I think we will come to see the year that the American Association of Museums met in San Francisco as a watershed in museum automation. Although there was a certain immaturity exposed in the balance of sessions (only one session devoted specifically to mainline automation concerns out of four overall, with the others dedicated to CD, digital color imaging, and artificial intelligence), automation emerged as a full-fledged partner in the professionalization of registrars and collections managers. The same immaturity was evident on the exhibit floor, where systems accessing videodisc or digitized images attracted more attention than a five digit numeric field in a database deserves. Yet some of the systems being displayed were practical and could be implemented essentially as is. And although it was clear that most AAM members are still lost in the present of automation, it was evident that an adequate number can not only see the outlines of a future but are taking actions to assure that it evolves according to a professionally dictated plan.

Sessions:

While many attendees will long remember the extraordinarily crisp color images presented by Howard Besser and the experimental digitizing systems at the University of California, or the delightful tour of artificial intelligence pastures in Stuart Dreyfus' keynote address, the museum profession will be influenced over the longer run by the dialog begun in three sessions devoted to discussions of evolving information exchange standards.

The one session explicitly devoted to museum information standards, chaired by Deirdre Stam, focused the issues. Toni Petersen (editor of the Art and Architecture Thesaurus) spoke to the very concept of a standard as something supported by numerous agencies and not subject to unilateral change. Recognizing that standards are expensive to maintain, she insisted that they are essential in our complex society. She then presented the AAT as a potential vocabulary standard. Jim Blakaby reported on the efforts of a committee involved in revision of Chenhall's Nomenclature, a topic to which he had previously addressed himself in a session devoted strictly to the revision process. He presented Chenhall as a classification system, which could be a standard while being open to local addition of specific terms at the lower levels. Angela Giral, of the Avery Art Index, noted that standards are only important if we want to communicate with anyone, and then reported on how the project she directs, AVIADOR, has make use of both AACR2 (library community cataloging standards) and the AAT. Eleanor Fink, reported in the use of MARC by the Index of American Sculpture and the development of shared data definitions between that project, the Index of American Painting and several other art database projects at the Smithsonian National Museum of American Art. The MARC format was serving well as a means for inter-system data sharing.

Lenore Sarasan of Willoughby Associates was to present the "devils advocate" position. But as she argued that standards could be developed by users of commercial systems, and presented the work her firm was doing in developing common data dictionaries and authority files for clients, it was evident that this sphere of standardization did for those users exactly what profession-wide standards might do for all museums. Lenore revealed that any argument about the need for standards in the museum community is no longer about ends, but about tactics.

Two promising tactical developments were featured in sessions. The first, already briefly mentioned, is the revision of Nomenclature by a committee (albeit self-appointed) representing a range of museums which have used the classification system . More important than the second edition itself (although it will improve on the first), is the self-conscious discussion which the group generated about the purposes of an hierarchical classification system and the difference between those ends and the aims of an indexing vocabulary. While the discussion at the session itself did not resolve the purpose of the new edition, it paved the way for increasingly sophisticated discussions of the aims of different types of standards.

The third session was a report to the community by participants in the spring Conference on a Common Agenda for History Museums, sponsored by the Smithsonian Institution. The meeting was held to identify actions required in four areas: collections, collaboration, interpretation and documentation. The working groups will be continued under a broader umbrella of the AASLH in the next year, but their two day output is itself exceptionally promising, especially in the difficult area of common databases and documentation where the group launched a survey of the data fields currently used by history museums as a first step towards normalizing the data in a data dictionary which could help museums to plan for common documentation. The modesty of the effort bodes well for its ultimate success, as

does the openness of the subcommittee to a permissive data standard. The conference report will be published in full by the AASLH this fall.

Systems:

Six museum information retrieval/cataloging systems were exhibited at San Francisco along with numerous membership/development, financial, ticketing, shipping and other applications. With the exception of a new offering from QL Systems (The Volunteer Management System), I will restrict this report to comparing these six.

First, because it is too good to pass up, I want to share my enthusiasm about the QL product ($399; $299 from AAM to AAM Members). The Volunteer Management System is, simply, the best human resource management package devoted to this tedious but critical function I've ever seen. It performs in an elegantly simple fashion, retaining information abour your volunteers and their skills, the availability and the cumulative experience, reporting with ease on who can fill a need. It even sends year end thank you letters with total hours neatly merged using a word processing function. Most archives and museums I know of can use this IBM/PC based product as is, and have needed it for years.

AAM exhibit goers could try out ARGUS (by Questar), ARTIS (by the Williamson Group), MYMSY (by Willoughby Associates), STAR (by Cuadra Associates), STIPPLE (by Erros Computing) and a U.S. National Park Service D-Base III system. [Sounds a bit like the monkeys got my keyboard for a moment there.] In addition, they could hear about MILAM and QUEXIS (both by Willoughby), but not use the actual products because QUEXIS is not yet developed and MILAM was not on-line to the exhibit floor.

The most interesting observations about this crop relates not to what they are, but to what they all aren't. None of them are essentially collections management systems. They all lack some minimal functionality in this respect (e.g. ticklers, generalized collections actions statistical reporting and life-cycle tracking). Instead they are information retrieval, or cataloging, systems. They tended to be thin for experienced users - none of them provide a direct command interface in addition to menu driven capabilities, none provided for easy record redefinition or systems administration within the confines of the application itself (features present, for instance, in the public domain ILS or MINISIS packages). Nor are any capable of processing or generating MARC records to interface with library based systems. None yet has an installation larger than about 100,000 items, the size of a tiny natural history collection or a modest history or archeology museum. Only the US Park Service DBase III System, which is too limited and unsupported to be taken seriously as an option for others, and the Oracle-based MYMSY (Willoughby) are built on commercial DBMS's. The others have a few startling weaknesses as a consequence.

At present the vendors of museum software products are not doing a very good job of differentiation themselves, in part, it appears, because they do not know the capabilities of their competition. Each claims to be better able to support a museum because of their experience in museums, but none can point to more than a handful of installations of their present system. There is a definite tendency to mislead naive users about the virtues of operating systems (Pick vs. Concurrent DOS, vs. MS-DOS vs. S-38 OS) which are irrelevant to evaluation of the applications. Questor made a splash with the videodisc link in ARGUS (they can have a five character field and a cable to a videodisc!) and Willoughby wowed the crowd with its digital images in the QUIXIS demonstration file (using Picture Ware for imaging in a demo which had no other real functionality). Obviously other vendors could have done the same. I was encouraged by one important shift - while some vendors demonstrated their systems on PC's, the commercial applications are designed to run on mini-computers for multiple users.

ARGUS, STAR and STIPPLE showed thesauri and Willoughby promised one in Super Mimsy and QUIXIS in the fourth quarter. This is encouraging, although must museums are probably not generally ready to use a thesaurus well.

Both Willoughby and Questor emphasized the importance of quick entry for retrospective data and their ability to custom make a data entry screen for clients (although neither demonstrated any functionality in this respect not shared by STAR and STIPPLE). ARTIS, apparently aware that its eight screens could inhibit data entry severely, promised to reform. What is more interesting to me is the insistence on the part of museum oriented vendors and museum staff on retrospective conversion, rather than on simply starting up the system and using it progressively. This reflects the fact that the systems are not very capable in collections management (and that information retrieval functions are suspect with incomplete databases) and the assumption made by museum staff that all the records of the museum will be (should be?) automated, rather than simply "pointed to" by the system. While both these assumptions seem to me invalid, this issue deserves a full treatment. I welcome comments for a future issue. I will address this

further in the Fall Technical Report, Requirements for Collections Management Software. A comprehensive comparative analysis of museum and archives software packages is scheduled for Archival Informatics Technical Reports, Spring 1988 (v. 2 #1).

VENDORS:
Cuadra Associates, Inc. (STAR)
11835 W. Olympic Blvd., Suite 305
Los Angeles, CA 90064 (213-478-0066)

National Park Service
Curatorial Services Division
P. O. Box 37127
Washingtion, D.C. 20013 (202-343-8138)

Questor Systems Inc. (ARGUS)
844 Colorado Blvd.
Los Angeles, CA 90041 (213-258-5174)

STIPPLE Database Services
Warren Farmhouse
Thame Lane, Culham
Oxfordshire, UK OX14-3DT (44-235-24676)

The Williamson Group (ARTIS)
129 Mount Auburn St.
Cambridge, MA 02138 (617-497-6848)

Willoughby Associates, Ltd. (MIMSY & QUEXIS)
2800 Sheridan Place
Evanston, IL 60201 (312-328-3284)

THE ARTFL SCANNER AND TEXTUAL DATA ARCHIVES

A recent announcement in the ARTFL Newsletter that the long established project on American and French Research on the Treasury of the French Langange had acquired a Kurzweil scanner and was offering ASCII encoding of printed or typed texts as a free service to members, suggested a number of opportunities for fruitful, if not artful, cooperation between archives and scholars interested in text analysis. It also served to remind me of the risk of overlooking textual data archives, and increasingly graphic and even performing art data archives, in our focus on the more traditional social science and government data archives.

ARTFL is a textual database of 17th - 20th century french language texts in literature, philosophy the arts and sciences, maintained by the University of Chicago and the Centre National de la Recherche Scientifique. The database, which now consists of over 150 million words, was initiated by the French Government in 1957, as as step in the creation of the Tresor de la Langue Francaise, a new dictionary. Since 1984, ARTFL has been organized as a consortium which can be joined by any degree granting institution ($750 p.a. for PhD granting and $400 for others); more than 20 major American Universities are members of the program, contributors of texts, and users of the ARRAS, (Archive Retrieval and Analysis System). ARRAS, a full-text system currently being converted from an IBM3081 to UNIX, does not analyze a text in the sense of interpretation, but it does provide statistics on word occurrences, concordances, occurrences within contexts, and indexes. It can produce graphic distributions of results and do proximity searches.

