

# ARCHIVAL INFORMATICS NEWSLETTER

## Part 1 of Archival Informatics

ISSN 0892-2179

SUMMER 1988

Volume 2, #2

### Remedy for a sick market

The market for archives and museums automation has never seemed so robust, yet I fear it is perilously close to death. On the surface, all the activity we are witnessing is evidence of health, why do I worry that it is sick?

During the process of compiling the Directory of Software for Archives and Museums, I found myself wondering why so many of the vendors of collection systems to whom I wrote had their mail returned unforwarded. After completing the publication I knew the answer - no one buys archives and museum collections systems. The sales figures returned by vendors of membership and development packages were relatively respectable - 88 to 500 installed. Generalized information retrieval products like STAR, MARCON and INMAGIC also have sales in the hundreds. But the sales of archives and records management systems ranged from none to 45 and of museum packages from none to 15 (excluding the MDA MODES data entry system and the ISIS ARKS zoo information system which have relatively captive markets). The mean number of sales of 6 purely archives and records management systems was 17 (even allowing for some healthy exaggeration on the part of vendors) and that of the nine specifically museum collections systems was 4.5! No vendor can survive in such a market.

The consequences are extremely worrisome. Not only are good ideas and substantial investments, being wasted in implementations that don't find a sufficient market to support them, but vendors starved for their next contract will be unable to provide the kind of support, growth and enhancement that buyers need. Investments made by buyers today are more likely than not to be in systems that won't be marketed two years from now!

Why do we find ourselves in this predicament? Some of it is that systems are easier to develop than they were in the past, and some is that everyone wants to recoup their inhouse investments by offering a system for sale, but the largest part of the problem is that archives and museums have been teasing the market for several years, but aren't buying. Three years ago they weren't buying because there were few if any systems made for them and those that were lacked necessary functionality. This is no longer true. Several of the new systems on the market this year do the job and do it well, and all are within what should be a reasonable price range. Archives and museums that aren't buying now either: 1) have unrealistic expectations, 2) aren't really interested, or 3) are scared.

I believe that fear is now a major factor. Now that archives and museums finally have choices, they are paralyzed by having not done adequate analyses of their own requirements to evaluate the options available. □

### TABLE OF CONTENTS

#### Articles:

Systems at the AAM by David Bearman	26
An Experiment with Research Data & Standards for MRR Reference Service by Thomas E. Brown	33
Copyright & Optical Disk by Brian Kahin	35

#### Regular Features:

Conferences	36
In-Box	38
Projects & Proposals	44
Software briefs	45
Standards	47
Technical Report Summary	48

---

## Software vendors at the AAM by David Bearman

The exhibit area at the American Association of Museums conference attracted a large number of vendors of automated systems. Each year the number of commercially available museum applications being shown increases, and the features of the systems on the market expand and mature. More than ever, buyers need a clear sense of their own requirements or they will find the array of systems competing for their dollars overwhelming.

Quite a number of the systems being shown at the AAM meeting this year were not displayed in previous years and have not previously been noted in this Newsletter. I am discussing them more extensively than usual because they are just being introduced. By doing so, I do not intend to slight the established vendors in the market whose products are being enhanced but do not represent brand new releases. Systems displayed at the AAM meeting, but not reviewed here, included enhanced products from Access International; Blackbaud Microsystems; Cuadro Associates (STAR); Erros Computing (STIPPLE); Master Software Corp. (MasterPiece and FundMaster); and QL Systems (QL Volunteer Management). With the exception of STIPPLE and QL, which was discussed in this Newsletter in conjunction with the AAM meeting last year, these products are detailed in the Directory of Software for Archives and Museums, Archival Informatics Technical Report vol.2 #1, 1988 [available for \$45 prepaid, \$50 billed from Archives & Museum Informatics].

The following descriptions are based upon demonstrations I viewed at the AAM meeting. Therefore, I can not say anything about how well these systems might operate with real databases and multiple users under real conditions. At the same time, I am not being non-judgemental. All the opinions expressed were formed as I observed the demonstrations and from vendor responses to questions I and others asked while I was present. Potential buyers are strongly advised to check these systems out for themselves.

