

# ARCHIVAL INFORMATICS NEWSLETTER

## Part 1 of Archival Informatics

ISSN 0892-2179

SPRING, 1988

Volume 2, #1

### EXCITING HERESIES

The RIAO-88 meeting provided an occasion to observe that archivists and curators have knowledges and skills desperately needed by designers of image and full-text retrieval systems. Two particularly exciting heresies occurred to me, which I had an opportunity to explore with information retrieval experts at that meeting (reported further on p.8).

As image retrieval systems become more widely available, they are being examined by users who are no longer just "wowed" by the appearance of an image on a screen. These users are asking how images can be indexed, using language and symbols, so that the desired image can be obtained. And they are demanding system capabilities to manipulate and analyse the images once they are found.

Curators have a knowledge, recently formalized in such products as ICONCLASS and the Art & Architecture Thesaurus, of the differences between subject and object indexing and they have experience, based on years of work with art historians, archeologists, natural historians and engineers, of the tools required to use an image.

Equally exciting is the realization by full-text document retrieval systems designers that they need to limit searches for texts to portions of the full archive of office documentation, and that the methods for limiting searches involve exploiting what they only vaguely understand about the sources of documents, the systems out of which they are generated, and the genre of the record. These concepts of provenance, series and form-of-material or document type are the stock in trade of archivists.

No matter how efficient information scientists make their full-text searches, methods to reduce the number of documents that must be searched, and the parts of those documents that require "reading", will play an increasingly important role in full-text retrieval systems. Archivists have long employed inferential logic based on their understanding of the way in which records are created and the structural means by which documents carry their messages, to find records with contents relevant to a user query. Now information retrieval specialists are looking for just the methods long employed by archivists.

Archives and museums have direct contributions to make to the design of image and full-text systems if they can formalize their knowledge of image and document access methods. This is a worthy challenge, I think.

### TABLE OF CONTENTS

#### Articles:

Advanced Revelation: A Review	2
J. Penny Small	
Machine Readable Views	5
Thomas E. Brown	

#### Regular Features:

Conferences	
ARLIS/NA	7
RIA0-88	8
Calendar, 2nd & 3rd Quarter	10
In-Box	11
Letters to the Editor	15
News	18
Projects & Proposals	19
Software	21
Standards	22
Technical Report Summary	24

---

# ADVANCED REVELATION

## A REVIEW

by Jocelyn Penny Small, Director, U.S. Center,  
Lexicon Iconographicum Mythologiae Classicae,  
Rutgers University, College Avenue Campus, New  
Brunswick, NJ 08903 (201) 932-7404

"It was the best of programs, it was the worst of programs, it was the age of transparent windows, it was the age of opaque documentation...in short,...some of [the] noisiest authorities insisted on it being received for good or for evil, in the superlative degree of comparison only." Advanced Revelation (AREV) provokes extreme reactions, and, frequently, simultaneously. The "power" of the program is extraordinary; figuring out how to tap that power is also extraordinary.

### The US LIMC Project & Revelation

I shall illustrate its features by describing their implementation in the Computer-Index of Classical Iconography at the U.S. Center of the Lexicon Iconographicum Mythologiae Classicae (US LIMC).<sup>1</sup> The US LIMC, as part of an international project to publish a multi-volume pictorial dictionary of classical mythology, is responsible for classical objects (ca. 800 B.C. - A.D. 400), with mythological representations, in American collections. Unlike most computerized projects, its cataloguing goes below the level of the title for each scene to record the individual figures and elements (animals, plants, architecture, etc.), their types, and what the figures are wearing or holding (attributes).

From the outset it was obvious that a relational database program with fully searchable, variable length, and repeating fields was necessary. Objects currently range from single figures

---

<sup>1</sup> The US LIMC is part of the Library at Rutgers University, and is very pleased to acknowledge not only the support of the University, but also that of the National Endowment for the Humanities, Research Tools Div. and the David and Lucille Packard Foundation. The US LIMC welcomes inquiries.

(statues and gems) to Roman sarcophagi with nearly forty figures. For microcomputers these capabilities appeared only in Revelation (Rev) by Cosmos, now known as Revelation Technologies Inc.. Rev also came with a full programming language.

While I at times keenly felt the absence of the bookcase of manuals-to-the-manual available for dBase and RBase, I was very satisfied with the results. So why have I switched to AREV? Why have I tortured myself with near paralytic trances over the keyboard? (They changed ALL the keyboard commands in AREV, and provided no table of equivalents to Rev.) Three things spurred me on: windows, being able to use lower case at the system prompt (TCL for Terminal Control Level), and, of course, the compulsive need to upgrade. The second release should appear this April.

### Windows

Windows! What a wondrous tool they are. And how marvelously they are implemented in AREV. First, some background on my use of them. The US LIMC has two core files, Objects which have Scenes, and nearly thirty satellite files<sup>2</sup>, which: (1) control the words used, and (2) check their spelling, (3) classify the words (Carnelian is a kind of Chalcedony which is a Quartz used for Gemstones, which are obviously Stone; Myron is Greek and a Sculptor), and (4) index data from the core files for fast retrieval and relating files.

Whether or not a particular satellite file does more than one task depends on its data. Furthermore, certain satellite files are linked to other satellite files. Thus Bibliography checks all bibliographic references no matter where they appear, while Cultures vets the entries for Culture in both the Objects and the Artists file.

The verification of fields in both Rev and AREV is very simple: put the name of the verification file in the input pattern for the prompt for a field, and the program automatically checks the

---

<sup>2</sup> Satellite describes the relative position of the file to the Core files. "Classification", however, is preferred to "authority" to describe these files, since they are more complex in what they record and what they do.

---

entry against the record identifier of the verification file. In Rev, by indexing (inverting) the data from fields in Objects and Scenes into satellite files, I could retrieve all objects made of stone or all scenes from the Trojan cycle, but the process could be cumbersome. Now the program "knows" and keeps track of the linkages.

### A Search

For a real example, an archeologist (from Brown University, naturally), wanted all vases in American collections decorated by the Providence Painter. We started in the Artists' file by first accessing the cross-referenced index to the individual words in the record identifiers (Artist Name) to find all painters with Providence in their name. From the list of four painters we chose the desired Providence Painter. Merely by pushing Alt-F5 we switched to the Object file with all of the objects by the Providence Painter saved in a list, available for browsing backwards and forwards. When a particular object was of interest, Alt-F6 brought up the Scenes on the object, with Alt-F8 taking us to further information about the Title in the Titles file and another Alt-F6 for a "second copy" of Scenes with only that Title and then Alt-F5 for a second copy of Objects for more detailed information about a particular object. Hypertext is here.

This design speeds corrections for fields where full verification files do not yet exist. I invert the data in Objects and Scenes into their respective classification files, select for all records with no record date (all records entered by human hands are automatically date-stamped) to check the data, call that selected list from within the verification file window so that I can switch to all the records in Objects and Scenes with that particular entry with a mere touch of the function keys. If you want to see the selected group of records in columnar form, with Control-F5 you go to Table Mode, although too many records with too many prompts will blow this function.

Two other aspects of windows are important. The windows can be larger than the CRT screen, up to 32K or 180 columns by 180 rows or any combination thereof. No longer are five

separate, but linked, screens necessary for the two core files. One window, or rather template in AREV terms, suffices for each file. All, nonetheless, is not sweetness and light. The implementation of "paging" is jerky, annoying, and absurd in this set up.

### Pop-ups & options

Pop-ups are windows that appear generally when you press F2 at pre-defined fields. While they can perform a number of tasks, I use them to provide lists of record IDs from small classification files. Techniques and Materials, with a little over a hundred entries, are feasible, but Artists, with nearly two thousand, is easier to consult via the related templates. Thus F2 in the Techniques field brings up the list of all possible techniques, and all the data enterer has to do is press the enter key at the appropriate value. Since this field is repeating (multi-valued), more than one technique can be chosen at one time with F9 saving the selections and automatically entering them into the field.

Pop-ups are also used extensively by AREV; HELP lists of your last one hundred commands at the TCL prompt, an ASCII chart, etc.. The system also allows you at any point to go to TCL, and thence to DOS. It always remembers where you are, and moves recursively back to the starting point. I should add that like Rev. the program has not lost any records in my five months of use; it has, however, sometimes "hidden" them from, me.

While the types of indexes in AREV (cross-reference, relational, and b-tree) are more varied than in Rev, allow for upper and lower case distinctions, as well as stop-lists system-wide or individually tailored to particular fields, I still use the Rev program, invert.all, because I can edit the fields containing the inverts. AREV incessantly updates all indexes (in background) by keeping a record of the changes in an indexing transaction file. Changes made at any level, be it template, an R/Basic (as they call their flavor of Basic) program, or separately invoked indexing, will update the indexes.

Here, as throughout the program, the choices for ways to accomplish a particular task are sometimes staggering. Seemingly simple decisions can have large effects. For example, at

---

one point with Rev, I was reluctant to reduce the number of fields in Objects by increasing the use of the classification files. Because Rev did not automatically keep track of the cob-web like connections between my thirty or so files, it was difficult to remember how to traverse them at the time of searching. Now, with the smooth movement between files (templates), I have no such hesitations since actual use is much easier. I should add that the maintenance of the connections and the addition of new ones has become a minor nightmare. A plain, run-of-the-mill pop-up requires the classification file, its template, the pop-up description stored in the pop-up file, and the call from the prompt for the field in the template.