Materials in the database range from troubador poetry and lyric poems in the old Provencal language, to a body of texts from the revolution of 1848 and modern works by Apollinaire, Bonnefoy and Meschonnic. It has been used for research into neologisms, the significance of certain streets in Parisian novels, the origin of concepts such as "opinion publique" and "administration" in intellectual and political history, and for teaching. Recognizing that the database could be more useful to scholars, ARTFL is currently rewriting the search systems to permit its use on smaller systems and cooperating with the Textual Information Retrieval and Analysis (TIRA) project at the University of Chicago in a cooperative software development effort.

The acquisition of the Kurzweil Optical Character Recognition device, which was paid for by the Packard Foundation, is part of this effort by ARTFL to serve its community in new ways. It seemed to me that this kind of project has potential for bringing those special items held by archives, items of historical significance, to the attention of researchers who can make use of them in new ways. As we begin to consider the role of primary materials in teaching and scholarship in the "hypermedia" learning environment many major universities are planning for the 1990's, the role of textual databases (along with the image bases and sound bases of museums), will become more critical. Machine-readable does not just mean social science.

[ARTFL, Department of Romance Languages & Literature, University of Chicago, 1050 East 59th St., Chicago, IL 60637]

David Bearman

NORTH OF THE 49th PARALLEL
By Thomas E. Brown

Events during the last few months have reinforced the leadership image of Canadian archivists in the management of automated information systems.

I-ASSIST

As it does every four years, Canada hosted the 13th annual I-ASSIST conference at the end of May. The International Association for Social Science Information Service and Technology, is the professional association of data archivists and data librarians. As one would expect, four days of workshops and formal sessions covered a variety of topics important to archivists in traditional institutions who administer computerized files. Topics included trends in the use of machine readable data, data interchange standards, advances in storage and dissemination technology, and tools to train data archivists.

The keynote speaker, David Nasatir of California State University--Dominguez Hills, outlined the history of the social science data archives. He contrasted two philosophies which have competed since the mid 1960's when these archives were beginning to flourish. One school argued that the purpose of the data archives was to facilitate and promote access. This group went so far as to propose that all data created with government grant or contract funds should be available to the research community for the cost of duplication. The other approach concentrated on trying to insure the proper use of the data. This meant restricting the information to social scientists associated with major research institutions, since only they would be in an environment in which the data could be properly analyzed. The tension between these approaches led to the demise of the Council of Data Archives, the first organization concerned with data archives, whose members were North American repositories and whose purpose was to establish standards for archives of computer files. I-ASSIST emerged to replace the Council, as an organization of individuals concerned with data archives and their operations. Nasatir argued that I-ASSIST has been instrumental in the progress which data archives have made in the last two decades. During this time, he noted: (1) the expansion of the scope of research use of materials from data archives, (2) increased use of data archives in social science research projects (3) additional ways of disseminating information, and (4) standards for data description and data interchange. All of these trends have improved access and thus seem to have decided the early turf war in favor of those arguing that a major purpose of data archives is to facilitate access. In this connection, Nasatir stated that there has been a convergence of data archivists and traditional librarians. He concluded that the influence of the librarians' inviolable commitment to access underpinned the progess which data repositories have made in promoting access to computerized information. Despite the keynoter's acknowledgement that other unnamed information professions had been involved, session chair and Association President, Judith Rowe of Princeton U., noticed some consternation among the audience and twice explicitly acknowledged those outside of the library community who have helped data archives advance. Following Nasatir's presentation, she noted the role of the SAA Task Force on Automated Records and Techniques and in her conference summary, Rowe detailed the contributions traditional archives have made to the data repository effort. She specifically mentioned appraisal criteria, standards for research use of administrative files, and preservation.

One of the conference workshops focused on CULDAT, the Canadian Union List of Machine Readable Data Files. Paula Mitchell and Edward Hanis of Tycho Research, Associates described CULDAT as a nationwide conputerized inventory of machine-readable data in Canada, created as a source for information products and to help researchers identify and locate machine readable files. CULDAT will support an on-line service available nationally and a reference periodical. CULDAT has standardized its record format based on the MARC format for machine-readable data files, developed cataloging rules with authority lists, and constructed an online cataloging system. Tycho estimated that up to 10,000 records will be in the database when completed.

The final of the program session was a 150-minute marathon of five papers. Organized by Sue Gavrel of the National Archives of Canada, it presented a glimpse of the successes and failures of traditional archives in dealing with automated information systems. In the first presentation a records manager for the government of British Columbia, Rueben Ware, outlined his aggressive program to effect the economical destruction of computer records of temporary value and to identify those records which have archival potential. Not suprisingly, the Province adopted the "systems approach" to inventorying and scheduling which the national repository has pioneered in Ottawa. Ware discussed some creative means to get the information system managers involved with records management but the

program illustrated a common disjunction between ideals and realities. The Provincial Archives may be able to identify records with archival potential, but British Columbia does not have a program to accession them. Thus despite the productive efforts of the records managers to inventory and schedule the disposition of the components of the province's automated information systems, conputerized materials with archival value may still be lost. In her paper, Margaret Adams took up the archival management of automated information in the Kentucky project which was discussed in my last column. She again called on records managers and archivists to rethink approaches developed before the advent of electrostatic copiers and microfilm, let alone computers. In the earlier era, records were static and information sharing was not possible. Since technology has advanced, she urged the archival profession to focus on information sharing in the present rather than reference service in the future. However, she had to report that no action has yet been taken about her recommendations for Kentucky to move in those directions. In the third presentation, William Deimer outlined his efforts to establish a municipal data library for the City of Los Angeles. Beginning around 1970, the Community Analysis Bureau within the municipal government acquired a variety of computer files relating to Los Angeles. Federal Census material was supplemented by a materials from city government, including tax information, crime reports, ambulance runs, morbidity, natality, etc. From a reservoir of over 1500 data tapes in the city's central computer center, the Bureau provided analysis for policy and administrative decisions. But in 1982, the Bureau suffered a severe budget reduction. As a result, use of the data declined. Because of the infrequent use, the data files were routinely blanked until none of the city's records remained. When Deimer contacted the municipal archives for assistance to prevent the destruction of the data, the response was sympathetic but ineffective. Undeterred, he is now working to establish a Statistical Bureau within the city government to acquire and analyze data related to Los Angeles.

After these three disheartening reports, Jorn Leipart of the Norweigan Social Science Data Service provided a welcome balance. He reported on his organization's effort to acquire and disseminate data produced by the Norway's Regional Development Agency (RDA). This agency provides economic assistance to businesses interested in locating in rural areas of the country and, in the process, acquires a variety of information on these businesses and their activities. While the information in computerized form was of great interest to social scientists, it was not easily accessible. First, some of the data remained in the hands of RDA's contractor and was unavailable for analysis by the agency. Since the data was organized for administrative purposes, it needed to be reorganized and adapted to different kinds of software for statistical analysis. Furthermore, it was evident that the information would be more valuable if it could be linked to Norway's Central Statistics Bureau. Leipart outlined how his data library acquired, reorganized, and linked the RDA data into a valuable information resource. In the final paper, Mario F. Lopez-Gomez of the National Archives reported on the researchers interested in the computerized datafiles accessioned by his organization. Using survey data for the period January 1985 to June 1986, he reported that most researchers did not write but used the telephone to obtain information about the records in the National Archives. The focus of attention was on financial, transportation, and military records. Finally, he reported that his researchers were equally divided among other government agencies, academic institutions, and private firms. While it would be quite unusual for most archives to have one-third of its clients come from private firms, it is not for repositories of computerized materials.

This session had a variety of themes and counter themes, currents and countercurrents. Each paper outlined a different function of the archival administration of computerized information -- inventorying , scheduling, accessioning , processing, preserving, and disseminating machine-readable data files. These different functions were reported from different perspectives-- records manager, archivist, policy analyst, and academic social science researcher. And the five presentations focused on different levels of government : two from the national level, two were from the state or provincial level, and one from the municipal level. Yet in this diversity, a coherent whole emerged. The first three papers candidly reported that while individuals can create a foundation for the archival administration of automated information systems, these will come to nought if, as in these cases, the responsible archival repository seemingly fails to respond. Just when we began to wonder whether to bother, the final two speakers demonstrated that with proper institutional support, the information had research value and could be of interest.

ASSOC. OF CANADIAN ARCHIVISTS

During the last four or five years, countless formal and informal meetings have discussed holding a conference on archives and automation. I have participated in no less than six or seven such meetings. Invariably, someone will comment that the last conference devoted to archives and automated records systems was held in Ann Arbor in 1979. The Association of Canadian Archivists changed that. At its annual meeting during June, every paper of every session concerned either the archival administration of automated records or use of automated techniques for the control of archives. We can now say that the last meeting devoted to archives and automation was the ACA conference in Hamilton, in 1987.

Under the theme "Archives and the Information Age", the program explored the implications for archivists, both in theory and in practice, of new, repidly changing information technologies. In the opening keynote, Hugh Taylor argued that the changing nature of information technology has transformed the archives through its impact on culture, records, the computer, the researcher, and the archivist. For the next three days, every paper of every plenary and concurrent session discussed Taylor's transformations and practical responses to them.