**Delaware Computing Services** [5700 Kirkwood Highway, Suite 205, Wilmington DE 19808, 302-996-0490] introduced a collection management system, written in COBOL for DEC Micro-vax and HP 3000 systems. The system is offered as customized software, built around the model system installed for an "unnamed", nationally known art museum in Delaware, to the specifications of its registrar Karol Schmeigel. Although it is available as a turnkey system or software only at present, Delaware will consider operating a service bureau and then migrating users to their own systems when and if their size justifies it. Delaware Computing is a 1981 spin-off of Hewlett-Packard, based in Belgium with annual sales of about \$12M and 180 employees. Most of its income is derived from custom programming, but it also sells a variety of packages that it modifies for buyers in the fashion planned for the museum application. Two full time and one part time employee currently support the museum application, including the conversion of data from cataloging cards using the Palantir Compound Document Processor character reader.

Access to the system is security controlled by function (menu) and separate permissions can be granted to read, read/write/update, and read/write/update/delete. Usual systems management functions (including audit trails) and escapes into locally resident utilities are provided from the main menu for systems administrators, but the screens are hard-coded.

The main file being shown (subject to customization for future clients) consists of 30 data elements regarding objects and a textual (notes) field that can be as long as 99 screens and searchable by character string including wildcards. The data is physical description, source and association, and any field may be controlled by a value table, constructed either outside of data entry or through the data entry process. Measurement conversion is provided between English and metric linear measures only. Screen and field specific help is available and user modifiable.

Separate records hold location history (and scheduling forward locations), material characteristics, and donor information. While these records are not linked to a user defined or system supplied process, accessioning data entry

---

can return user defined prompts and require specific fields to be present.

Retrieval is a reporting function, rather than a separate online catalog interface, but it provides for some enhanced features like phonetic searches for donor names, a four level thesaurus with scope notes, and the facility to define a set by means of a prior search and execute a scheduled location change for such a set, so that exhibits can be planned or conservation actions scheduled involving a number of objects at a time with the system checking to assure their availability and make the necessary reservations. The search itself displays the data elements and allows the user to identify up to three values with EQ and NE operators for each element. A Boolean OR is imputed between elements. Only columnar reports are printed in response to queries, but the elements to be printed and their table location and length are user specified. Up to five levels of sort are permitted and the reports can be run in foreground or background with the query printing out as a separate page on printed reports. Search results may be saved in order to narrow a search, but not search logic.

On the whole it is a very usable basic system with features like whole/part and object/group relations and relations to reproductions, multiple former owners, makers or associated objects, but without many features of true collections management systems that would be of benefit beyond the registrar's office.

**Milestone 1, Inc.** [2625 N. Meridian St., Suite 43; Indianapolis, IN 46208; 317-926-4545] reintroduced the COBOL software for WANG VS systems previously marketed by Comsolutions, Inc. The partners, Burland Brown and Tim Prendergast, have enhanced the six modules (Fund Accounting, Membership & Development, Collection Information Management, School & Community Tour, Class & Theatre Registration, and Resource Center) and are offering them as software packages, with customization of screens but not of the underlying data definitions and architecture.

Fund Accounting supports 99 funds in its basic implementation, but has been modified for the Baltimore Aquarium to support up to 9999. The system supports traditional accounts payable,

receipts processing, monthly financial reporting and chart of accounts maintenance functions using separate vendor files and transaction (order/fulfillment) files. Report writing to a pseudo-spreadsheet view is very flexible but quite complex. Budgetting is supported by copying the year's budget and modifying it by applying increases or decreases in funds lines by percentages or \$ amounts. Some standard outputs, like 1099's, are defined.

Membership & Development is built around the person file (member, donor and the collection module person file share this). Person records include an alternate address, a spousal segment, and the usual interest code/membership category data. Codes are user defined, outside of the M&D data entry process, and exercise validation. Some aspects of the system are more rigid than we would expect from applications software. The module assumes batch processing of membership renewals and the issuance of cards and notices. Similarly, one can calculate payments due on multiyear gifts, but a fixed number of payments must be set up for all gifts rather than a number defined interactively for each gift.