### Shortcomings

Unlike Cicero, I cannot pass over the fact that a number of bugs are due to too early a release and a striking lack of communication between different teams of programmers or, worse yet, no attempt at the reconciliation of obviously changing methods.<sup>3</sup> For instance, the routine that imports Rev screen builds a call to a record in the HELP file that ends with the same field name rather than the prompt number for the template, as used for all new templates. Petty complaint perhaps, but the absence of all help and the concurrent creation of four hundred some odd unreachable entries in the HELP file were no joy to discover and amend. Some routines, such as "preserved fields" in the window design, are not merely half-baked, but downright raw. The intricacy of "paint", their program for designing entry and record templates, can be gauged by the fact that each prompt has forty-three parameters which can (they don't have to be) filled in, and some of these, like pattern matching, allow for multiple entries. Pattern matching can be customized to do just about any pattern you can possibly dream up.

All of which brings us to documentation! Six, count them six, manuals with the poorest indexing seen this side of the mainframe. Not only does the overall index leave out main

references, it still has no entry for "back up". When I asked why, Technical Support told me that since "back up" was not one of their commands, there was obviously no need for an index entry. With that approach, imagine the quality of the documentation. With the complexity of the program added to the sheer innocence of what documentation is about, imagine scrambling through the four main manuals. With the changes between actual implementation and the earlier dates for printing the manuals, think of the results. But do not even contemplate the fact that since Release E, with which I started, every single set of instructions for installation has had errors. Nor should you forget that help is not at the corner bookstore or your local computer center. If the program is so good, and it is, why the poor documentation?

But, let us return to what else AREV does do. It can be either menu (very easy to set up) or command driven or a combination of both. Since all fields are variable in length, only display lengths are defined. In templates, the text either tickertapes for single line displays, or scrolls. You never run out of space. It has a mechanism for doing forms, reports, or whatever you call those things that are not columnar lists. Alas, its previous forms, while not splendid, are better than the current incarnation, which does not adjust dynamically for changing amounts of data, such as the "brief" statue or the "long-winded" sarcophagus. You can, however, print any template (screen) with a mere Alt-P. (The Rev program GFORM is on the compatibility disk). A new and improved report program is promised.

### Utilities

AREV acknowledges that no one is perfect, and has provided a set of powerful routines for global updates. An intermediate transaction file is built so you can compare the original with the changed version. One can also do global swapping of strings within a field as in word processing; but, of course, it took me an afternoon to figure out how it worked. Moreover, much, like the pop-ups and the related windows, it is fully automated so that no programming is needed. The scholar,

---

<sup>3</sup> A wise and witty book on software engineering is: Frederick P. Brooks Jr., The Mythical Man-month, Boston, Addison-Wesley, 1982

---

if like me, can twiddle to his heart's content until he gets it just right or righter.

I have not used their network version, which is supposed to be excellent. Nor have I had any reason to implement any of their provisions for security. You can define separate users with any kind of ideosyncratic access you might want. The system automatically resizes files for optimum retrieval -- a real plus-- but, at the same time, the data now resides in only two files that can grow to enormous proportions. My Scenes files are each over 1800K. You must use a DOS program that can split files between disks. The program now comes on ten diskettes -- some of which you won't need forever (tutorial, Rev compatibility), but do not even kid yourself that AREV can be used solely on a floppy system.

While I have focussed on the new things that AREV does, that Rev does not, it should be noted that both programs use dictionaries to hold field definitions. "Symbolic fields" are R/Basic programs up to 32K. I use them for plucking the city, museum, and inventory number from the Objects file into the Scenes file. Such "joins" are not real joins. That is, no transfer of data need occur (it can, if you want it), and no third file is built (a great saving of time and space). These fields can also be used to do calculations and special formatting, like underlining.

### **The Bottom Line**

If you are already using Rev, should you upgrade? Emphatically yes. But emphatically steel yourself for an excruciating experience. If you are using something else or not using anything else, should you buy AREV? Absolutely. Buy AREV at the cheapest price (list \$950) you can find. Then invest your savings in the long distance telephone calls to Technical Support and/or one of the independent consulting firms. (I do both). AREV really will do what you want it to do (all mortals are reasonable, you are mortal, therefore. . .) Getting there, however, is half the agony.

J.C.P

---

## **MACHINE READABLE VIEWS**

Thomas E. Brown National Archives & Records Administration Washington, DC 20408
---

For the past several months, I have spent a lot of energy as chair of the program committee of the forthcoming conference of the International Association for Social Science Information Service and Technology (I-ASSIST). As anyone who has ever served on a program committee can attest, it is both time consuming and rewarding. The rewards stem primarily from the opportunity to see the cutting edge of one's profession. Because I-ASSIST brings together professionals interested in the acquisition, administration, use and preservation of computer data files, the last three months have offered me insights into the evolving interests of colleagues concerned with the archival administration of electronic records.

The proposals and final program reflect the growing role of microprocessors or personal computers in this, as in other spheres of contemporary society. One session is devoted exclusively to the analysis and dissemination of data using the microcomputer and several sessions dealing with archival administration or use of computerized data include at least one paper on the applicability of microcomputers. In one series of papers addressing the use of information on individuals in modeling for social policy analysis, two authors propose tutorials on how to use microcomputers to analyze microdata for policy analysis.

The I-ASSIST program also reflects the growth of automated cartographic systems. While geographic information has long been digitized, in this decade such systems have been riding the tidal wave of the advances in computer graphics. Over the past few years, as developments have taken place in the CAD/CAM (Computer Assisted Design/ Computer Assisted Manufacturing) arena, the I-ASSIST program has routinely included one or two papers on cartographic systems. The upcoming conference will devote two full sessions to digitized cartographic

---

systems! The first set of papers will focus on managing these systems. One paper will examine the technical problems associated with their creation and maintenance and the need for standards. The second paper will discuss one such standard emerging within Canada, the Map and Chart Digital Interchange Format (MACDIF), a telecommunications standard for cartographic information. The last paper outlines a current system within the U.S. Government, and relates the previous observations to a specific mapping system.

The second session on automated cartographic systems was designed to be provocative. Each paper will discuss whether reference services should provide data to the user in cartographic form rather than (as is now customary) as raw data or crude tabulations. Michal Peleg of the Social Science Data Archive of Hebrew University in Jerusalem argued in her proposal that: "clear and meaningful representation of data is a universal requirement from every scientist, student, and executive. Quantitative data in the social sciences is usually presented by tables of rates, frequencies or percentages, while the more advanced data presentations involve statistical charts. However, geographically oriented data needs more sophisticated presentation techniques. Otherwise the spatial character of the data is largely obscured. The map then provides the necessary medium for presenting areal relationships of spatial data, which are so essential in comparative and regional studies. Until the eighties, most of the statistical maps were manually produced by a few geographers and large agencies. Now, due to the decreased cost of computers and graphic peripherals, as well as the availability of sophisticated graphics software, this challenge is within the reach of public use in a large number of applications." As a result, she argues, data repositories that have data sets with spatial characteristics such as census samples and small area statistics, household surveys, government macro-data and election results provide their users with information in cartographic form.

Peleg's proposal to use cartographic presentation of archival information relates to a far broader question. Today, archives and data

libraries are generally providing their users with raw data, primarily on magnetic tape, and with the necessary documentation to interpret that raw data. Certainly, providing the data on magnetic tape really does not help the PC user. So what are the alternatives? What should be the standard for reference service in the future? What level of service should our profession provide? These questions remain unanswered. The discussion of cartographic output as a reference vehicle is one solution to this broader and increasingly complex problem which we have not really addressed or adequately defined.

The I-ASSIST program reflects a new emphasis within the profession on reference service. If we divide the functions of data archives and data libraries into three broad categories, acquisition (records scheduling and disposition, appraisal, solicitation policies and accessioning); administration (description, cataloguing, technical documentation standards, and preservation); and reference (dissemination technology, confidentiality, and end-user needs), conference proposals relating to acquisition were nearly non-existent. Proposals relating to actual administration of computer files within a repository continued to hold their own. The growth area, to the point of eclipsing acquisition, was reference service. And this emphasis on reference seemingly concentrated on the research use of the information. What does this say about our profession?

It may reflect the growing influence within I-ASSIST of librarians who are realizing the importance of machine-readable records and who have traditionally had a strong interest in reference service. Originally, I-ASSIST attracted archivists with responsibility for automated records within a traditional archives, and only later, data librarians who usually found themselves administering a collection of machine-readable data files outside of traditional libraries. As a result of the recent influx of professionals working in traditional libraries, management issues may have been overshadowed by reference service. The profession continues to expand to include new people with new emphases.

---

## CONFERENCES

### Art Libraries Society/North America

The (ARLIS/NA) annual meeting in Dallas, February 4-11, reflected the growing interest among visual resources managers in automated systems, indexing, intellectual access approaches, and the MARC-VM format.