One theme which ran through several sessions was information resources management (IRM). Presentations examined whether or not archivists have a role in information resources management and if so what should that role be. This discussion exposed the disparity of meanings of the term, and while no clear definition emerged during the conference it was clear that working definitions of IRM were reshaping policy. Peter Gillis, of Canada's Treasury Board, outlined how the Canadian government was revamping its government-wide information policies after deciding that "the first crucial decision that we had to take was that most of our current policies in this area are obsolete. This was somewhat difficult to admit since the most recent policy, records management, dated only from 1983." (How many of us with records management responsibility will conclude that our current policies are likwise obsolete as a result of the information age?) A representative from Canada's Department of Energy, Mines and Resources outlined how she had implemented the new requirements into a system for control of the information resources within the department. Her theme was echoed in another session entitled "Automating the Life Cycle." In it, the speakers outlined automated records management control systems -- from federal, municipal, and corporate perspective. While each of these control systems implemented Gillis' prerequisite that organizations know what information is under their control, not one of the systems had incorporated their respective archival repositories into the design of the control systems. As one person observed from the audience, automated control of the life cycle had stopped at the archives' front door!

Sessions covered a variety of other topics dealing with automated records. One session presented case studies by two archivists involved in the appraisal, acquisition, and preservation of valuable information generated in automated offices. Their perspectives were dramatically different; one shaped by a large corporate structure and the other by the needs of a small operation. I participated in a session on appraisal in the information age. Five separate sessions dealt with automated control of materials in various sizes of institutions. These dealt with both microprocessor and mainframe systems, strategic data design information system planning, and approaches to networking and descriptive standards. One interesting sessions on archival automation asked how researchers view such systems. Do our systems meet the users' needs? Do the users even care?

The program committee collected about half of the papers into an informal set of proceedings. For information on how to acquire a copy, write: Association of Canadian Archivists, P. O. Box 2596, Station D, Ottawa, Ontario, K1P 5W6.

ACA FOOTNOTE: REPORTS AVAILABLE

In his ACA paper, Jay Atherton (Director General, Historical Resources Branch) reported that the Public Archives of Canada has been experimenting with ways to control the disposition of records in automated information systems and has tested several pilot projects. The findings indicate a need to take a fresh look at how records are scheduled and suggest a methodology based upon a system approach: analyzing and scheduling all data and information used and produced within a definable administrative process. A report and other information on these pilots is available from John McDonald (as reported by him, pp.14-15]. Also available through Mr. McDonald is a fascinating study on interchange standards for a variety of different purposes and for records containing a variety of different types of information. It is called "Data and Document Interchange Standards". This study, produced under contract to the Canadian government, paints a clear portrait of the complex family of interchange standards.

LETTERS TO THE EDITOR

Frank Burke (NARA) wrote on April 30, that in spite of my reservations about what OPTIRAM claims to be doing (v.1 #1)

"The National Archives has been impressed with the results of its tests of OPTIRAM on a limited number of documents that it has subjected to the process. These consisted of holograph letters, 19th century ships passenger lists, printed forms with handwritten insertions, typewritten 3x5 cards in French, and a variety of typed, printed and handwritten documents. The transcription accuracy rate was very high. The Archives is now trying to find out if anyone else has developed a comparable system for scanning these mixed formats." [On May 29, NARA released Solicitation NASP-N2-P-0046 for an "Indefinite quantity contract for optical character recognition for handwritten materials", with bids due on 7/29/87, so we may find out whether OPTIRAM or anyone else has the technology it advertises. ed.]

Elizabeth Betz Parker (LC) wrote on May 17 that the "LC Thesaurus for Graphic Materials: Topical Terms for Subject Access, compiled by Elizabeth Betz Parker, introduction by Jackie M. Dooley, will be available (I hope) in late June. 617p."

"The Prints and Photographs Division has begun entering records for groups of photographs in the MARC format... The records will be distributed to subscribers of MARC tapes and will therefore be available in RLIN and OCLC. (However, none of our records have been verified yet so as to initiate distribution. Should be pretty soon)"

Glen McAninch (KY State Archives) wrote on June 16 that he would welcome contributions of references and articles to be used in compiling a revised and updated bibliography on archival automation incorporating some citations (perhaps as many as 300) from Richard Kesner's previous bibliographies and all post 1983 materials. Glen can be reached at the Kentucky Department of Libraries and Archives, P.O.Box 537, 300 Coffee Tree Rd., Frankfort, KY 40602

STANDARDS

Common Command Language:

Louise R. Levey reports, in the Bulletin of the American Society for Information Science v.13 #5 (June/July 1987) on the status of the proposed standard for a common online command language. The proposed language, which uses the typical verb-object structure of most command languages (DISPLAY, FIND etc.) contains 20 primary commands and the syntax rules for them. The standard will be submitted for review and vote shortly. If it passes, a future in which a user could reasonably expect to search a variety of databases without having to learn each of their command languages is imaginable. To review the standard, contact Pat Harris, Executive Director, NISO, National Bureau of Standards, Administration 101/Library E-106, Gaithersburg, MD 20899.

MARC FORMAT:

At its June 1987 meeting, MARBI approved the SAA proposal for defining a subfield in 851 (Location) and 853 (Location of Originals/Duplicates) to permit sorting repositories by country or state/province of the U.S., Canada, U.K. and Soviet Union. It accepted a proposal to drop some form of reproduction and media codes in 008/22 and 008/23, keeping those which indicated how the information was to be played back or read. In a move of importance to archives and museums, MARBI adopted, with modifications, a proposal to use the 583 (Actions) field to record information about preservation actions. Now that libraries have adopted this approach to collections management data recording we can hope to see expanded support for it by the bibliographic networks and extensions of the concept in local software systems. Two days of preliminary discussion by MARBI of the complex proposal for "format integration", produced equally preliminary consensus around the proposal to view "seriality" and "archival and manuscript control" as ways of looking at cultural materials which are distinct from bibliographic item description. Further discussion can be expected in January, but I see little propect for a final agreement before July 1988.

DESCRIBING MICROFORMS AND THE MARC FORMATS

A Discussion Paper

Lisa B. Weber, Society of America Archivists

I. INTRODUCTION

Both archivists and librarians are having difficulty using the MARC family of formats to describe microforms. Some of their difficulties reside with the format. But, underlying the technical problems are deeper confusions and thornier issues which need to be addressed before we can develop satisfactory solutions.

Libraries and archives approach microforms differently. Ultimately, the reason lies in the fundamental concepts which distinguish archival from library materials. For a discussion of microforms, we need to make two distinctions; library and archival materials differ both in the intention associated with their creation and the methods we employ to control them. Library materials result from deliberate, intellectual acts, with a creative purpose. Archival materials are the residue of disparate activities which often span long periods of time. Library items are usually in a single format. Archival materials are often collectivities comprised of materials in a variety of formats. This paper outlines the issues as a first step towards a solution which both the archival and library communities can accept.

Historically, what distinguished archival from library materials, has been the concept of publication. Librarians collect items that are published and exist in many identical copies while archivists and manuscript curators collect unique records created during the course of daily activity. These "clear" distinctions are becoming increasingly obscured by technological changes in the production and distribution of information. And, although archival materials are usually unique, modern collections often contain "near-print" and other published materials such as books written or collected by the creator of the collections.

Yet, answering the question of whether the materials in hand are published or unpublished (library or archival) is crucial to using the MARC family of formats for describing microforms. Only when this slippery question is answered, can we ask how to use a specific format (AMC, Books, or Serials) to describe the microform.

Although the focus of this paper is on microforms, these same issues present themselves with respect to material copied by other techniques, including xeroxing. Archivists need to be able to use the MARC formats to describe these kinds of materials consistently, and archivists and librarians must face these same descriptive issues associated with technologies, such as machine-readable magnetic records and the various optical disk formats. The National Archives and the Library of Congress are already experimenting with optical disk technology as means of preserving information. And, an increasingly large number of indexes and publications are being distributed on CD-ROM. Therefore, any solutions to handling microform within the MARC family of formats must take these, and future, technologies into account.

II. MICROFORMS: LIBRARY POINT OF VIEW

Not surprisingly, the kinds of microforms librarians encounter are primarily "micropublications" or microforms that are created for wide distribution. Their emphasis is, therefore, on bibliographic control. Unfortunately, changes in the cataloging rules have confused the issue of library microform cataloging.

The Anglo-American Cataloging Rules, First Edition (AACR1) established the principle that a microform was to be described in terms of the original work, so microform publication details were relegated to a note. This rule assumed that microforms cataloged by libraries were primarily copies of already existing published entities. The second edition, AACR2, took a different tack. Since a microform requires special equipment for its use, under AACR2 rules microforms are regarded as a special type of library material. AACR2 rules require that the cataloger describe the microform in hand and include information about the

original item in a note.

The AACR2 approach to cataloging microforms provoked much controversy in the library community. To understand why, it is important to know that librarians distinguish between reproduction and original microforms. A "reproduction" is a microform which is a copy of a pre-existing bibliographic entity (i.e. a book or serial). An original microform is more difficult to define. Glenn Patton, of OCLC, explains that original microforms include items or collections without a previous bibliographic identity. For example, items brought together specifically for the purposes of producing a microform publication would produce an "original microform."

The Library of Congress took an official stance contrary to AACR2. Its rule interpretation 11.0A states that for microforms that are reproductions, LC continues to follow the AACR1 principle of describing the original in the title and statement of responsibility, edition, publication, distribution and physical description areas. Information about the microfilm is placed in field 533 (reproduction note). For original microfilm, LC follows AARC2. The motivation behind LC's position stems from the benefits of derivative cataloging and the economic inefficiency of AACR2 for "reproduction" microforms. Following AACR1 allowed the cataloger to use the orginal cataloging record to derive a new copy, and merely add a note. Following AACR2 would require that the cataloger create an entirely new record or modify the record of the original work extensively.

Whether they follow AACR1 or AACR2 rules, library catalogers must record the same information in the record. The difference in approach will be apparent only for "reproduction" microforms.