Unfortunately I didn't have time to explore the events management modules, but I did check out the so-called Collection Information Management System which consists of 99 fixed length (40 character) repeating fields. The information retrieval technique, because Boolean query is not available, is to list one argument, obtain an interim result and refine the list. Accession numbers are 16 characters and no special sort logic is provided. No collections management functions or transaction records are provided. At best this is a very primitive facility.

**Performing Arts Technology** [2115 A Fourth St., Berkeley CA 94710; 415-548-6500] demonstrated their EXPLORER System, consisting of seven modules designed for events management, membership and development and accounting, that run in a very integrated fashion on the DEC Vax and Micro-vax. The first four modules (Walk-up Ticketing; Advanced Sales; Group Sales; School Group Management) provide powerful tools for active cultural repositories to manage admissions, ticketing for special events, scheduling of groups and educational programs for groups, and the management of cash registers

---

and billing systems associated with each of these functions.

A combination of ticket printing, visitor attendance/event preference reporting, and cash register management, makes the walk-up ticketing module especially attractive. Special discounts are calculated based on user defined types of visitors (seniors, children, membership categories) as well as on event combinations (multiple shows on the same day). Venues are booked as far in advance as desired and interactive ticketing assures that space available (including by seating section if desired) is always known to all cashiers. Each cash register is monitored separately and cash removals during the day by managers are security controlled. Users control refunds policies and the daily "set-up" of all activities (which can be copied from a template and modified according to the date). Well conceived variables in the daily set-up procedure allow managers to compare prior experience for any date (like previous year), the day of week, the weather, and by other variables that might affect attendance. A visitor survey data gathering window permits the cashier to administer a visitor questionnaire to track zip codes, how the visitor learned of the program, and any other information desired by management or the marketing department. The interface to membership enhances the value of reports on when each member visited and what special events they attended. In addition to gross attendance and income figures for the day, standard reports detail all categories of visitors by all events, distinguishing each show and each theatre or exhibit hall.

The advanced sales module adds name and address recording (for billing and mailing of tickets) to walk-up ticketing and can impose handling fees based on per person or per transaction costs. Group sales streamline ticketing by recording characteristics of the group (contacts, special needs, lunch plans etc.) and allowing for issuance of either separate tickets or a group voucher. It also permits the repository to store information about the composition of the group (minorities, handicaps etc.) to qualify for special funds. The school reservation system permits the institution to schedule groups, in advance, for educational

programs that occur only when so scheduled, and to locate such programs as are scheduled for one group, and reserve space for other groups to join them at the same time. It tracks programs that can be offered, interest levels, facilities in which the programs can take place, and instructors, and reports on each of these as well as issuing the necessary tickets or vouchers. The museum or archive can define a variety of educational offerings, identify instructors, and define appropriate spaces for conducting these. A reservation clerk can then schedule school groups to take advantage of these offerings. The education office can control its schedule by blocking out certain programs, instructors, or facilities for given dates, and monitor the actual schedule in advance with structured data about each group (grade, experience etc.) available to the instructor. The special requirements of a program (such as room set up requirements) can be taken into consideration in scheduling based on previously defined limitations, including such interaction effects as timing between programs required for set up.

The entire events management function, from walk-up ticketing through group reservations, is superbly crafted. One is struck by its elegant simplicity and power, and by the flexibility it provides for users to define their own programs without sacrificing functionality to support almost every conceivable reporting need or the demands of on-site ticketing.

I did not have time during the AAM meeting to examine Explorer's Membership & Development or Accounting modules except to see such interfaces to the ticketing function as on-site, walk-up membership cards, production of financial reports on attendance related income, and member interests tracking made possible by recording membership number when tickets are issued.