Helene Roberts, Curator of Visual Collections at Harvard, sounded what was to be a theme of the meeting in the first session when she criticised existing reference works that provided subject access to visual collections for not being sufficiently complex, consistent or convenient. Her use and explication of ICONCLASS suggested that it is a system that is sufficiently complex to represent works of art, but she presented no assessment of its consistency nor any mechanism by which it might prove more convenient. Jack Robertson also addressed the convenience of art reference data in his well thought out guidelines for reviewers of reference books, but a presentation by Marshall Lapidus on the merger of RILA and the RAA which followed revealed the nature of these problems in all their concrete ugliness. The attempt by the Getty Trust and CNRS to merge the two principal bibliographic resources in the field is more than three years in the making, and will represent at best a compromise.

Compromise was the subject of the next session I attended on indexing and search strategy using the Art and Architecture Thesaurus (AAT). Amy Lucker presented the use of the AAT in the AVIADOR project at Columbia University in which only single terms were assigned, simplifying the indexing but leading to false drops and sacrificing the hierarchical structure of the vocabulary. Bethany Mendenall examined the implementation of expressions, or sets of AAT terms in specific systems, and Murray Waddington presented the way in which the AAT has agreed to be represented in MARC using indexing strings which can be sets of expressions (see also Standards, p23). In a spirited

discussion led off by commentaries from Pat Molholt and Jim Anderson, the audience debated the desirability of complex representation versus simpler, more easily implemented approaches and the barriers to national implementation of hierarchical faceted indexing vocabularies.

The meeting provided an opportunity for ARLIS members to experience some systems used or developed by their colleagues. Jeanne Keefe Watkinson presented the RPI online public access catalog to 65,000 slides. Joy Alexander and Jo Schaeffer discussed their experiences automating slide collections with standard commercial packages (dBase and SMART). The organizers provided an afternoon to see other ARLIS/NA member developed systems, including a dBase auction catalog file by Beth Dinoff (Dallas Museum of Art), an index to the illustrations in architectural books by Henry Pisciotta (CMU), and a PC/Mac Videodisc driver from the University of Iowa. The modesty of these implementations was contrasted with extravagant claims made by computer systems vendors at Infomart ("the world's largest computer showroom") where I visited CPT's full text retrieval, IBM's Infowindows, and Image Sets showing off their use of Pictureware on an ATT system, and was impressed only by the hype.

One late night session was designed to capture high-technology imaginations. Karl Miller (U.T. Austin) led off with an enlightening history of the CRT and the potential of electronic imaging. Madeline Nilsen conducted a tour of projects that combine image and text ranging from Athena, the Aga Khan, and Boston Architecture Projects at MIT through museum videodiscs to such commercial efforts as the Chadwyck-Healey publications on American architecture. Merily Snow presented the U.C. Museum (Berkeley) prototype digitizing system described by Howard Besser in a recent issue of Museum Studies Journal. For me, the most exciting event of the conference was the final fifteen minute talk by Marc Rorvig of Project ICON, at the University of Texas, Austin. Dr. Rorvig introduced the concept of a visual thesaurus - images linked in hierarchies with defined relations - as a means

---

of conducting a "wordless search" and of achieving greater consistency between indexers. He and his colleagues are constructing such a visual thesaurus for NASA. Rorvig suggested that it be examined in conjunction with the AAT, bringing the meeting around in a delicious full circle. Whereas we began by struggling with how to use language to index images and make them available to new users from diverse disciplines, we ended by using images to index concepts in disciplines whose language was alien to us! Although what he described was far from the image recognition capability some in the audience mistook it for, it was elegant for the simple use it made of human intelligence to transport the user from one image, through a series of related images, to the one needed.

---

### **RIA0 88: Conference on "User-Oriented, Content-Based, Text and Image Handling"**

Over 700 people from 21 countries gathered at M.I.T. March 21-24, for RIA0-88, a conference organized by the Centre des Hautes Etudes Internationales d'Informatique Documentaire. (Late in the week, I discovered that RIA0 stood for Recherche d'Information Assistee par l'Ordinateur).

Participants were, like myself, largely first time attendees drawn to the conference by its emphasis on hypermedia and full-text. Nor were we disappointed, although the organization of sessions left much to be desired. The 75 twenty minute papers were delivered much as they read in the three volume conference proceedings, with no opportunity for real discussion, and the demonstrations of systems (almost all prototypes) were often hard to locate. Nevertheless, the meeting displayed the range of activity in these areas by researchers from the Artificial Intelligence, Information Retrieval, and Image Access communities. The emphasis on demonstrable systems made the reports broader than those at similar ACM and IEEE sessions, and together with a ten page limit, this made the papers seem somewhat superficial. I left excited by many of the ideas being explored but

discouraged by the evidence that in hypermedia as in full text, our prototypes are far from implementation and scaling up to production exposes unresolved problems.

Karen Spark-Jones, in her keynote address, characterized the problems as inherent in heterogeneity:

- We have objects of different kinds. Can they be characterized in common ways?
- We have data with different levels of granularity. Can they be connected?
- We have different functional uses. Can they be supported by the same representation?
- We have different relevance criteria. Are the apposite?

In sum, she asked, can we go beyond aggregative systems to integrative ones?

A week of looking at large screens, many windows, bright color images and full-texts, large optical stores and attractive user interfaces, could not obscure some very serious problems. We still can't capture very large amounts of text or images from paper formats very efficiently. Automatic analysis indexing and analysis are still in their infancy and content/structure relationships are not very well understood. Freeing users to wander gets them lost quickly, and even when they find things, we have too few useful analytical tools attached to retrieval systems to meet users' needs. I admit a certain production oriented bias in stating that I found the crude tricks more exciting than the subtle methods.

Among the practical, implemented, concepts was Lee Hollaar's Utah Retrieval System which keeps an index of words in full-text documents only. It doesn't identify relative position or specific location, but string searches are fast enough, once the document is located, to execute such sophisticated search, and the overhead for simple indexes is only 15%. Yaacov Choeka reported on a crude method of identifying phrases, and names, for indexing, which simply locates all occurrences of 2,3,4 and 5 word phrases in millions of words of text, and provides an indexer with all phrases occurring more than x times. The resulting lists identified

---

potential indexing phrases with a precision greater (often much greater) than 50%. Susanne Humphrey reported on the National Library of Medicine's automation-assisted indexing, emphasizing the useful assistance over any notion of automatic indexing. NLM and OCLC displayed "orienting" interfaces that locate "hits" on tables of contents of large texts so that the user can judge their likely interest based on location and clustering of occurrences.

This is not to suggest that the prototypes lacked useful ideas. Howard Besser showed off some very nice and functional image analysis routines developed at the University of California at Berkeley and applied to art, architecture and geography. Andreas Dangel displayed an automatic document structure analyser that exploited simple rules about basic structural components of documents to locate areas of documents with specific information contents (cf., addressee of letter). And Y. Tanaka of Hokkaido University discussed a creative approach to coding characters that can't be OCR'd with a simple 3 bit per character scheme that produced surprisingly high retrieval precision for words of more than five letters. A number of clever AI techniques were presented as components of other systems.

An exciting aspect of the conference was the number of vendors, especially those new to the US market, whose products are being turned to the retrieval of image and full-text. Almost everyone showed a prototype developed with a museum in order to show off videodisc or optical disk stored images in conjunction with their information systems. (We joked at one point that soon every museum in the world would have a system, each would be unique, and none of their vendors would still be in business. Sophiatec, an offshoot of the University of Nice, displayed English and French versions of **Sophiadoc**<sub>TM</sub> (C-DOS for the PC-based) and its pilot project with the Louvre (optical disk) and the City of Nice (videodisc). **Sophiadoc**<sub>TM</sub> boasts an integral thesaurus generator, variable length fields in a relational model, and the image link. **SIGMINI**<sub>TM</sub>, a product of the Ecole Nationale Supérieure des Mines showed off a declarative data

structure in which a record consists of numerous parenthetical declarations of the form element=value, that enabled them to describe all and numerous parts of complex mosaics, including non-hierarchical relations (all tied to a videodisc of mosaics). Honeywell-Bull displayed **MISTRAL**<sub>TM</sub>, the software at the heart of **IMAGO 2**<sub>TM</sub>, the system used to manage the French national radio and television archives. GECI International exhibited **Hyperdoc**<sub>TM</sub>, its multimedia documentation package in C for PC's with CD-ROM and the Centre de Recherche en Informatique de Nancy displayed a visual thesaurus of mycology knowledge called **Mycomatic**<sub>TM</sub> to be marketed in July by FUTUR\*VISION.

A number of US academic offshoot products were shown as well. Robert Kraft showed off the \$25 CD-ROM with full-text of the Bible in Greek, Latin, Hebrew, and English as well as much other material for Septuagint Studies being distributed by the University of Pennsylvania. Cognetics Corporation demonstrated **Hyperties**<sub>TM</sub> version 2.3 (\$249), developed by Ben Schneiderman at the University of Maryland for "hypermanual" editing. Brown University demonstrated Intermedia, its hypermedia courseware environment, but similar projects seen at EDUCOM meetings, from M.I.T. itself (Athena), CMU (Andrew), and numerous other U.S. campuses, were absent.