MICROFORMS: THE ARCHIVAL VIEW

Archival concerns vis-a-vis microforms are very different. First, since archivists administer "unique" materials, derivative cataloging has never been an issue in the archival community. Related to this situation is that fact archivists have not, until recently, been concerned about universal cataloging standards. Previously, each archival repository "cataloged" its collections as it saw fit.

But the primary difference is that archivists view microfilming as a process in the life-cycle of information, as such, microform is a means to an end (or a tool), not a new item to catalog. Often archival repositories either run their own microform laboratory or have easy access to one. (Of course, librarians may also produce microforms and many reasons cited here for archival microfilming also apply to libraries, but I'm making distinctions to contrast the approaches). As such, archivists use microfilming (and other forms of "copying") in a variety of ways that are not mutually exclusive. It is not unusual for archivists to film the same materials for several reasons.

A. Filming for Users

1. Reference or scholarly copying

To provide offsite researchers with unique (and therefore non-circulating) materials. For a fee, it is not unusual for a repository to microform part of all of collection for a researcher who cannot travel to the institution.

2. Copying for publishing

This is the category that most directly relates to the library dicussion of micropublications . Commercial publishers are interested in archival materials to make them available for wider dissemination at a profit. Sometimes micropublications are partially funded by a granting agency (most often the NHPRC) and distributed by a micropublisher. Commercial microfilm publishers film collection editions (a single collection from a single repository such as the Draper Manuscripts from the State Historical Society of Wisconsin) or collected editions (materials about a topic, event, or person that are gathered, selected, and filmed from a number of repositories.)

3. Publication of holdings

Closely tied to commercial micropublications (and often overlapping) is in-house (or out-of-house) microfilming to disseminate a repository's own holdings. This category is also related to reference or scholarly copying only the purpose is for wider distribution and the products are more polished. These "publications" may be no more than duplicates of a master negative retained from a reference request. But established programs such as the ones at the Library of Congress and the National Archives, issue these kinds of products are of high quality. This may

confuse matters further, as some catalogers then treat LC and NARA as commercial publishers.

B. Acquisition copying

To acquire archival materials which are owned or housed at other repositories, institutions, or are in private hands.

C. Preservation

To preserve the intellectual content of deteriorating original documents and to protect original documents from the wear and tear of use.

D. Bulk Reduction

To save space. The bulk of many modern archival collections is so tremendous that archivists choose to film the records and destroy the originals.

Other archival concerns:

Because they see microfilming as a process suited to any or all of the above ends, archivists often film, or acquire on film, only parts of collections. Scrapbooks, clipping files, case files, and other forms of materials are routinely filmed as part of processing to preserve information that exists on fragile media, reduce the bulk of the collection, or acquire part of the collection held by another institution. These microforms are very different indeed than materials produced by micropublishers and intended for wide distribution.

Finally, in addition to making microform copies in the course of managing their collections, archivists administer archival collections in which microforms are the "original" documents. Many government agencies and large institutions with voluminous documentation responsibilities create COM (computer output microform) as the original record. (When I first heard of the library distinction between reproduction and original microforms, I immediately thought of COM as "original" microforms).

Reflecting their orientation, archivists are concerned the following data elements when they describe microforms.

Custody (Location of materials)
1. Own repository
2. Another repository
3. Creators of material

Disposition (where are originals)
1. Kept
2. Destroyed
3. Returned

Who made the reproduction
1. Inhouse
2. Out-of-house

Availability source
1. Own repository
2. Other

Ownership of master negative
1. Own repository
2. Other

IV. DESCRIBING MICROFORMS USING MARC - THE CURRENT SITUATION

Having discussed the background to the library approach to cataloging microforms, and the archival use of microforms, let us examine how librarians and archivists currently catalog microforms of archival and manuscript materials using the MARC formats.

A. Librarians:

1. Which format to use?

The choice of "which format to use?" should be straightforward. The MARC Formats for Bibliographic Description offer clear guidelines in Leader/06 (Type of Record) which states that:

"Microforms, whether original or reproductions, are not identified by a distinctive type-of-record code, i.e. the type of material characterisitics described by the codes take precedence over the microform characteristics of the item."

Therefore, according to MFBD, all microforms of manuscripts materials should be described in the AMC format. However, this is not the practice of library catalogers. Both OCLC and RLG users are cataloging some microforms of archival and manuscripts materials in BOOKS, SERIALS, and AMC formats.

a. Library catalogers use BOOKS and/or SERIALS format for "commercially" generated microforms (see discussion concerning the question of what is a "commercial publisher"). Proponents say these belong in either BOOKS or SERIALS because they are "published" (i.e. not unique) and those formats contain the necessary "publication" fields (e.g. 265, and the series blocks-- 4xx and 8xx). On the other hand, these formats lack some note fields available in AMC (555, Cumulative Indes/Finding aids note; 535, Location of original/duplicate note; 351, Organization and arrangement note; and 524, Preferred citation of described materials) are not available in BOOKS or SERIALS. And

530-Additional physical forms available is not present in BOOKS. Catalogers using BOOKS or SERIALS either leave this information out of the record or put it in the 500, general note, field.

b. Library catalogers who use the AMC format, following the MFBD's, encounter problems with this format as well. Because commercially generated microforms often exhibit "publishing" kinds of information, fields not valid in AMC are needed. For example, commercial microform publishers often produce series of microforms and therefore catalogers need the 4xx and 8xx series fields. (Although not valid in MFBD, OCLC makes 4xx and 8xx series field available in AMC).

2. Reproduction vs. original aspect

Regardless of which format the cataloger chooses, the question of whether the microform is a reproduction or an original must be faced. OCLC, based on LC policy, gives their users guidance to distinguish between the two. RLG has not issued any guidelines, although their Archives and Manuscripts Task Force is examining the issues. RLG has, however, been more involved in helping archivists describe reproductions in AMC (see below).

"OCLC advises users to catalog microforms of archival collections, or parts of archival collections, as reproductions using the AMC format when the collections exist prior to filming and as originals when the collections are brought together to generate original editions in microform. LC uses similar guidelines, but tends to catalog more of these items as originals. For example, in cases where part of a collection is filmed, OCLC has advised libraries to catalog the item as a reproduction while LC would probably advise libraries to catalog the item as an original."

OCLC tells its users that, within AMC, if it is a reproduction, describe the original in 245-300 with a note in 530, 533, or 535. If it is an original, describe the microform in 245-300 with appropriate notes in 53x. These guidelines are based on LC guidelines. Parenthetically, LC does not presently use the AMC format and describes microforms of archival and manuscript materials in BOOKS so their rules about use of 53x notes differ.

B. Archivists

1. Which format to use?

With the advent of AMC, archivists were faced with cataloging archival materials in accordance with library standards to create descriptive records that could be integrated into library catalogs. Their concerns in cataloging microforms were, however, different from those of the library community. For example, archivists are more often faced with cataloging microforms created for a variety of reasons (see discussion on page 4-5) . Cataloging micropublications of archival materials is just one aspect.

The four archival repositories taking part in the initial RLIN AMC implementation (Cornell, Hoover, Stanford, and Yale) developed a series of guidelines for cataloging reproductions in AMC. These suggested guidelines have become the de facto standards and are what SAA teaches in the MARC AMC format workshops. The RLG implementation group decided to catalog micropublications of archival materials in AMC. However, they were much more concerned with cataloging the other kinds of microforms and concentrated on developing guidelines for these cases.

2. Reproduction vs. original aspect

Instead of concentrating on reproduction versus original aspects, the RLG archivists' major concern was to develop standard ways to use 530/533/535 within the library context and at the same time to answer archival needs. The kinds of information archivists want to include in the records answers the question "does my repository own or have custody over the originals that were filmed? if not, who does?"

3. Use of 53x note fields

Following the RLG guideline, archivists using RLIN AMC catalog microforms in the following way:

530- Additional physical form available

If all or part of the archival material is available in a different physical format (microform, photocopies, published book?) and your repository holds the originals, describe the originals in 245-300 and use 530, to note the additional physical form.

533- Reproduction note

If you own only the reproduction and the originals were 1) owned by you but destroyed or 2) owned by a different entity and either still extant or destroyed, then describe the material in hand in 245-300 and note the reproduction informa-

tion in 533. Although dictated by the description of field 533, archivists do not always describe the originals in 245-300. This field is used in conjunction with 535.

535 - Location of original/duplicates

Use only in conjunction with 533 when describing location of originals (because you use 530 when YOU have the originals). Note the location of the originals or note if the originals were destroyed in 535. If a repository tracks the location of duplicates, note those locations in 535 as well.

C. Issues and Problems

1. The distinctions between "original" and "reproduction" is difficult to apply to microform materials of manuscripts collections.

2. Confusion stems from trying to use fields originally developed by library catalogers for other formats (533 and 530) in archival ways. For example, archivists distinguish between the use of 530 (additional forms) and 533 (reproductions) both by who owns (or has custody) of the materials and whether the materials still exist. The RLG guidelines are not logically consistent because library practice isn't; therefore, they are confusing to apply.

Specific problems occur when you try to describe the following:

- a collection which a repository owns, has filmed and has destroyed a part of;
- copies of your material, available for purchase from another institution
- the distinction between preservation negatives and positive copies.

3. There is inconsistency in what archivists are describing in 300. Some repositories are describing originals and microform and some are describing just originals.

V. POSSIBLE SOLUTIONS

The two principal issues discussed in this paper are which format to use to describe "commercially produced microforms" and how to describe various kinds of reproductions within the AMC format. Two possible solutions are discussed below.

A. Catalog all "commercially produced" microforms of archival materials in BOOKS or SERIALS

Problems:

1. LC would need to validate some of the AMC fields for BOOKS/SERIALS.

2. How to define "commercially produced?"

3. This solution violates current MFBD distinctions.

4. It would confuse searching the data bases if researchers/reference librarians search under rather than books

5. How would this solution effect an archivist who has responsibility for cataloging this kind of "commercial" material but has software that only creates AMC records (e.g. MicroMARC:amc).