**Questor Systems Inc.** [1005 East Colorado Blvd., Pasadena, CA 91106; 818-356-0808], was showing two additions to its Argus system: a Slide Libraries module that makes reports, labels and backing cards for slides and provides for check-in/check-out, using bar-codes if desired, and a Site Files module that provides for recording of authority data about architectural and archeological sites with appropriate links to

---

the Argus lexicon controlled vocabulary. Questor also announced the opening of an Eastern regional office at P.O.Box 860, Upper Montclair, NJ 07043 (201)-746-7333.

**Stelwagon Management Systems Inc.** [114 Forest Avenue, Narbeth, PA 19072; 215-667-6776] debuted a museum system consisting of four modules: Accounting, Development, On-line Reservations, and Curatorial Management, running on an IBM System 36 in COBOL and RPG. Stelwagon will sell a turnkey system or software only, and expects to customize the product for each client. Company President, Jack Stelwagon, emphasizes integration and shared systems as the selling point and promises to bid aggressively while declining to suggest any specific prices. I had an opportunity to review the Development and Curatorial management modules only, so there was little opportunity to verify the integration of the parts around a shared database.

Development security control is exercised at a File, Menu and Field level, but from a user profile that must be set for each individual so field level security overheads may prove prohibitive. The donor record consists of 640 bytes in fixed length fields, stored as fixed length data, but a view of the donor record also incorporates all transactions associated with each donor from separate, short, transaction records. Codes may be recorded for mailings and campaigns in the transactions records, allowing reports on donations associated with specific drives, but there is no facility for setting up a campaign and managing it as a separate function. Gift receipts and donor history reports are the workhorses of the Development system. Other reports can be written using Executrive, a relatively straightforward query language, and Stelwagon has made an easy query screen with limited templates for one client.

Curatorial management provides an object record linked to an artist (creator) authority record and transaction records for exhibition and preservation histories. No part/whole relations are reflected in the object record. The catalogue entry field is tied only to a text editor and there is no query yet available from within the curatorial management module, although Executrive can be invoked from outside. Access points (subjects) are 24 characters in length

and string searches are planned. Exhibition, preservation and bibliographic references are all text fields without associated functionality. The system clearly has a long way to go before its Development and Curatorial Management (read catalog) functions begin to be competitive with others in the market, but the firm was launched with accounting and reservations as its expertise, so I may have missed its strong points.

**The Williamson Group** [129 Mount Auburn St., Cambridge, MA 02138; 800-992-6848] demonstrated the latest release of ARTIS - Collections Management which permits users to define a custom data entry screen consisting of fields they need in the order they need them rather than requiring them to use the nine screens and see all 45 data elements as in the past. This is a substantial improvement, however, these screens are limited to 15 lines each and must therefore be joined to build a full data entry screen. Also, each screen is named by a single character, severely limiting the number of processes that can be customized and the ease of use. Other enhancements include the ability to browse up a list without going out of the function, user defined one line displays for hits, user defined loan/exhibition data records, and true transaction based loan/exhibition history. Now users can set up the loans for an exhibit, copy the record and change all relevant data once when the exhibit travels to a new site, rather than having to update each loaned item with a new venue.

ARTIS still records all textual data in word processing documents pointed to by name from the field to which they apply. Not only can't users invoke the word processing document directly, but calling the document in WP puts the user outside the ARTIS system, from which he must log back on and into the record that was being updated. This limitation is so severe that the improvements demonstrated at this meeting, although substantial, seem insufficient.

**Vernon Systems Ltd.**, [P.O. 6909, Auckland, New Zealand; tel. 649-395-230] attended its first AAM meeting and gave me a full day demonstration of COLLECTION. Recognizing that a demonstration is always less exacting for a system than a hands-on review and that it

---

employed a single workstation on a small test data set provided by Vernon (and cannot therefore tell us about performance), I was immensely impressed. COLLECTION is, very simply, an elegantly designed system and its "procedural control" features provide it with the most powerful collections management capabilities of any system on the market. My only serious criticism of COLLECTION is that it tries to do too much, but except for the demand that places on sophisticated users, this criticism should be read as praise, since COLLECTION very largely succeeds in every case.