Now that digital image storage systems are here, the problems of digital data capture and the absence of digital analysis capabilities (or needs?) invites us to re-examine some assumptions. For the first time in a few years, it was respectable again to question the need for digital storage of images. Quite a few of those to whom I spoke admitted to reconsidering videodisc and even microform storage of images with digitization on demand, either at the local workstation or from a file server. While archivists and museum staffs need to remain alert to changes in image storage and transmission technologies, those who are not acquiring systems immediately can use this time to analyze the uses their clients make of images.

---

**CONFERENCE CALENDAR  
2nd & 3rd Quarters 1988**

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

**April 29-30** Library Descriptive Standards: An Introduction for Archivists, a workshop sponsored by the Society of American Archivists in Chicago [600 S. Federal St., Suite 504, Chicago, IL 60605; 312-922-0140] Also to be offered in Atlanta, prior to the SAA annual mtg.

**May 15-18** Artificial Intelligence, the ASIS mid-year conference, Ann Arbor, MI [American Society for Information Science, 1424 16th St., NW, Washington, DC 20036]

**May 22-26** Basic Videodisc Design & Production Workshop sponsored by the Nebraska Videodisc Group in Lincoln, NE. [P.O. Box 83111, Lincoln, NE 68501; 402-472-3611]. Also offered July 17-21.

**May 23-24** "The Coming Age of Electronic Text"; Study Group on the Structure of Electronic Text (SGSET), Conference, Carnegie Mellon University, Pittsburgh, PA [Peter Capell, 412-268-8599]

**May 26-29**, IASSIST Annual Conference - "Public Data: Use it or Lose it", Washington DC. [Pat Doyle, Mathematica Policy Research Inc., 600 Maryland Ave. SW, Suite 550, Washington, DC 20024]

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

**June 6-10** Association of Canadian Archivists, Annual Meeting, Windsor Ont. [Secretariat, Learned Societies Conf. 1988, Room 2129 Windsor Hall North, Windsor, N9B-3P4, CANADA; 519-253-4232]

**July 8** Management Strategies for Disaster Preparedness, a workshop sponsored by ALA

Resources and Technical Services Division, New Orleans, [RTSD, 50 East Huron St., Chicago, IL 60611; 312-944-6780]

**July 19-23** Microcomputer Applications in Visual Resource Collections [Fine Arts Contin. Education, Fine Arts Building 2.4, Univ. of Texas, Austin, Texas 78712; 512-471-8862]

**July 20-23** National Association of Government Archives and Records Administrators (NAGARA) Annual meeting, Annapolis, MD [Stephen Cooper, Maryland State Archives, 350 Rowe Blvd., Annapolis, MD 21401; 301-974-3914]

**August 21-28** Preservation of Black & White Photographs, two workshops [Rochester Institute of Technology, College of Graphics Arts & Photography, Tech. Ed. Center, One Lomb Memorial Drive, Rochester, NY 14623; 716-475-2757]. "Identifying image forming processes, handling & storing" August 21-25; "Copying and duplicating", August 26-28.

**September 14-17** American Association for State and Local History, Annual Meeting, Rochester NY [AASLH, 172 Second Ave., North, Suite 100, Nashville, TN, 37201]

**September 21-24** Second Annual Museum Documentation Association Conference, "Terminology for Museum Documentation", Cambridge, UK [MDA, 347 Cherry Hinton Rd., Cambridge, CB1-4DH, England]

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

**CALL FOR PAPERS**

Proposals for papers, accompanied by abstracts, will be accepted until April 30, for the Museum Computer Network Annual Conference, to be held in Los Angeles, October 26-28. Contact David Bearman, Program Coordinator, 5600 Northumberland St., Pittsburgh, PA 15217 (412-421-4638)

---

## IN-BOX

### REPORTS

Bilfinger, Monica; Buysens, Danielle; Jost, Karl; Meles, Brigitte & Zurcher, Ronald. Vers une banque de donnees culturelles et artistiques suisse: Conception de l'informatisation et de l'echange de donnees dans le domaine des beaux-arts et des arts appliques, Berne, Academie suisse des sciences humaines (ASSH), 1987, 117pp. also published in German

This Swiss working group report on the possibility and value of information exchange in the arts (commissioned in February 1986 and completed in June 1987) sounds many themes that are familiar from similar "towards national systems" studies elsewhere. While its conclusions are, appropriately, tied to Swiss circumstances, the systematic manner in which it addresses options, and its application of rigid criteria of technical and political implementability, should be emulated. As is often the result of the best of such efforts, the recommendations are modest and focus on common standards for description and the exploration of specific cooperative ventures, ranging from information sharing consortia to shared videodiscs. They emphasize the need for improving local inventory control, and creating a union list over the allure of an online database, and the desirability of shared vocabularies over interactive authority control. [Available from Mrs. Anne Christine Vogel-Clottu, Swiss Academy of Humanities, Hirschengraben 11, P.O.Box 2523, CH-3001 Bern, Switzerland]

Horvath, David; The Acetate Negative Survey: Final Report, Louisville, KY, University of Louisville, 1987, 91pp. \$10.00 from the Photographic Archives, U.Louisville, 40292

The final report of this National Museum Act funded project will be of great value to those who have responsibility for historical collections of cellulose acetate film (many of whom participated in the survey that informed the study). Following a detailed history of the

problems associated with degradation of "safety" negatives dating from 1925-55, and an account of the various manufacturing processes involved, the report summarises the technical literature on cellulose acetate film stability and factors in its degradation. Survey findings are summarised and some conclusions are drawn leading to tentative recommendations. Excellent bibliographies and notch reference identification information are appended.

Our Memory at Risk: Preserving New York's Unique Research Resources. A report and Recommendations to the Citizens of New York by the New York Document Conservation Advisory Council, [Albany], 1988, 56pp.

Building on the foundation of a 1984 report "Towards a Usable Past", which stated that "preservation may be regarded as the most important historical records issue in New York today", this report advances a practical program of action at all levels and by an astonishing range of tactics. Handsomely and clearly presented, without jargon or hyperbole, its first recommendation is to "complete the initial Historical Documents Inventory project, maintain an automated statewide database for historical records collections and repository acquisition policies, and provide for updating the system"!

Vision 2000: A Cooperative long-range plan for the Maine State Library, Maine Historic Preservation Commission, Maine Arts Commission and Maine State Museum, [Augusta ME, 1987, 80pp].

A needs assessment and cooperative planning document for cultural agencies in Maine, similar in intent to the 1984 plan developed in New York State, but without the synthesis. Here each agency has developed plans to extend its current functions, and juxtaposed them, but there is no strategy. I was struck by the absence of any attention to documentation issues, and by the claim that "no cost" is associated with the objective of the museum to "pursue an aggressive program of collections management" or the objective of the library to study "information acquisition and data handling within state government".

---

## SPECIAL JOURNAL ISSUES

AICARC: Bulletin of the Archives and Documentation Centers for Modern and Contemporary Art, (#25 & 26, 1986/2-1987/1); ; Special issue on "Computers and the Future of Art Research: Visions, Problems, Projects", 61pp.

This issue of AICARC brings together the views of leading figures in art documentation from throughout the world, in a series of tantalizing, short but stimulating essays on where we are and are going, extending and in some cases revisiting AICARC issues 21/22 "Automation takes Command" published following the 2nd International conference on Automatic Data Processing of Art History held in Pisa in the fall of 1984. Hans-Jorg Heusser, the issue editor, remarks that this issue is a more sober, if not somber, view as a consequence of the hard work of the past several years. While Jacques Thuillier, Oskar Batchmann, Vil Mirimanov and Dimitri Pertsev still herald anticipated revolutions in historiography and art critical methods, most authors are cautious, choosing, like Salome Schmid-Isler to assess experts systems or like Helene Roberts to explore the requirements of research databases, without predicting that they will radically alter theory. Several contributions focus on the critical need for more data in automated form, expressed by Karl Jost in his article "Quantity is Quality", and reflected in projects reported by Laura Corti, Margrethe Floryan-Pedersen, and others. A useful contribution by Brigitte Meles describes the "Databases available to Art Historians Today", and suggests how much useful research we could be doing with automated assistance already. A final set of articles addresses how scholars use, and could use, automated systems, and includes an exciting testimonial by Marilyn Aronberg Lavin of intellectual breakthroughs she made by using such a database as well as an announcement by Marilyn Schmidt of the availability of the Getty Trust/Brown University research on the practices of art historians, published under the title Image, Object and Inquiry (available from the Getty Art History Information Program).

Humanistische Data, 3-87 is a special issue devoted to optical media projects. English language articles include a report on "The BBC Advanced Interactive Video and the Domesday Discs" by Phyllis Gove, "The North West Educational Computing Project" by Ian Robertson and Mike Picciotto, "Recording the Italian Cultural Heritage" by Ernesto Bartolozzi, and "A CD-ROM based Geographic Information System" by Erling Maartmann-Moe. Bartolozzi presents the plans for implementing a variety of optical media systems through the Italian National Library Holding System and National Photographic Archives and discusses systems architecture. Maartmann-Moe defines the functional requirements of Geographical Information Systems in place at the Norwegian Computing Center and explores the implications of having such systems on CD-ROM.