B. Put all microforms of archival materials in AMC.

Problems:

1. LC would need to validate fields for publication information including the series fields.

2. This goes against fundamental archival principles.

Possible solutions to the problems of describing reproductions in AMC are not as straightforward. To redefine fields 530 and 533 so that they are more logicallly consistent, is not possible because 530 is valid in the VISUAL MATERIALS, SERIALS, and AMC formats and 533 is valid in the BOOKS, VISUAL MATERIALS, MAPS, MUSIC, SERIALS and AMC formats.

An option suggested by Max Evans is to create one large field (535 since it is only valid for AMC?) and use it to hold all the necessary subfields archivists require.

VI. CONCLUSION

It is easy to get bogged down in the complexities of this topic. What is most important to keep in mind, (and what we tell the SAA MARC-AMC workshop participants) is to create a catalog record that expresses to the user what you want it to. However, this is often easier said than done.

SAA NEWS

The NEH awarded the SAA most of the funds it requested to continue offering the MARC AMC workshops, a workshop on descriptive standards for the next two years, revise Steven Henson's archival cataloging manual and publish a book of descriptions in AMC format, as well as to support the SAA Automation Information Center.

By now, most American archival repositories have (hopefully) returned questionnaires concerning automation inevtments to Lisa Weber at the SAA office. Data is being entered into an SAA database where it will become part of a clearinghouse on archival automation.

PLANS & PROPOSALS:

ELECTRONIC RECORDS AND THE NEW NATIONAL ARCHIVES OF CANADA By John McDonald

After five years of consultation and review, the National Archives of Canada Act was proclaimed on June 11, 1987. This significant legislation, the first archives legislation since the Archives Act of 1912, changes the name of the institution from the Public Archives of Canada to the National Archives of Canada and enables the Archives to carry out the following responsibilities:

* collecting and caring for records of national importance and making them accessible to researchers and the public from all parts of Canada
* serving as the permanent repository of records of government institutions and for ministerial records
* providing professional technical and financial support for archival activities and the archival community

It also establishes the National Archives' role as advisor on the management of government records. Furthermore, it stipulates that no record under the control of a government institution and no ministerial record is to be destroyed or disposed of without the consent of the National Archives.

The Act also requires institutions to transfer those records having historic or archival importance to the care of the National Archives under certain conditions and few exemptions. These transfers are to be accomplished in accordance with schedules or other agreements.

The broad scope of the National Archives mandate is based, in part, on the definition of the term 'record'. According to the Act, 'a record includes any correspondence, memorandum, book, plan, map, drawing, diagram, pictorial or graphic work, photograph, film, microform, sound recording, videotape, machine readable record, and any other documentary material, regardless of physical form or characteristics, and any copy thereof.'

As a result the National Archives has a legislated mandate to appraise the archival value of any form of recorded information (including machine readable or electronic) and to arrange for its acquisition, preservation, and dissemination. Based on its additional mandate to facilitate the management of records it is also in a position to participate in activities that relate to the overall care and handling of the valuable data resources generated in federal government institutions.

Recognizing that this broad mandate confirmed a potentially large role for the National Archives in federal government data management activities, a study was commissioned to determine how the National Archives should position itself with respect to its potential relationships with the informatics communities of the federal government institutions as well as with those other areas that are responsible for the management of machine readable records. The results of this study, which has involved extensive consultation with central agency and departmental officials, will be available in August 1987.

During the coming months, steps will be taken to test a model approach to the implementation of data scheduling and data conservation functions in selected federal departments and agencies. The proposed model is associated with systems and surveys (i.e. structured data that is created, used, retained and disposed of on a systematic basis - essentially a data management environment supported by the tools and practices associated with the field of data management).

Past experience has demonstrated that the greatest challenge to the establishment of these functions is securing senior level support and building the functions into the mainstream of the systems development and survey design process. In this respect, it has been found that an understanding of the institution's system development life cycle (including the use and potential applicability of data dictionaries, systems audit checklists and their openness to incorporation of data retention and conservation issues, existing data conservation practices, if any) and, above all, the political and organizational characteristics of the informatics and/or research and statistics areas, is essential before any procedures for the actual establishment of data retention schedules can be developed.

In addition, and whenever possible, the issues that give rise to the need for data

scheduling should be incorporated into issues that have already been identified within the institution. These can range from the space problems associated with tape storage, to the impact of legal and other accountability requirements, to the need by senior management to establish comprehensive views of their institution's information holdings (possibly through the linkage of corporate finding aids such as automated records systems and data dictionaries).

Given that the data holdings of an institution are often scattered across a variety of diverse program activities (operational and administrative), it has been found useful to build these data scheduling and data conservation functions within those areas that will offer the greatest return for the energy invested (i.e. through the selection of highly valuable corporate data holdings which best reflect issues that can be identified and owned at the senior corporate level of the institution). Finally, if data scheduling and data conservation considerations are to figure prominently in the design steps leading to the installation or major modification of systems/surveys, then archivists and records managers will be required to view themselves as corporate users - users who have a right to express functional requirements that ought to be respected by systems designers to the same extent as the requirements expressed by the primary users of the systems or surveys.

With respect to the nanagement of electronic documents, a set of draft (very preliminary) functional specifications for the management of information in integrated office support systems has been developed. In contrast to the data scheduling efforts described above, these specifications focus on the management of documents (unstructured text created, transmitted, used, retained and disposed of in a non-systematic manner - essentially a 'document management' environment supported by the tools and procedures associated with the field of records management).

The rather simplistic division of the information universe into data management and document management worlds raises a number of questions concerning the manner in which archivists should view the information universe. For the order one gives to this universe, will determine the policies, and procedures, that govern the management of information in a given organization. Suggestions concerning the establishment of (a) model(s) of the information universe from the perspective of the issues raised in this article would be useful.

Among other activities related to the National Archives' involvement with federal government data holdings, steps have been taken to develop guidance concerning the application of General Records Disposal Schedules to data in automated information systems. Similarly, a project is underway to produce a retention and disposal authority for so-called transitory records or those inconsequential records that are normally of temporary value and are considered to be neither corporate nor part of the official information system of the organization. This latter project has been particularly challenging and any ideas or suggestions regarding the criteria that could be used to define this body of records would be welcome.

Finally, a recently completed study has presented options for potential National Archives involvement in national and international data and document interchange standards activities. This study describes the objectives, structure, and responsibilities of national and international standards organizations and assesses the applicability of certain standards to the concern of the National Archives for the ongoing care of digitally recorded information, particularly as it is managed in federal government institutions.

Copies of this report, as well as other information associated with the issues raised in this articles, may be obtained from: John McDonald, Director, Automated Information Systems Division, Government Records Branch, National Archives of Canada, 395 Wellington Street, Ottowa, Ontario, K1A-0N3 or call (613) 996-0969.

CANADIAN DESCRIPTIVE STANDARDS

It is my privilege as editor to introduce a new Canadian colleague -- the Planning Committee on Descriptive Standards.

Following the publication of Toward Descriptive Standards, in 1986, the Bureau of Canadian Archivists established a committee to ensure that the recommendations of that report were acted upon. It consists of

two representatives each of the ACA and the Association des archivistes du Quebec plus the Secretary General of the Bureau of Canadian Archivists (Jacques Grimard) and a PAC observer. When it met for the first time early in 1987, it defined its objectives as:

Establishing descriptive standards and rules for the intellectual control of archival materials a) at the fonds d'archives level (to accomodate all media), and b) by medium at the series, file unit and item level.

Working with the Canadian Committee on MARC and the National Library of Canada to adapt the format. [CanMARC is only slightly different from USMARC, but it is politically quite another beast.]

Studying a) existing names authorities and rules and b) the problem of subject indexing of archives and archival finding aids, and adapting them to archival principles and practices.

This breathtaking agenda is supported by the Canadian Council of Archives which provided a grant to enable the committee to hire a project officer and to hold its meetings. Two short term projects were undertaken. The first is to advise the Canadian Committee on MARC about the view of Canadian archivists towards MARC format development. The second is to produce a guide for archivists on construction and use of name authorities (drafts will be ready by the end of summer).

The main project for the year has been assigned to a subcommittee of media specialists who will establish the rules for control of archives at the fonds d'archives level. They are expected to report by the end of March 1988. The second group has also been appointed, and while it will begin its discussions this year, it cannot proceed far on rules for description at the series, file unit and item level until the first group reports.

The Planning Committee on Descriptive Standards will keep interested parties informed of its progress through an occasional newsletter. For subscriptions, write to:
Diane Beattie, Project Officer, Planning Committee on Descriptive Standards, c/o Public Archives of Canada, 344 Wellington St., Room 4101, Ottowa, Ontario K1A-0N3, CANADA or call (613)-995-2372

CONFERENCES

July 12-14
International Conference on Databases in the Humanities & Social Sciences, AUM, Montgomery, Alabama

July 22 - 25
National Association of Government Archives and Records Administrators, Colony Square Hotel, Atlanta, Georgia

August 2-5
Recognition Technologies Users Association, Hyatt Regency, San Francisco

September 1 -5
Society of American Archivists, Grand Hyatt, New York. SAA, 600 S. Federal St., Chicago, IL 60605 (312-922-0140)

October 4-8
ASIS 50th Anniversary Conference, Sheraton Boston Hotel, Boston. ASIS, 1424 16th St., NW, Washington, DC 20036 (202-462-1000)

October 12-13
Museum Computer Network, Royal Sonesta Hotel, Cambridge, Mass. $90, incl. 1 yr. membership, to P.O.Box 111, East Winthrop, ME 04343

October 19-22
ARMA Annual Conference, Anaheim, CA Int. Rec. Mgmt. Council, 22243 Miston Dr., Woodland Hills, CA 91364

October 21-23
Local Television News Archives Conference, Madison WI.
Sponsored by the National Center for Film and Video Presentation, American Film Institute, with funding from the NHPRC. For archives, libraries and museums that preserve local television news and public affairs broadcasts. $25, registration deadline August 1. AFI, 2021 North Western Ave., Los Angeles, CA 90027 (213-856-7637)

November 11-13
Optical Publishing and Storage, Penta Hotel, New York
Learned Information, 143 Old Marlton Pike, Medford, NJ 08055

IN-BOX

REPORTS:

American Association of Colleagiate Registrars and Admissions Officers (One Dupont Circle NW, Washington, DC 20036), Retention of Records: A Guide for Retention and Disposal of Student Records, 45pp., 1987 ($8), is an exceptionally valuable report of a very productive AACRAO Task Force. Not only does it treat the legal issues and requirements for a records retention program, it provides excellent advice on micrographics, computer readable records and electronic imaging systems. A fascinating appendix reports, state by state, on the policies governing disposition of academic records of closed schools.