After a security controlled (encrypted) log-on, the main menu can identify up to 10 functions, beginning with Entry/Acquisition, and proceeding through Cataloging to Loan Processing, Conservation, Exhibits, Movement Control, General Inquiry and Reports, and finally the Systems Management and Administrative functions. A nicely conceived prompt identified the next level of the menu at the cursor position. Help is available both for each menu and every field, and may be fully tailored by the user. ESC consistently ends any function and returns the user to the main menu. Any process can be suspended at any time to PERFORM another activity and will be recommenced where it was left off after the detour is completed.

**ACQUISITION:** The Entry/Acquisition process displays the full power of the system. Any field may be set up to have a value table; the table may restrict entries to terms on the list, or it may be added to by entries in the field, and then made to control the new values after it is completed. Value tables may be simple term lists or more complex authority files with full name, display name, notes, broader/narrower relationships, use for and use instead indicators, related terms, and status of term. The hierarchical relations are shown when the data is entered, so that distinctions between Paris, Virginia, United States/North America and Paris/France/Europe are immediately evident. However, since the data is stored in only one place, no added space is occupied by the additional information, no matter how many levels of hierarchy (as in biological taxonomy when this can be many) are employed. Some of the implemented authority files are incredibly powerful; the person file

demonstrates complete descendant/ancestor linkages that display a full family tree, automatically making reciprocal updates (child records generated when mother is defined, for instance). In addition to such inherited data augmentation, data may be carried forward from prior entries, or defaulted to a template.

**CATALOGING:** The Cataloging function from the demonstration menu provides for cataloging of objects from 15 disciplines (archeology, costume, decorative art, ethnography, etc.) and 11 areas of interest (identification details, production details, field collection details, association, provenance etc.). These screens are generic templates intended to be customized for buyers of the system according to the nature of their collections. It is unlikely anyone would have all 26 sub-menus.

Cataloging itself can support very complex descriptions. In addition to the eleven detailed "areas of interest", a hierarchical structure can be used to define groups of objects, with components of each object and fragments of objects as well as pieces of components. The multiple independent parenthood of different parts allows properties of the whole group, the object, or a component to be inherited by all pieces. In the demonstration, I established a keyholder with keys as an object, each key and the keyholder as separate components, and assigned the object to the groups of my collection and of household objects, the keys to a group of metalwork and the holder to a group of leatherwork. The characteristics I described for each group (provenance for my collection, material for the metal and leather assignments) carried appropriately to the object, its parts and their fragments or pieces.

Vernon Systems has developed a rich variety of dating schemes permitting objects to be cataloged by a single date or a date range to any level of precision (year, month, day) as well as using the names of seasons, a number of years/months or days ago, decades, centuries, circa and such approximation qualifiers as pre-; post-; early-; late. All these dating methods are transformed within the system to a consistent internal format that can be searched (in general this is very elegant, but ranges are reflected as their mid-points which is not quite right).

---

In cataloging, groups of fields can be entered with alternative attributions, allowing the catalog to reflect different scholarly opinions. Linear and volumetric measures are converted from English to metric and back. Links can be made to other identification numbers, to reproductions (both images and models) and to association histories. And true transactional records are created for all exhibit, loan, conservation, and movement history and can be made for any other named action.

**PROCEDURAL CONTROL:** The exceptional innovation in this system is called Procedural Control. Procedural control enables any repository to define a process, for instance acquisition, by specifying its steps and the requirements of each step both in terms of data and reports. In practice this means that if a repository uses an existing form and procedure it may specify that certain data (the form) must be completed before it is forwarded to the acquisition committee for review, and that the committee wishes to see the data in a certain format. This report will then be completed by the system, awaiting action by the committee. When the result of the committee deliberations is entered, the next step is defined as notifying donors. Letters will be written merging data from the committee decision and acquisition information, including completed gift agreement forms with donor terms specified. Upon return of the forms, the repository could have further processes scheduled. Reports of costs in acquisition would be generated automatically to the accounting office and of insurance needs to the insurance office, etc. The entire process is completely at the discretion of each repository by completing a Procedure Control definition consisting of: 1) naming the procedure, 2) identifying next steps, 3) identifying steps to reverse to, 4) identifying data to be checked at the step and corrective actions if data is missing, and 5) identifying reports to be sent and writing any special calculations to be conducted.