Library Hi Tech #20 (Winter 1987) was devoted to space planning for cultural repositories and the implications of information technologies. Its a hot topic; Museum News promises to focus on architecture in general in its May/June issue. The Society of American Archivists is seeking an author for a special publication on the topic. And I recently received a review copy of Richard W. Boss' Information Technologies and Space Planning for Libraries and Information Centers (Boston, G.K.Hall, 1987, 116p. + index).

Boss begins by discussing all the relevant hardware and communication systems and their space, power and communications requirements. Chapters are devoted to automated library (cataloging and circulation) systems, microform, optical media, telefacsimile and compact storage. A variety of formula's used by different institutions for calculating space requirements are presented along with numerous useful details throughout the text but the concluding chapter, which recommends that a library should retain qualified architects to help it plan, include space for computers and terminals, use standard ratios to plan to storing printed materials and for floor loads and carrel space and allow 175 ft per staff member for work space, suggests that information

---

technologies really don't alter space planning at all. The only "technology forecast" based recommendations are to allow for limited growth of microform areas, use no specially constructed desks, try to minimize glare and use dry pipe sprinkling systems to minimize water damage (not even halon!). Did we need another book to tell us this?

Fortunately, authors in Library Hi Tech are more critical. John Kountz presents the cost comparisons of traditional shelving and industrial shelving using automated storage and retrieval techniques and Michael Gorman assesses the pros and cons of movable compact shelving. David Michaels presents a case for discarding those very traditional planning formula's that Boss recommends. Two superb "roundtables" of librarians and designers address how to build the "forgiving" building, one that permits altered internal space utilization and accommodates changing systems requirements. Case studies of Infomart and other intelligent buildings, the new OCLC headquarters and the remodelling of the California State Library round out an exceptionally useful issue.

Not surprisingly all the authors agree that a multi-expert team is required to design well, what they don't propose explicitly is how that team should record its working assumptions and the design conclusions that follow in order to subject them to scrutiny and critique. Implicitly, any argument from function to form suggests that each design element should withstand such scrutiny, but which assumptions about potential technological developments should be factored? After considering how to cable every possible location and all exhibit floor space, should we ask ourselves what wireless communications within a building would mean? After planning for stacks that require power to move and communications outlets for handheld terminals to transmit inventory information back to a central system, should we consider what implications communicating shelves and containers, or robotic retrieval would have? Designing buildings for longterm futures is in most respects a standard risk management activity - it is odd, therefore, to see no rigorous risk management methodologies in any of the discussion cited. Builder beware.

Museum Studies Journal, vol.3 #1, Fall-Winter 1987; Special Section: "Information Technology & Museums", pp.41-110

Michael Templeton's introduction to this special feature section of the Museum Studies Journal raises questions about institutional behavior in response to technological change that are addressed in only one paper in a collection that seems to have been randomly associated. James Beninger and Georgia Freedman-Harvey take up Michael Templeton's questions in what initially promises to be a study of the impact of interactive technologies on exhibit success and visitor understanding but degenerates into a diatribe against museums the authors call "children's zoo's", that use techniques intended to involve visitors and fails to report any research results.

Two articles in this issue are important reading. Lenore Sarasan's sensible "What to Look for in an Automated Collections Management System" can be read with profit by novice and expert alike. Howard Besser's discussion of "Digital Images for Museums" is an insightful introduction to the potential of this technology. Associate Editor Barbara Thompson has compiled a useful, if slightly ideosyncratic and incomplete, "resources" directory.

On the other hand, the editors made some dubious decisions. Hubert and Stuart Dreyfus have collaborated with Renee Dreyfus in a philosophical discussion of artificial intelligence that bowdlerizes their own book, Mind over Machines and doesn't try very hard to be applicable to museums. The editors have chosen to reprint David Williams' relatively unsatisfactory history of museum automation from his more unsatisfactory Guide to Museum Computing.

## ARTICLES & BOOKS

Bender, Avi; "Optical disc technology for records management: A user's perspective", The Electronic Library, 1987, vol.5 #5, p276-81

Reports on the partially implemented, or pilot, Nuclear Regulatory Commission system for storing text and image of its licensing records on WORM.

---

Genoways, Hugh H.; Jones, Clyde & Rossolimo, Olga L., eds.; Mammal Collection Management, Lubbock, TX, Texas Tech University Press, 1987 219pp.

These papers, given at a workshop held in August 1985, are arranged in three sections. The first three papers define the importance of mammal collections, the need for their management, and the history of mammals preservation. Seven papers in the second section are devoted to computerization issues. The final section reports on mammal collections in Australia, Hungary, Spain, Latin America and India. Terry Yates' introduction to the value of collections as evidence is a useful contribution to the literature of museology and as valid for other disciplines as for mammal collectors. A paper by Daniel F. Williams on computer hardware selection is surprisingly fresh (given its age), but won't remain valid much longer. A few useful insights about aspects of mammal collections and their automation can be salvaged from the reports on systems implemented at the Carnegie Museum, the National Museum of Natural History, the Royal Ontario Museum, the Field Museum and Texas Tech, but the gulf between 1983/84 and the present is only magnified as we read of batch query systems, punched card entry and mini-computers operating with 64K.

Johnson, Susan; "The Birth of an Information System", Information Retrieval & Library Automation, Oct 1987.

The LSS is characteristic of the large, integrated, multi-agency, on-line information systems that are proving so problematic to archivists; the history of its design and implementation is, therefore, of special interest.

Moffett, Jonathan, "Computing in the Department of Antiquities of the Ashmolean Museum", Archaeological Computing Newsletter, # 13, Dec. 1987, p. 15-20

This account by the Chairman of the Museums' Computing Group of progress in his own institution details what an alert professional can do with few resources.

Veccia, Susan; "Full-Text dilemmas for searchers and systems: The Washington Post Online", Database, April 1988, vol. 11 #2, p. 13-32

Nine different versions of the Washington Post are available electronically in as many different searching systems. Understanding what works, and doesn't and why is critical to anyone designing full-text systems; incidentally the article may help librarians and archivists make decisions about what formats to keep newspapers in.

## NEWSLETTERS

Quite a number of the "Newsletters" I receive are serially issued promotional literature, available free from commercial firms. Some of these contain quite valuable information on a regular basis. Among those received recently are:

FAX Facts [Paper Manufacturers Company, Office Products Division, 24 Triangle Park Drive, Cincinnati OH 45246]

Government Publications News [Bernan Associates, 4611-F Assembly Dr., Lanham, MD 20706-4391]. New in 1988; the first three issues include quite useful general information, as well as order forms.

MAPS Newsletter [Mid-Atlantic Preservation Service, Lehigh University Mountaintop Campus, Bethlehem, PA 18015] MAPS is a microfilming service organization, and its newsletter includes a "Technical Issues" section of general interest to preservation microfilming officers.

The Scanner; A Barcode Newsletter [WKM/NIDI, 88 Westpark Rd., Dayton, OH 45459]

Other "Newsletters" issued by commercial firms that often include information of general interest are user group publications. Among those recently received are:

Up & Running [Questor Systems, 1005 E.

Colorado Blvd., Pasadena, CA 91106]; MUSE News [Julian Humphreys, Corson Hall, Cornell University, Ithaca, NY 14853]; Society of American Archivists NOTIS Users Group Newsletter [Patty Cloud, Northwestern

University Archives, University Library 1935

Sheridan Rd., Evanston, IL, 60208]; Network News [Conservation Information Network, 4503 Glencoe Ave., Marina del Rey, CA 90292], recently announced availability of it up-load and down-load local systems software and the compilation of an E-mail directory for conservation professionals.

NOTE: SCOPE suspended publication at the end of 1987 (vol.5). According to editor Joseph Raben, it will resume in another form soon.

#### EPIHEMERA:

Directory of Federal Historical Programs and Activities, Washington DC, Society for History in the Federal Government, American Historical Association & National Coordinating Committee for the Promotion of History, 1987, 84pp.

A list of phone numbers and addresses for an astonishingly diverse set of people and projects.

1988 Computer Salary Survey and Career Planning Guide, 14pp. [free from Source EDP offices throughout the U.S. and in Ontario] is a useful source of job descriptions and salary data for employers.

Archival Informatics Newsletter is a quarterly publication of Archives & Museum Informatics, 5600 Northumberland St., Pittsburgh, PA 15217; (412-421-4638). It is edited by David Bearman, whose authorship can be presumed for all items otherwise not attributed. Subscription to the Archival Informatics Newsletter (ISSN 0892-2179) is available for \$24.00 per year pre-paid, US addressed, \$30.00 pre-paid foreign addresses, and \$40.00 p.a. billed. A combined subscription to both the Newsletter and a companion quarterly publication, the Archival Informatics Technical Reports (ISSN 0894-0266) is available for \$160.00 p.a. in the US; \$180 abroad; and will be billed at no extra charge. Individual technical reports are available at \$45. each, prepaid.

## LETTERS TO THE EDITOR

### MARCON

**Ted Durr** [AIRS Inc., 335 Paint Branch Dr., College Park, MD 20742] writes:

"The review of MARCON by David Bearman in the Winter 1987 issue of Archival Informatics Newsletter reminded me of the Broadway play **Fiddler on the Roof**. Remember the lines, "On the one hand..., on the other hand..."? That was David's review. There were some lines we would like to shout at the world; other we would just as soon mumble. Of course MARCON is not a play; users intend it for serious business.