Coopers & Lybrand, Information & Image Management: The Industry & The Technologies, Coopers & Lybrand, NY, 1987, 67pp. & glossary. This study, commissioned by the Association for Information & Image Management, is a forecast of the market and trends in this dynamic field. Significantly, the study concludes that micrographics will not be made obsolete by other means of dense data storage in the near term. Not surprisingly it looks towards greater integration of imaging with other office capabilities. The report is loaded with pretty graphs but the sources of the data are usually AIIM itself, so one is forced to wonder whether the purpose of the study is to legitimate Information and Image Management, and if so, for whom (given that AIIM wants $495 a copy!) Do you believe that in 1986 the U.S. produced 1.175 Billion pages of paper (or film equivalents) and that 34% of this was on COM?! or that scanning and OCR costs will fall 20% a year for the next decade?!

Library of Congress, Prints & Photographs Division, Descriptive Terms for Graphic Materials: Genre and Physical Characteristics Headings, compiled and edited by Helena Zinkham and Elisabeth Betz Parker (available from the Cataloging Distribution Service) Washington DC, 1986, is the long awaited list (familiarly known as gmgpc to your subfield 2 of MARC 655 and 755 which have been hankering after it). It is everything we waited for and more. Not only does it contain most of the terms we could want, it is and excellent thesaurus with helpful scope notes for the public and catalogers. The introduction contains good clear advice and reasonable suggestions for level of specificity and exhaustivity in indexing. Perhaps the most important information in this book is the address to send updates and recommendations. Because the Library will maintain the list, we will all benefit; we need to put aside other quibbles and adopt the language. We can work to make it work.

Miller, David C., Special Report: Publishers, Libraries & CD-ROM: Implications of Digital Optical Printing. A report for the Fred Meyer Charitable Trust, March 1987, 99pp.
Millers' style is very chatty and the structure of the report is somewhat telegraphic. If you already know the basics, Millers opinions are intriguing, if debatable.

New York State. Governor's Office of Management and Productivity, State Archives & State Education Department; Computer and Audio-visual Records in State Government: Preliminary Report of the Special Media Records Project, April 1986, 69pp.
This is a report on a cooperative study of 19 NY State Agencies conducted by the authoring organizations in 1985/6. Their findings, that the use of computer and audio-visual information was increasing dramatically and that the state government was ill equipped to handle them are not surprising but their examples of the importance of such records (their non-routine quality) and the range of actions proposed, are fresh and worth further study beyond the state boundary.

Vogt, Diane; Smithsonian Archives Photo Survey Project: A Draft Photographic Thesaurus, 3/87, 117pp. is the bi-product of three years effort in cataloging over 6,400,000 photographs in 1,500 Smithsonian collections for The Finders Guide to Photographic Collections at the Smithsonian Institution a planned five volume work to be published beginning in 1988. It is likely to be more important to photographic archives and repositories with photographic collections, than even the impressive volumes it was constructed to index. The draft, billed as incomplete and promised soon in an edited and widely available version, is, in its present state, the most complete, authoritative and usefully organized reference work for photographic cataloging which exists. Its importance lies in its definitions of photographic proceses, formats and techniques, where the discussions are extended and provide clear criteria by which to distinguish different types. The draft contains a few modest errors and occassionally differs in use of preferred terms for document types from the LC Descriptive Terms for Graphics Materials and the AAT. Hopefully, these differences can be worked out in final draft making this a truly definitive work.

Walch, Victoria Irons, Information Resources for Archivists and Records Administrators: A Report and Recommendations, Albany, National Association of Government Archives and Records Administrators, 1987, 42pp. [Available, free, from NAGARA, NY State Archives, Room 10A75, Cultural Education Center, Albany, NY 12230] For the past two years, NAGARA has been studying how best to meet the information needs of archivists and records administrators; this report presents three options ranging from status quo to full-service information center. It advocates the middle course of establishing a modest clearinghouse function within the NARA library. Although the report is well written and its conclusions well supported, its very conservative recommendations are somewhat anticlimatic.

NEWSLETTERS:

Access Reports: Freedom of Information Newsletter (ISSN 0364-7625) is published bi-weekly by the Washington Monitor Inc., 1301 Pennsylvania Ave. NW, Suite 1000, Washington DC 20004 $250 p.yr. has been carrying a series of reports on computerized government databases and privacy issues and closely following legislative hearings on "computer matching".

Charles Babbage Institute Newsletter (Center for the History of Computing, 103 Walter Library, 117 Pleasant St. SE, Minneapolis, MN 55455) reports on a wide variety of activities having to do with the history of information processing, most recently on the formation of a "National Archives for the History of Computing" in the U.K. (by John Pinkerton at Manchester University, Dept. of History).

Library Conservation News (ISSN 0265-041X) is a free quarterly publication of the Preservation Service of the British Library, Great Russell St., London WC1B 3DG. I found October 1986, which reported on the digitization of sound recordings at the British National Sound Archive, particularly intriguing, but it is always full of useful information and is very international in scope.

Library High Tech News (published 11 times p.a. by Pierian Press, P.O.Box 1808, Ann Arbor, MI 48106, $65 p.a.) continues to be the best source of basic bibliographic references for archives and museum automation, even though it includes no archives or museum periodicals per se. The bibliography is printed in a machine readable form (Softstrips) in each issue as well.

Traveling Exhibition Information Service Newsletter, published bi-monthly by The Humanities Exchange Inc. (P.O.Box 1608, Largo FL, 34294) is more a classified listing service than a journal, but its lists of 50 or more available exhibitions from its members is very useful.

ARTICLES & BOOKS

Blake, Monica; "Aspects of Electronic Archives", Electronic Publishing Review, vol.6 #3, 1986, p.151-158, is a report of a study by the British National Bibliography Research Fund on establishing a national archive of electronic publications similar to the National Sound Archive and the National Film Archive in concept and form. The paper reports on the extent of electronic publication in the UK by format (videodisc, CD-ROM, on-line databases, videotext) and discusses some archival implications of dynamic change in data/imagebases which are only available electronically. The problem of archiving Prestel is discussed. The paper does not propose actions.

Cloud, Patricia, "RLIN, AMC and Retrospective Conversion: A Case Study", Mid-Western Archivist, v.11(2) 1986, p.125-134, is the first real discussion of the costs of doing a retrospective conversion in RLIN for MARC AMC. It is exceptionally useful for that reason as a planning framework for others. The results, simply, are that it took Northwestern 2.7 hrs per record with one full time project coordinator, one part time archivist and a part time student assistant. Considerable time was devoted to authority checks and reviewing the records themselves; in addition it took 41 minutes to code and 22 minutes to enter a record.

Cook, Michael, "An Introduction to Archival Automation: A RAMP Study with Guidelines", UNESCO General Information Program, September 1986 49pp.
As the title implies, a primer, complete with a basic bibliography.

Johnasson, Stig; "Machine Readable Texts in English Language Research", Humanistiske Data, #3-86, p.27-34, reports on the status of the International Computer Archive of Modern English (ICAME) and the authors' use of the Lancaster-Oslo/Bergen Corpus which, together with the Brown University Corpus, the London-Lund Corpus and several minor corpa is distributed by ICAME.

Logan, Robert K.; The Alphabet Effect: The Impact of the Phonetic Alphabet on the Development of Western Civilization, New York, William Morrow & Co., 1986. It probably takes having co-authored with Marshall McLuhan to write that: "It is the transition from the impressed tablets with a vocabulary of two hundred token signs to the incised tablets with a proliferation of pictographic signs, created with the use of a stylus, that marks the advent of writing, mankind's greatest breakthrough in data processing." Other observations are equally fascinating and equally facile, adding up and infuriating in the manner of Michel Foucault.

Marx,Peter; "State Public Records: Database Goldmine or Landmine", Information Times, April 1987 p.21,28, explores the prospects for commercial distribution of state databases including legislative and regulatory data, concluding that efforts to date have been hampered by position taken by the states with respect to commercial use of public data. Urges the development of "uniform and reasonable" standards for access to state data as an impetus for investment by private firms in its dissemination.

Nath, Sandra; Research Study for a National Documentation Organization in the United States, May 1987, MS. Thesis, Museum Studies, John F. Kennedy University. 172pp. It is a coincidence that the 50th anniversary year of the American Documentation Association (now ASIS) is the tenth anniversary of the Museum Documentation Association, the British group which inspired this thesis, but the fact (which went unnoticed by the author) might have contributed to this study which examines the MDA and the Canadian Heritage Information Network and asks, in effect, how we can make it happen here. The author concludes that what is required in the US is an organization which maintains a database, data standards, consultative services and training programs, as CHIN and the MDA do in their respective (but quite different) ways. One might also conclude that the failure of that model in the U.S. in the past bodes ill for it in the future.