Other aspects of the system also help support collections management. Locations are attached to time of day, not just date, so that many moves in the same day will record and display in order. Lists can be constructed by a search, and the entire list may then be attached to a process. As a

result, Exhibitions can be set up, cataloging data for objects in the exhibit imported, labels and catalogues defined, packing instructions attached, and the whole can be moved through multiple venues, shippers, insurers, etc. by rekeying only the elements of information that change. The entire history is retained. Separate steps in complex procedures, such as conservation actions, may be scheduled in advance to take advantage of equipment and people available, and when conducted can be recorded en masse.

**RETRIEVAL:** Finally, but not least, Collection supports full Boolean queries, building large expressions and executing them rapidly. The sets that are retrieved may be saved and researched, but at the moment the expression itself cannot be saved and the interim display of number of hits per term disappears from the screen too rapidly to be used. In image searching, Collection uses a special application of relationships to get associated biographical information for "portraits of X" while retrieving appropriate geographical data for "landscapes of Y". The linkages made in this kind of secondary search routine are user defined, and thus extensible.

The entire system comes complete with the systems management and administrative tools used in its own construction, giving the sophisticated user a powerful systems development environment restricted only in its ability to undo COLLECTION itself. Indeed, the very power and complexity of COLLECTION prevents me from recommending it without caveats. Vernon Systems is a small (3 person), young (2 yr old) firm in New Zealand (far away) that has yet to make a sale. The enthusiasm of the founders and their evident skills, are not a substitute for support, which will be hard to get unless they open a U.S. office. An IBM PC/386 running Revelation under DOS is not a large system, although it will support multiple users; only benchmarks with a real database will give us an idea of how large a database it will support. Revelation is a successful product, but a minor actor in the PC dbms market; its longevity is an open question. Which adds up to COLLECTION being a superb system for a sophisticated user able to assess the risks and provide support.

---

**Willoughby Associates Inc.** [266 Linden St., Winnetka, IL 60093; 312-501-4540] previewed their about to be released QUIXIS collections management system. This multi-user system, running on an HP3000, is to be Willoughby's top-of-the-line system. Like their PC product MIMSY, QUIXIS is licensed as a software package, without need for customization and with all its capabilities included for one price. Willoughby continues to be willing to bid the conversion of manual files into its system and encourages new clients to compare the economies of entering the data themselves with those of contracted input.

**ARCHITECTURE:** QUIXIS is built around an object file that supports object groups and parts. Separate records for associations, attributions and loan/exhibit transactions permit multiple independently occurrences of sets of reference data and object histories. Other fields are all variable in length and occurrence and may be indexed to separate or combined indexes without limit. Rights and reproduction data is driven from a separate menu to maintain its security and an even more limited menu (visiting scholar system) is provided for view only access to a portion of the database. Some features of the system reflect Willoughby's long experience in museums:

- \* the accession number is built from four data elements which can have independent sort rules so that all the different accessioning systems within one institution over many years can be accommodated and properly sorted;
- \* similar segmentation of locations into six fields with a note allows the most complex locations to be acted upon in groups
- \* attributions records for people and organizations are full reference databases complete with a generalized facility for recording events in the life of the individual or organization and relationships or roles.

Willoughby intends to distribute its system to art museums with a considerable number of these records already built. This will be a great advantage to them without binding them into any particular form of name since the user elects the form to be used and the alternates.

**RETRIEVAL:** QUIXIS will improve on MIMSY's QUE (Query, Update, Entry) screens, using the same concept to allow for integrated access and entry, but providing both interim hit counts and full search expression hit counts. Text fields may be searched by character string, with delimited keywords, or for adjacency, with or without stoplists. Wildcard searches are fully supported as is the full range of Boolean expressions. One particularly nice feature is the ability to define a display format for records retrieved in a set and to then view a different display for the same set, so that a query for records having some intellectual characteristic could display a brief name, title, subject, curatorial department entry. The user can then switch to the location view to see what is on exhibit and where.