Overall we feel that David was objective, as usual. We have produced, and are continually upgrading, a product that, as another reviewer, also from Pittsburgh, said, "does everything". From a software developer's point of view the truth of that overstatement is not that it does it all but that it tries to do so many things that sometimes, in layperson's terms, it trips.

AIRS spends about half its customer support time teaching users the ins and outs of database management as applied to archives and records management. The other half is spent by our users teaching us - either about deficiencies or about things they would like. We look on this as a partnership and we thank our many friends.

David's review contained some helpful suggestions. His point about the help screen and manual "mess" done by an external firm has been noted and new help screens and a whole new manual are now being shipped. A helpful options bar will appear in a future version of MARCON.

David did make one error. He states that in data entry the screen does not default to the next blank template. If you press Shift-F10 (instead of F10), the data is entered and another blank template appears.

The comments about reports were accurate and this difficult module has been corrected. AIRS suggests that sometime soon David might write a review just of report generators, comparing several packages. Our major report advance is to give users two options: (1) to print reports as record display appears online and (2) to allow users to "paint" reports with row and column

---

formatting. Sorting is possible for multi-valued fields, but, as noted, intra, not inter collection.

The global edit feature was in the version David reviewed but is not in the current version now being shipped. There are some problems of both design and consistency in this feature which we are addressing.

Finally, we expect to have a MARC interface (both for entering and receiving records) in beta testing by the summer. We will introduce it at the SAA this fall in Atlanta. We will also offer workshops in Atlanta prior to the annual meeting. One will be for basic MARCON, and the other for using the U.S. MARC:AMC interface.

### **MICROFORM vs. OPTICAL DISC**

**C. Lee Jones**, previously of the Council on Library Resources, and now President of the Mid-Atlantic Preservation Service [ MAPS, Lehigh Univ. Mountaintop Campus, 118 Research Drive, Bldg. J, Rm. 120, Bethlehem, PA 18015], writes:

"At least for black and white and some mono- and duotone print media, the question is not whether or not to chose optical media for storage and access, but whether or not it meets the needs of users. There is an option that allows the capture of print media in more traditional micro-optical (read microfilm) format with conversion to whatever user format may be required. While the technique is not widely known, it is used by some of the largest micro- and disc-publishers.

A unit known as a composing/ reduction camera will take any of a variety of input streams, including analog or digital or film, and create outputs in an equally broad spectrum of formats. The price tag runs to \$2 million and the units are custom built for the needs of the owner and his services.

MAPS as you may know was brought into being with capital support from the EXXON Education Foundation for the express purpose of developing and demonstrating preservation content capture strategies in the not-for-profit environment. The first step was to evaluate options and after looking at the various options including optical media, it became clear that the most rational current approach was microfilm, a decision that caught many, including EXXON, by surprise. We

continue to keep our ears to the ground and our eyes peeled for any indication that our selected preservation format may be replaced. For the first few months, we felt certain that we would end up using some disc technology for capture, preservation and use. However, a combination of these strategies now seems the most prudent course to follow.

35mm. archival film has been demonstrated to be extremely stable when stored under proper conditions and systematically monitored; some claiming as much as 500 years life. While I would be happy with 300 or 400 years, the point is likely to be moot over those time frames. 35mm is also the dominant microformat from preservation filming, with 16mm a distant and fading second. Microfiche, while lacking the archival life of roll film (100 vs. 300-500 years), is becoming increasingly popular with librarians wanting to provide the most useful formats for their readers. Consequently they now believe that original filming has to be done in a fiche format or 16mm and jacket loaded. These issues forced us to consider what if any alternatives there might be for production of fiche from roll film. Clearly, there are relatively inexpensive machines coming on the market soon that allow us to produce fiche on 105mm roll film, but there are then some very serious material handling problems on the library end of the system.

The composing/reducing camera (CRC), as noted above, is capable of handling a range of inputs and creating outputs in a variety of formats as well. With the professional pressure to continue to explore "electronic/ optical" preservation techniques and the material handling problems of 105mm roll fiche, the CRC starts to become an appealing alternative to the limits of one content capture format or another.

It is likely that MAPS will put together a requirement statement for such a machine along the following lines. The primary input stream is likely to be 35mm film, probably second generation in order to preserve the integrity of the archive master (the original camera film). However, we are also looking at capturing some images in a digital form prior to laying the image on film, so we would also like the option of an analog or digital input stream. This particular

---

feature would be a deduct alternative should it drive the price higher than resources available.

The output streams required from 35mm roll film will include microfiche, 16mm roll film (another early deduct alternative feature), a CD-ROM format (specific technical requirements to be determined later), and preservation photocopy output (yes, even a paper option). The engine of this camera is a high speed minicomputer which may soon shrink to a super micro. Despite that sort of shrinkage, the cost is very high. The first copy will be in the \$2 million range and the second and third in the \$1 million range.

With costs like this, it is not possible for the present MAPS to justify such an investment. However, if the Commission on Preservation and Access is successful in raising \$200 million in funds to do the content capture work in libraries, archives and museums, then we are talking at a much expanded MAPS (with four or five other regional sites) or several partners interested in the same business (preferably not-for-profit partners in order to keep the tax situation as simple as possible) or a combination of these two. The latter is the more likely circumstance. In any case, a CRC will be more readily amortized if it can be operated two or more shifts per day, in which case, it will be capable of handling the output of many 35mm cameras.

All of this just to indicate that there are some options that we expect to come on stream in the near future that do not force one to decide on a limited archival life in order to provide the power and flexibility of the user format. It is possible to have a combination of stable film for archival purposes, even archival photocopy for hands on use, and/or CD-ROM for a "high tech" retrieval and access format. Another virtue of such a combination system is that one does not have to consider double capture if both archival film and CD-ROM is required for material.

The issue of color will be addressed experimentally only after the CRC (and its mates) are in operation. There is some light at the end of the color microform (archival) tunnel with one group claiming far more stable dyes. Actually, if the dyes were archival for as much as 100 years, they would probably be deemed an archival product with a plan to duplicate every

75 years or so. Should that occur, MAPS would be interested in using the CRC in a color mode. So far as I know, there is nothing like that in existence yet."

#### **SAA Automation Program Officer**

**Donn Neal** [Executive Director, Society of American Archivists, 600 S. Federal St., Suite 504, Chicago, IL 60605], writes:

"As you may know, Lisa Weber has accepted a position with NHPRC and will be leaving SAA in Mid-May. While we are celebrating with her about her new opportunity, we are also aware of the gap that her departure will leave in the SAA office: we will be without a Program Officer to direct the remaining months of our NEH Automation Project, to perform staff duties in the areas of automation, and to help SAA plan for the future of this very important initiative.

I am writing to you for assistance in identifying candidates for this position . . .

The NEH grant runs through September of 1989. Our first choice, of course, would be to have someone simply pick up where Lisa is leaving off and see the project through to completion while helping us to chart our future course. We realize, though, that few people may be willing to pull up stakes and move to Chicago for this purpose, so we are quite open to other, creative solutions. These might include having someone on leave from a regular position, the negotiation of at least 50% released time so that SAA duties might be shared with regular ones, a short commute, and others . . ."

#### **Museum Computer Network Director**

In a similar vein, **Suzannah Fabing** [Chair, MCN Executive Director Selection Committee, National Gallery of Art, Washington, DC 20565] could have written me that MCN is seeking a full time executive director, willing to secure funding for MCN projects, double membership next year and expand member services. The MCN office will move to the incumbents' location. A dedicated individual is sought, especially since the salary is only \$30K and the future of the job depends upon the success of the appointee. Applications accepted to mid-May. Interviews to be held in Pittsburgh following AAM meeting.

---

## NEWS

### NARA CHALLENGE

Don Wilson was sworn in as archivist of the United States on December 4. I look forward to seeing how he will act on the challenge he identified in his confirmation hearings on October 20, when he said:

"The ability to administer public records in the computer age depends on the ability of the National Archives to adapt to the changes. It must assume a national leadership position. Old archival theories and practices must be reexamined and adapted to present day needs. The passive role of archivists waiting for old files of paper to be transferred by reluctant agencies must be replaced by active participation in records management."

### ADVICE FROM NARA

As if in response, NARA has announced the availability of two free information packets: "Managing Electronic Records" and "Records Management Software Packages". From the Records Administration Information Center (RAIC), Agency Services Division, National Archives, Washington, DC 20404. Or phone 202-724-1471.

### OPTICAL MEDIA IN GOVERNMENT

Such advice is going to need to address the impact of write-once optical digital media in government according to Karen M.G. Howell, "Federal Government Applications of Write-Once Read-Many (WORM) Optical Disk Systems in Library Hi Tech News, January 1988). In addition, government archivists need to pay attention to CD-ROM publications. E.J. McFaul, in CD Data Report Feb, 1988, pp. 32--36, lists U.S. Government applications agency by agency, and there's hardly an agency untouched. This on top of Linda Helgerson's report in January on the USPS Zip+4 CD, and Robert Williams' report in December on the U.S. Navy Paperless Ship program should wake Government archivists up!