Perry, Meg Woollen, "An Inside look at a LAN Data Archive System", Byte, July 1987, p.169-176, reports on a home grown LAN archiving (not backup, but true archiving) system in use at Brecton Dickinson Research Center in Research Triangle, NC. Useful flow charts illustrate user and system decisions in a real time environment with an archiving option. A working model of how data archiving can work in an OA system.

Welsh, Peter H. and Steven A LeBlanc, "Computer Literacy and Collections Management", Museum News, vol.65 #5, June 1987 p.43-51, is a solid, general introduction, to the kinds of questions which should be asked in a computerization project. The authors do not rigorously address the requirements of a collections management system per se nor, I'm afraid, do they fully declare their interest in Questor Systems and its software ARGUS which is their exemplar throughout. They insist upon the common (& dubious) view of museums that it is essential "*to enter all records completely and as they stand*" (their emphasis) rather than build the collections management database through use, because they equate collections management functionality with information retrieval and are really discussing cataloging/registration systems rather than collections management. They seem overly impressed by the importance of operating systems for end-users (and with their selection of a PICK operating system based application), but their observation on the importance of system tools certainly holds true for developers and if multi-user, multi-tasking functionality is not considered a concern which would be voiced at the application level, their discussion of it is well taken.

EPHEMERA

Kerr & Downs Research; ARMA International Membership Survey, 1987, 22pp. is based on a 41% response of members polled. ARMA is perceived to be meeting the needs of members. What might interest archivists and museum staff is that well over 80% of ARMA dues are paid entirely by employers with exceptions being almost all retirees, consultants (who are also their own employers) and government workers. When a trade certifies practitioners, and employers hire the certified, the employers in effect gaurantee the continued membership of the association.

Derwent Guide to Patents , free from: Derwent, Inc. (USA), 6845 Elm St., Suite 500, McLean, VA 22101, is a useful introduction to the patent process and to patent records and their contents.

D&O, Directors'& Officers' Liabilty: A Crisis in the Making. A study of Museum Directors, Peat Marwick, 1987, discusses the changing responsibilities of directors of non-profits and reports on a museum survey which documents the extent to which they have not adjusted to the new context of legal liability. Applicable to all kinds of cultural institutions this study complements

other Peat Marwick studies of the liability of University Presidents & Boards, Government officials, and others. Available free from Peat Marwick, try your local office.

We Are Losing Our Past is a popular pamphlet calling attention to the Preservation Needs of State Archives published by the National Association of Government Archives & Records Administrators and available free from the Council of State Governments, Iron Works Pike, P.O.Box 11910, Lexington, KY 40578

As part of an NHPRC grant which enabled the transfer of the archives of the National Federation of Abstracting and Information Services, to Temple University Archives, NFAIS published A Model for Donor Organization and Institutional Repository Relationships in the Transfer of Organizational Archives, by Miriam Crawford, NFAIS, Philadelphia, 1987, 25pp. [NFAIS, 112 S. 16th St., Philadelphia, PA 19102].

UNESCO's ICOM Documentation Center (Maison de L'UNESCO, 1 Rue Miollis, 75732 Paris) will be glad to send, upon request, the fourth edition of its Basic Museum Bibliography (1986).

The NARA Archival Research and Evaluation Staff has circulated Prospectus for Access by Function and Process for the National Archives and Records Administration, 23pp plus numerous appendexes, within the National Archives. The document discusses the potential use of a functions/processes vocabulary to support user access to materials across provenance in line with the proposals being discussed in the RLG Seven States Project and often advanced by this editor.

Financial Executives Research Foundation has published "EDGAR: The SEC's Pilot Program and Its Impact", an introduction to the electronic filing project and a discussion of its potential. ($8, from FERF, 10 Madison Ave., P.O.Box 1938, Morristown, NJ 07960)

Technical Data Publishing Company (91 North Bertrand Rd., P.O.Box 458, Mt. Arlington NJ 07856) is distributing a free electronic publishing glossary including the terminology used by printers, typesetters, writers, editors, and terms from data communications and computing.,

SOFTWARE

Mason Barnett (Assistant University Archivist, Duke) reports that after considering MARCON, Micro-MARC AMC, and In-Magic, the Duke University Archives selected Revelations by Cosmos and is in the process of writing the routines to create MARC AMC leader/header and record AMC descriptive data within that package.

Caesar Iacovone (Director, Div. of Archives & Records Management, NJ) reports that the NJ micrographics accounting system is proprietary, but for the benefit of those considering similar applications, he describes it as consisting of:

"System Code Tables - for identifying employees, labor hours, supplies, fixed assets, and hard copy/microfilm location; codes are used to drive data entry and reporting on all system modules.

Job Costing - individual and aggregate cost totals for seven-phase production process

Job Specification - complete service specifications for microfilm jobs

Productivity Tracking - based upon entry of daily labor hours and supply usage; used for employee productivity evaluation, budgeting, supply inventory depletion, and billing

Work Flow Tracking - reporting modules for monitoring work-in progress and turnaround performance

Bill Out - itemized billing based upon input of productivity data

Accounts Receivable - driven by billing; includes client budget status and open amounts (aged)

Accounts payable - records payments due to vendors; tracked by vendor, account, and minor object of expenditure

Inventory - status tracking of supplies; includes reorder report and current inventory level /value report

Fixed assets - records information about equipment and fixtures; includes depreciation tracking and book value reporting

General ledger - general purpose application for recording and reporting on all fiscal & financial transactions; includes various financial reports - e.g. balance and income statements"

Peter Sigmund (Director, Rijks Archiefschool, the Netherlands) sent me a description of MAIS, a new Oracle based, PC system developed by the Dutch Ministry of Interior.

"MAIS stands for 'Micro Archief Inventarisatie Systeem'. MAIS is a fully menu-driven program designed to support the description of archives and the production of inventories and guides.

Functions of the System:

Data Entry: The screen lay-out contains fields that cover all relevant elements for arranging and describing archives, such as 'form of material', contents, dates, etc. Value tables speed up the input and control the uniformity of the descriptions. In the contents text field, words can be marked for indexes of persons or subjects. Extra descriptors can be added. There are special functions for linking series descriptions or splitting series descriptions into sub-records. The system keeps track of correct numbering.

Arrangement: Once entered, descriptions may be automatically arranged for a guide according to classification codes assigned by the archivist. The descriptions are sorted by classification, according to criteria chosen by the archivist. The system sorts descriptions in chronological order within the arrangement selected, although one can choose other fields for a secondary sort as well. The system provides several methods for assigning classification codes and designing a classification scheme (such as record groups, classes, series, subseries) . For instance, it is possible to store classification codes with their definitions in separate tables in advance or during the description process. A special function enables the archivist to view the distribution of descriptions by class, or to get a synposis of the documents that are not yet classified.

Reports: MAIS produces six standard reports:
1. A preliminary inventory (including both the preliminary and final numbering)
2. A finished inventorty (in which only the final numbering is printed)
3. Indexes by subjects with keywords from the contents field, as well as attached keywords
4. Indexes on names (with the same options as 3)
5. Concordances
6. Lists of annotations made up by the archivists while arranging and describing can be printed with reference to the number of the description.

Technical specifications:

MAIS runs on microcomputers with MS DOS version 2.1 or later on the relational database management system Oracle which requires at least 512KB memory and a 10MB hard disk. A facility is provided to edit MAIS reports in Wordstar or Wordperfect.

The Museum Documentation Association (Building O, 347 Cherry Hinton Rd., Cambridge CB1-4DH, ENGLAND) has released its Museum Object Data Entry System (MODES) which supports entry of the standard MDA data records on an IBM PC or a system running CP/M 80 version 3.0. The system contains data entry features (such as carry forward of field values and validation by MDA vocabulary and syntax) and prints 3x5 cards. MDA claims that the input and output formats can also be tailored. It sells the system, without updates or support, for 115 pounds sterling. One year of software and documentation updates cost an additional 33 pounds .

Copies of the ADAPSO/EDUCOM brochure "Using Software: A Guide to the Ethical and Legal Use of Software for Members of the Academic Community" which includes the text of the statement on software and intellectual rights, are available free, from ADAPSO, 1300 North 17th St., Suite 300, Arlington, VA 22209.

Paradigm Press announces that DiscSCOPE, a version of SCOPE: Humanities Computing Update, which includes selected items from Software, Courseware and Software Reviews sections of recent issues, "has been placed on a floppy disc with a sophisticated menu program to provide access. The disc may be freely copied for circulation amoung colleagues and uploaded to bulletin boards. Updates are scheduled every four months. Copies of discSCOPE can be purchased for $5. Organizers of conferences and user service personnel at computer centers who wish to distribute copies (with their own labels if the desire) may request free masters on their letterheads." Prepaid orders or requests for free masters should be addressed to DiscSCOPE, P.O.Box 1057, Osprey, FL 33559.

IN CONGRESS

Proposals to ban Digital Audio Tape recorders without copy protection systems are still being pressed by the music industry and optical publishers. If they succeed, digital audio tape (Whats DAT?) will not do to CD's what videotape did to home videodiscs and what audiotape threatened to do to phonographic records before CD's got there, and archives & museums will lose a potentially valuable technology. (see Nancy Herther, "Much ado about DAT", Database v.10 *3, June 1987 p.,116-120)

At this writing, Congress had overwhelmingly passed legislation to remove authority for control of non-classified but "sensitive" information from the military effectively rescinding President Reagan's notorious executive order of November, 1986, although it remains to be signed.

Capturing Rich Context
by David Bearman

For the past several months I've been preoccupied with two problems which arise from the goal of attempting to document culture accurately and richly. The first issue is how to identify and acquire evidence of processes, not just of products. The second is how to represent processes and relationships, not just entities. As usually happens when we get preoccupied, I would like you to take these concerns as seriously as I do.

We have long relied on tacit knowledge of how things are done (by humans, in particular types of social situations which we can envisage ourselves in) as a surrogate for documenting how things are in fact done. This luxury is receding in a number of areas as a result of automation. For example, since the advent of bureaucracies and postal systems, archivists have assumed with reason that if a document is in my filing system, I received it shortly after the date it was written and referred to it in connection with the topic under which it is filed. Based on such inferences from the evidentiary remains of the office of a policy maker, we can make reasonable guesses about how a policy decision was formed.

In fact, the validity of the inference began to break down with the widespread use of the photocopier, because many records in my files arrived at times long past the dates on them and may never have been referred to by anyone. It is clear that in electronic office environments our assumptions completely break down, either because the document isn't filed with me at all or because my query for information recovers numerous documents which I subsequently do not examine. The computer attached printer, especially laser printer, and CRT viewed documents, have merely exacerbated the problem by eliminating any clues we might have had about what constitutes the "original" (a problem which Lisa Weber addresses at length elsewhere in this issue). The result is that if we want to know what went into a decision, we may need to have contemporary, explicit, process oriented documentation. This kind of analysis is also the foundation for the efforts described by John McDonald.

This problem is more serious with electronic databases. In its last issue, the Archival Informatics Newsletter published a technical leaflet on how to transfer records from a DBMS to a software independent format for retention. This solves some physical problems, but it leaves the fundamental question of what we are documenting in the air. If I want to document what the file actually looked like (what data it contained) when I made an important decision based on it, I need not just a snapshot of its contents at some later date, and not even that snapshot plus a complete audit trail (with which to back in and out all the changes), but also an actual record of what I looked for and how the software processed it. Processing rules are increasingly replacing procedure manuals and regulatory interpretations as automated systems take on the tasks of deciding who is eligible for what treatment. *If we don't document the software itself, we cannot reconstruct the process.*

Perhaps some radical examples can help. To document electronic music, or computer art, do we want the compositions or the programs or both? If the bridge collapses, and I don't have both the stress modelling algorithms which operated in the system in which the bridge was designed and the assumptions which the designers actually fed into the system, where will I assign the blame for the collapse? Would the calculations (the record) help? This last example is typical of the kinds of issues which are arising in the evolving case law of software liability; issues which are familiar to archivists concerned with documented accountability. In the past, an engineer who used a faulty method of determining stress would clearly have been culpable, as would a government official who mistook the intent of a regulation and misapplied it. It is not clear today, and what is more, we cannot know without documenting aspects of process we have previously been able to live without.

Some of these problems inspired the work I recently completed for the Computer Museum in Boston and the Technical Report on Collecting Software which grew out of that study.

My second set of challenges arises from a need to represent the contents of our cultural collections, be they paintings, shards, or documents, and is closely associated with the first problem, sharing a focus on the wholeness of social processes. The inadequacy of keyword indexing, which leaves undefined the relationship between keywords, is well known. A variety of methods of avoiding false associations have been adopted in many systems, including keywords in context and keywords qualified (as they are in most two level back of book indexes). The computer has introduced an easy variant on qualified keywords - keywords in rotatable phrases - which permits one keyterm to qualify another and each to appear first in the proper place in a printed index. Finally, there is increasing pressure to adopt natural language searching, with proxim-

ity of terms serving in lieu of constructed headings.

Unfortunately, none of these methods represents relationships between terms very well. This problem is one which increasingly attracts artificial intelligence researchers who can now take advantage of the substantial knowledge of semantic relations built by linguists over the past two decades. And it is engaging me, in some work I am currently doing with the National Security Archive (a non-partisan organization which acquires declassified documentation of American foreign and security policy for analysis and dissemination). The challenge is to use insights from semantic analysis in a practical indexing system which must drive printed indexes but which is able to retain as much of the context of term use as possible both in print products and when used as a part of an on-line searching system. Indeed, the ultimate aim is to move beyond the limitations of natural language, full-text, based searches by employing context defined "frames" particular to the foreign affairs/security knowledge domain and derived semi-automatically from the indexing system.

All this sounds heady, but it is really quite simple. Imagine that we write a sentence (or multiple sentence abstract) describing an object or document. The sentences are comprised of words which belong to specific potential vocabulary lists. Such a set of linked vocabularies, for example, are being constructed by the Art & Architecture Thesaurus which has 27 hierarchies of terms, each controlling a single "facet" of description. The relationship between facets in a description within any one domain of discourse is limited. In art criticism, we can not logically talk about Rococo smelting or about terra cotta architects. We could, logically, talk of materials and processes, even if feather smelting is not an actual combination, because the definition of processes is that they act on materials. What this means is that we can construct vocabulary controlled descriptions with complex relationships and qualifiers. Can we derive printed indexes from these and can we search them automatically in such a way as to retrieve on relationships as well as on entities?

At the NSA we are launching a project to do both, as far as possible, with off-the-shelf small system software and document analysts whose training is in the subject area, not in indexing. The indexing issue becomes how to semi-automatically parse a complex sentence whose semantic structural variants are known, in order to generate index phrases consisting of 3 or at most four lead terms which can then be rotated in a printed index. The search issue becomes how to identify "frames", or social situations, which are implied by various semantic models and allow for searching by frame matching. That is to say, how to recognize that People (Samuel Adams) with Social Roles (Citizen) participate in Events (the American Revolution) which occured in Time (date ranges) and were perceived by persons (George III) with Social Roles (King of England) as Concepts (Uprisings) etc. in such a way that if I have an item associated with a frame called uprisings I will be able to provide values to fill the implied process roles of instigators (Person participating in Event of type uprising), of public authorities (Person or Corporate Name against which Event of type uprising is directed), of consequences (Event or Concept resulting from occurence or anticipation of Event of type uprising), etc. so that if my query invokes the uprising frame, and my document description invoked the uprising frame, I will retrieve the document even if I did not request any of the facets specifically indexed! This is similar to the extension which takes place with thesauri when I use a different term (not preferred or broader) for the same concept, and can retrieve the correct document, except that instead of operating simply on the term level, we can now move to the level of the meaning of the document description and the query as a whole.

As we struggle with designing document and object surrogates in our information retrieval systems, we should consider how to better capture the relationships between descriptive terms in order to support researchers with a variety of perspectives and more subtle research problems. (Instigators need to be searched without getting victims and objects of revolutions!) I should note that there is some impressive applied research going on in these areas. For a taste of it, I recommend a May 1987 article by Suzanne Humphrey and Nancy Miller of the National Library of Medicine in the Journal of ASIS, and a July 1987 article by Natasha Vieduts-Stokolov of BIOSIS in the same journal. I am personally indebted to Pat Molholt whose insights in a recent unpublished paper on the "Development of term relationships for the enhancement of semantic networks and hierarchically structured thesauri" convinced me that practical application of research findings in this area was well within our grasp and that thesauri such as those developed by the AAT can best be used as vehicles to this richer context and relationship capturing approach to document and object description. Indeed, Pat graphically represents the AAT as authority control over each term in a complex description.

TECHNICAL REPORT ON COLLECTING SOFTWARE

For forty years software has been an important creative product of our society. Its intellectual, social, economic and political impact has shaped the contemporary world and lent its name to an age, yet the community of culture preserving institutions has failed to document the evolution of software. Not a single archive or museum devoted to software exists. No substantial collecting of software history has taken place. Yet software is being written every day which defines the way in which we work. Bureaucracies (including governments) are entirely dependent upon software to faithfully execute the policies (including laws and regulations) which they have established, yet archives, those guardians of bureaucratic accountability, don't retain software. Popular culture and the arts have both been transformed by software, yet museums have yet to collect it. To archives and museums, software is still alien and insubstantial.

This report examines the history of software and its influences on our society and addresses the barriers to collecting software as a cultural record. It identifies essential policy distinctions which administrators will need to consider between software collections and other collections of archives and museums. It examines the ways in which software can best be described, made available to researchers, and exhibited and it proposes a framework for a descriptive vocabulary. And it identifies the physical requirements and management issues associated with the retention and storage, retrieval and use, of software in cultural repositories.

An earlier draft of this report was prepared for the Computer Museum in Boston as the framework for a discussion with staff of the Smithsonian Institution and the Charles Babbage Institute on establishing a national software collecting consortium. It became clear in the course of that project that no single institution, or even group of institutions as prestigious and well situated as the sponsors were, could expect to collect the entire corpus of software related materials, to support research on technical, social, financial and cultural impacts of computer programming. The report therefore considers approaches to multi-institutional collecting issues such as collections policy, cooperative acquisition and information sharing. Most importantly, it provides the concrete guidance needed by every cultural repository, for all archives and museums which document any part of modern culture should be considering acquisition of some software as part of their collections, since software is an integral part of that culture.

Available in August. Archives & Museum Informatics, 5600 Northumberland St., Pittsburgh, PA 15217
$45 pre-paid or by subscription to Archival Informatics Newsletter and Technical Report

To subscribe, send mailing and billing address with subscription request to:
Archives & Museum Informatics, 5600 Northumberland St., Pittsburgh, PA 15217

ARCHIVAL INFORMATICS NEWSLETTER & TECHNICAL REPORTS

Part 1 only: **ARCHIVAL INFORMATICS NEWSLETTER**
US$24 prepaid, US$30 foreign address airmail, US$40 billed

.Parts 1 & 2: **ARCHIVAL INFORMATICS NEWSLETTER & ARCHIVAL INFORMATICS TECHNICAL REPORTS**
US$160 billed or prepaid; US$180 foreign address airmail