### LIABILITY FOR ELECTRONIC MISINFORMATION

The New York Times reported on March 6 that Terry Dean Rogan, arrested five times in Texas and Michigan for crimes he did not commit, was awarded \$55,000 from the City of Los Angeles which had failed to remove his name from their police files where it had fed into a national crime network. The network is a California based version of the New York State system that Alan Kowlowitz has appraised, and which he has written up for the fall 1988 Archival Informatics Technical Report. At \$55K a shot, we had better devise good retention and disposal procedures!

### NEWS FROM STATE ARCHIVES

The NAGARA Clearinghouse reports that the Delaware State Office of Information Systems is funding a Machine-Readable records appraisal planning project. Hawaii's new archivist Jolyn G. Tamura is exploring automation to improve records management. Mississippi has received a grant to automate the tracking of newspaper conservation data during its statewide newspaper project. North Carolina is adding folder level data on over 200,000 land grant records to its FAIDS system. The State of Washington is integrating all its records management, archives and microfilm management systems using GenCat<sup>TM</sup>, software developed by Eloquent Systems.

### HISTORY OF INFORMATION TECHNOLOGY

Background papers from the Symposium on the Historical Context of the Evolution of Information Technology held at the National Museum of American History Sept. 11, 1987 are available from the Museum upon request. A summary of the meeting was published in "News of the Information Exhibition" a quarterly newsletter from the staff of this large new exhibit planned to open in 1990. Contact David Allison, Curator, National Museum of American History, Washington, DC 20560 (202-357-2038).

---

## PROJECTS & PROPOSALS

### RLG/State Archives Project

The Research Libraries Group project in which the State Archives of Alabama, California, Minnesota, New York, Pennsylvania, Utah and Wisconsin have been involved for the past three years has now completed its NHPRC and CLR funded effort to incorporate state and local government records into the RLIN database, utilize RLIN for collections actions management, develop a list of functions terms to describe the activity of governmental agencies and test RLIN as a mechanism for assisting in appraisal of governmental records. RLG will submit a final report on the three year project to the NHPRC in June, along with a proposal for an even more ambitious second phase.

Having demonstrated the potential benefits of RLIN use by State archives, RLG is expanding the project to include the National Archives and municipal archival agencies, and pursuing the development of form-of-material terminology and authority lists to be used in conjunction with archival appraisal, description and retrieval. Their proposal is, in effect, to augment the database being built by the initial project participants and explore the full value of a national system for archives. One would expect the conclusions of this second phase, if funded, to determine not only the costs and benefits of such exchanges but also provide the standards for data beyond those being developed by Steve Henson for archival descriptive cataloging. In particular, the project should test the viability of inter-governmental records strategies, such as those proposed by the States and echoed recently by Frank Evans at the National Archives.

### Scholarly Text Encoding

The NEH funded conference on the requirements of scholars for text encoding standards held at Vassar College in November (vol. 1, p.62), produced the following framework for the preparation of text encoding guidelines:

"1. The guidelines are intended to provide a standard format for data interchange in humanities research.

2. The guidelines are also intended to suggest principles for the encoding of texts in the same format.

3. The guidelines should

- define a recommended syntax for the format,
- define a metalanguage for the description of text encoding schemes,
- describe the new format and representative existing schemes both in that meta-language and in prose.

4. The guidelines should propose sets of coding conventions suited for various applications.

5. The guidelines should include a minimal set of conventions for encoding new texts in the format.

6. The guidelines are to be drafted by committees on

- text documentation
- text representation
- text interpretation and analysis
- metalanguage definition and description of existing and proposed schemes, coordinated by a steering committee of representatives of the principal sponsoring organizations.

7. Compatibility with existing standards will be maintained as far as possible.

8. A number of large text archives have agreed in principle to support the guidelines in their function as an interchange tool. We encourage funding agencies to support development of tools to facilitate this interchange.

9. Conversion of existing machine-readable texts to the new format involves the translation of their conventions into the syntax of the new format. No requirements will be made for the addition of information not already coded in the texts."

Based on this declaration from the planning conference, the Association for Computers and the Humanities, Association for Computational Linguistics and the Association for Literary and Linguistic Computing, along with a large number of other scholarly organizations, have proposed to the NEH a three year project to draft and approve standards for:

- documentation of encoded texts
- representation of texts at the typographical level

---

- representation of scholarly analysis and interpretation in an encoded text  
- formal descriptions of the syntax of this and other encoding schemes.  
[contact; project director, Dr. C. M. Sperberg-McQueen, Computer Center (M/C 135), University of Illinois at Chicago, Box 6998, Chicago, IL 60680]

### **Cornell Univ. AAP Slide Library**

For several years, Cornell University's Architecture, Art & Planning (AAP) Slide Library has been studying computerization of access to its holdings of approximately 350,000 slides. It recently issued a report on the nearly completed "specifications" phase in which it documents its decision to use videodisc over optical digital disc and a commercial full-text retrieval package.

[contact: Nancy Humphries, consultant, E-TECH, 607-539-6220]

### **Art and Architecture Thesaurus**

In January 1988, the AAT distributed its 17th hierarchy, FUNCTIONS, to its user community, together with revised versions of:

Architectural components  
Single built works and open spaces  
Built complexes and areas  
Settlements, Systems & Landscapes  
Document Types  
Visual Genre  
Drawings  
Styles and Periods  
Materials  
People and Organizations  
Design Elements and Attributes  
Processes and Tools  
Disciplines  
Associated Concepts  
Events

The AAT is now nearly half completed, although it will never be finished in the sense of being frozen. Plans are underway to publish the completed hierarchies in a number of formats and to make them available on-line.

The structure of the FUNCTIONS hierarchy as drafted is quite simple. Seven sub-categories are defined beneath which all functions terms are arrayed. These are:

< analytical functions >  
< collections management functions >  
< economic and financial functions >  
< governmental and legal functions >  
< information handling and communications functions >  
< organizational functions >  
< travel and recreational functions >

AAT Project Director Toni Petersen has recently proposed that the AAT Functions hierarchy be broadened to include terms required by archivists. She has also made the offer to expand and then maintain the AAT's Document Types hierarchy (similar to archival form-of-material), and the Processes and Techniques and Drawings hierarchies need by users of the MARC:VM format. One possibility is to mount all the AAT hierarchies on RLIN

### **Archival Descriptive Standards**

The Canadian Planning Committee on Descriptive Standards has announced a free, "occasional" paper, entitled "Developing Descriptive Standards: A Call to Action". A newly appointed working group on indexing as applied to archives met in January and plans to submit a final report by the end of March. Other groups to meet in 1988 will develop descriptive standards and rules for graphic, sound and moving images at the series, file and item level. Now that the archives format has been integrated into the Canadian MARC Communications Format (new edition available from the National Library of Canada) and the Canadian Committee on MARC has accepted the ACA as a permanent member (represented at present by Hugo Stibbe), descriptive standards activity has an added urgency.

In the U.S., concern expressed in a resolution passed by the Description Section at the SAA annual meeting over the absence of archival descriptive standards has led to the submission of a proposal to NHPRC for a working group to identify where and why standards are required, the scope of standards to be developed, criteria by which to define benefits, and mechanisms for monitoring and changing standards. The proposal asks that the one year project be coordinated by Larry Dowler at Harvard University.

---

## SOFTWARE BRIEFS

### MUSE

MUSE is described by its creators, Julian M. Humphreys and Barry Chernoff as a microcomputer collection management database system. It was developed for ichthyological collections in which numerous specimens are acquired from one locality, taxonomic hierarchies are important, and loan management and label generation are a must. It was designed to be used as a multiuser system with large databases (up to 300 MB) on PC/AT or 286 class machines on a software foundation developed by Softcraft Inc. of Austin, TX. MUSE has been around for quite a while as a piece of consultware (essentially free but requiring the services of the creators to make it run). Now it has a newsletter and a users group. Contact the authors at Cornell University, Ithaca NY 14850.

### MARC Catalog Records Data Entry

Ultracard/Marc<sup>TM</sup> and MITINET/marc<sup>TM</sup> belong to an unusual category of software that facilitates data entry of MARC records on PC's, but doesn't do much with them (beyond printing cards) once they are made. Intended to allow institutions to export records made off-line from the utilities, these packages both accommodate all the formats. MITINET/marc<sup>TM</sup> has substantial validation and "expert" help built in; Ultracard/marc doesn't. Contact:

Ultracard/Marc: Small Library Computing Inc.,  
619 Mansfield Rd., Willow Grove, PA 19090  
MITINET/marc: Information Transform Inc.,  
502 Leonard St., Madison, WI 53711

### Artist Authorities with Images!

Chadwyck-Healey Inc. has announced publication of the New York Public Library Artists File on microfiche, with a name index to be published in machine-readable form (presumably on CD-ROM). Details about the data to be contained in the name index are not yet available. I suspect that Chadwyck-Healey Inc. could be influenced by interested potential buyers if there was an expressed need for an artists name authority file and/or for incorporation of existing artist name files.

### COLLECTION

Vernon Systems has announced several enhancements to its COLLECTION system since the initial September release. Among these are that all data entry and query screens now have the ability to search via multiple cross reference indexes, group as well as individual permission sets may be defined, users may define their own data fields, and the extended ASCII character set is supported. The most extensive enhancement is a parameter driven procedural "harness" which permits a local definition of procedures within a generalized framework. Specifically, the procedural "harness" is said to permit an institution to:

- 1) name a step (and make user notes)
- 2) identify next step permitted
- 3) identify reversal steps permitted
- 4) check for presence or absence of data in a defined field (approval code or prior action)
- 5) check to assure integrity of transaction
- 6) define forms generated by completion of step
- 7) define other actions required by taking step
- 8) update status and action history

While I haven't seen this feature, the concept is similar to the generalized action processor I defined for CMASS (the Collections Management Action Support System) at the Smithsonian in logical models developed in 1983/4 but never implemented there. It could potentially resolve the problems created by local procedural differences within categories of actions such as loans, conservation treatment, or acquisition committee review and will definitely be worth examining closely.

### Willoughby Associates

Willoughby Associates announced the opening of a new office, staffed by Jane Sunderland, Marcy Reed, and Lynn Remington at 11619 Ohio Ave., Penthouse, Los Angeles, CA 90025 (213-444-8994). The Chicago office is still reached at 312-866-7996.

---

## STANDARDS

### FACETTED CLASSIFICATION IN MARC

At its meeting in December, MARBI, the advisory committee for the MARC formats, gave a generally favorable hearing to one of the most far reaching proposals it has ever received. The proposal, by the Art and Architecture Thesaurus (AAT), and supported by the National Library of Medicine, permits catalogers to assign post-coordinated indexing terms in place of pre-coordinated Library of Congress Subject Headings in MARC records.

In this way, a cataloger using the AAT, could construct an "expression" (to use some examples proposed by Murray Waddington), like 'constructivist architecture', or 'Gothic rose windows', or 'Embossed vellum bookbindings' and modify them by a geographical modifier and a chronological modifier in the equivalent of an LCSH heading, while keeping each term component in its appropriate facet or hierarchy. A complete string, as these headings are called, might be: **Modernist industrial buildings - France - Paris. Adaptive reuse.**

or

**Gothic Churches - Italy. Tympana. Photographs - 20th century.**

The proposed format of such a field (tentatively designated as a 65x field in MARC) would be:  
65X

Indicator 1 undefined  
2 thesaurus id.  
\$a focus term  
\$b other term  
\$c facet/hierarchy designation  
\$y chronological term  
\$z geographical term  
\$2 thesaurus code

A proposed coding pattern would be:

**\$a\$c\$b\$c\$b\$c\$2**

but the "focus" term need not occupy the first position, as in the example of "Gothic churches" where churches is the focus. Another plausible pattern would draw terminology from a variety of thesauri, and would be coded:

**\$a\$c\$2\$b\$c\$2\$y\$c\$2\$z\$c\$2**

If this proposal is adopted by MARBI and the Library of Congress, its initial implementation would be as a variant of LCSH and other pre-coordinated subject headings, but its implications are far more revolutionary. Such a system would have the benefit of preserving hierarchical placement of each term, thus in principle permitting narrowest possible facet assignment but searches for terms on parent values. Thus, if I were interested in churches, and a term "cathedrals" or "cloisters" was a facet of some assigned indexing terms, a retrieval system could be implemented that searched down the hierarchy from churches, identified these narrower terms, and retrieved them together with my request for churches, without redundant term assignment on the part of the indexer/cataloger. The potential "explosion" of terms provides many more access points for users and, if implemented for both "preferred" and "non-preferred" uses, obviates the need to standardize language in index term assignment or in query.

While this proposal still has to be approved, the issues it raises for archives and museums are sufficiently important to warrant introducing the possibility here. Comments and suggestions can be addressed to Toni Petersen at the AAT, Phyllis Bruns at the Library of Congress, or Lisa Weber at SAA.

### GEOGRAPHICAL AUTHORITIES

The J. Paul Getty Trust Art History Information Program Vocabulary Control Group, directed by Eleanor Fink, is exploring the sources of machine-readable geographical authority files in order to build on some existing database the International, historical and multi-disciplinary vocabulary for access to humanistic data. A recently completed internal study has identified some plausible bases for such an effort and negotiations are underway to acquire the rights to an appropriate foundation. Because no extant databases have the historical cross-references required, or systematically locate such humanities geo-concepts as linguistic regions, ethnic groups, literary/mythologic locations, building names etc., much elaboration of this basic file will be required for it to serve all the needs of art and cultural research.

---

## ARCHIVES

The Society of American Archivists has begun offering a workshop on "Library Descriptive Standards: An Introduction for Archivists". The workshop is intended to teach archivists about the Anglo-American Cataloging Rules and Library of Congress Subject Headings, issues they are confronting as they implement the MARC:AMC format in library utilities. The workshop will be offered for the first time April 29-30 in Chicago, and again in Atlanta in September in conjunction with the SAA Annual meeting. SAA is still holding its "Understanding the MARC AMC Format" workshops, with the next ones scheduled for June 2-3 at Old Sturbridge Village, MA and June 20-21 in Jackson, Mississippi.

## RARE BOOKS & MANUSCRIPTS

A very useful list of all the rare books and manuscripts standards and how to get them has been published by John B. Thomas III. See his "Standards and Guidelines Prepared by the Rare Books and Manuscripts Section of the Association of College and Research Libraries", Rare Books and Manuscripts Librarianship, vol.2#2, Fall 1987 p.109-112.

## GOVERNMENT RECORDS

NICLOG, the National Information Center for Local Government Records, is now available toll-free on 1-800-284-5456 to give advice on standard methods for managing historical records. NICLOG will also respond to mail sent to NICLOG, AASLH, 172 Second Ave. North, Suite 102, Nashville, TN 37201

## HISTORY MUSEUMS

The "Common Agenda" project, including its task force on documenting collections which is developing approaches to definition of standards for historical collections is seeking input from the profession. Write or call the project coordinator, Mary Alexander, at the National Museum of American History, MBB-66, Smithsonian Institution, Washington DC 20560, (202) 357-4573 or contact the documentation task group chairman, James Blackaby, curator of the Mercer and Fonthill Museums. The project will report on its progress at the AAM meeting.

## LIBRARIES

The Linked Systems Project: A Networking Tool for Libraries, Dublin OH, OCLC, 1988 is a complete history, technical introduction, and applications/implications analysis of the LSP protocols (which support computer to computer communication of MARC data) by leaders in the field. It should be considered must reading for library and archives technical services personnel.

## PUBLISHING & DOCUMENT MARK-UP

Although they've been available for some time, I just received copies of three very useful texts on SGML by Joan M. Smith. These studies were conducted for the British Library and published as British National Bibliography Reports #22, 26, & 27. They have the titles: The Standard Generalized Markup Language and Related Issues (#22); The Standard Generalized Markup Language (SGML): Guidelines for Editors and Publishers (#26); The Standard Generalized Markup Language (SGML): Guidelines for Authors (#27)

The Guideline volumes are appropriately nitty gritty, with all the necessary details and lists. The issues volume covers the gamut from history (with a distinctly British slant), through the participants, and on to implications. All three are recommended reading.

It is increasingly clear that SGML will become an important constituent of all electronic texts and that as text creators and users, it will be useful to understand them even though the actual embedding of codes will, increasingly, be done without direct knowledge of the writer, as it is in such word-processing packages as SoftQuad Publishing Software<sup>TM</sup> by SoftQuad Inc. [720 Spadina Ave., Toronto, Canada, M5S-2T9].

Those interested in the latest updates should consider attending MarkUp'88, the sixth International Conference and Showcase on SGML sponsored by the Graphics Communications Association in Ottawa, May 24-26. Contact: GCA, 1730 North Lynn St. Suite 604, Arlington, VA 22209-2085 or call (703-841-8160).

---

# DIRECTORY OF SOFTWARE FOR ARCHIVES & MUSEUMS

Archival Informatics Technical Report  
Vol.2, #1, Spring 1988

The Directory of Software for Archives and Museums is organized in three sections with a short introduction to the role of such a directory and its use.

The first section describes systems, in alphabetical order by system name, according to general characteristics such as hardware and software environment, size limits, standards, integration, support, utilities and types of applications supported based on information provided by vendors. Pricing and acquisition options, including availability of consultative support and customization, are described.

Section two compares systems by application and sub-systems or utilities. Application comparisons tables are presented for: cataloging & description, collections management, conservation management, education & interactive training, events management, exhibits management, information retrieval, membership/development & fund raising, publications management, records scheduling & disposition, space management, travelling exhibits and volunteer management. Utility or sub-systems tables are presented for: authority control, data dictionary, data entry, graphics, help, query & retrieval, report writer, screen writer, security and text editing/word processing. Each table compares data supported by the system, its functions and features of the application.

Section three indexes the systems by vendor name, hardware and operating systems. A bibliography cites selected published reviews of the products and discussions of systems selection processes.

Individual copies of technical reports are available from Archives & Museum Informatics, 5600 Northumberland St., Pittsburgh PA 15217 for US\$45. prepaid. Subscriptions to the Technical Report and four issues of the Archival Informatics Newsletter are US\$160 p.a., for addresses in the U.S. & Canada, US\$180 p.a. overseas.