**COLLECTIONS MANAGEMENT:** Although history tracking is supported, no generic actions capability is provided, only loans and exhibits. Nevertheless, QUIXIS is a true collections management system, not just a specialized information retrieval system. Routing for loans and exhibits, including specific flights, insurances, and packing materials (including cost data), is supported. Although an on-line tickler, person specific and sensitive to referred actions/authorization is still lacking at this stage of development, QUIXIS will notify users of today's scheduled collections management tasks at log-on and could notify them at log-off of incomplete tasks. Willoughby promises that the system will come complete with over 200 defined reports; those I have reviewed are exceptionally well conceived, reflecting the long experience of the principals in museums and the use of their prior systems, MILAM II and MIMSY. QUIXIS which will be available for installation in August has been worth waiting for.

□

---

## MACHINE-READABLE VIEWS

[The editor wishes to congratulate Tom Brown on his election as the incoming President of IASSIST. Correspondence regarding this column should be addressed to him at: Administrative Staff, Office of the National Archives, Washington DC 20408]

### An Experiment with Research Data

At the recent conference of the International Association for Social Science Information Service and Technology (IASSIST), a bold experiment in the preservation of empirical economic research data was unveiled. In empirical economics, a researcher acquires a primary data file or files from a data repository. Analysis of the information frequently involves recoding values, collapsing variables, inputting information, creating indexes, and altering values for certain variables. This information is called derivative data because it has been derived from the primary data sources. The researcher analyzes this derivative data to generate the final statistics that appear in publication. While access to the derived data is critical to validating the methodology, this derived data has not been generally available.

To remedy this, the Economics Program of the National Science Foundation has asked all grantees to indicate the public availability of derived data in a footnote each time they publish empirical research. In addition, journal editors are being encouraged to adopt editorial policies requiring placement of derived data in public depositories prior to publication. As a final inducement to researchers, the Inter-University Consortium for Political and Social Research (ICPSR) has agreed to accept deposit of derivative data. This effort to place derivative data in an archival repository raises some interesting issues. As would be expected, a variety of questions remain to be answered.

The first broad issue concerns archival appraisal. The predominant concern in the appraisal of automated records systems has traditionally been informational value. However, derivative data is being deposited with ICPSR in order to permit analysis of the methodology behind a published article. Thus it is being acquired for its evidential value or

evidence of the validity of the research process. If such acquisitions extends beyond the social sciences, to include also natural science data, the evidence becomes invaluable as a source for the history of science. This argument suggests the value of some of the derived data, but not all of it. Yet ICPSR is currently accepting, at least temporarily, all derived data sets submitted. Clearly, one complex appraisal question is how long to keep what data. One possible answer is to save data related to seminal publications. Unfortunately the characterization of a particular study as "seminal" is a judgment that evolves over time. Another possibility is to save only those derivative data sets with high reference demand. Conceivably, a repository could accept all derivative data sets; anything not used within a certain number of years would be removed from the holdings to allow for the acquisition of new data sets. While reference use frequently reflects the informational value of the records, archivists have traditionally been reluctant to base appraisal judgements on current reference trends.

The second broad issue concerns the extent of processing. ICPSR has developed four classes to describe the level of processing a primary data set receives. In addition, acquisitions are announced in the quarterly ICPSR Bulletin, and incorporated into the annual Guide to Resources and Services. Such descriptive efforts provide for intellectual control associated with processing an accession. For derivative data, ICPSR has added a fifth class (Class V). This class is for data sets or related documentation that the repository has decided not to validate, describe or otherwise process. At the present time, most of the derivative data is received on microcomputer diskettes containing both the data and the technical documentation. The depository is acquiring a collection of diskettes but has little or no knowledge of their technical specifications or intellectual content. Given this lack of processing, the derivative data seems to be at the periphery of ICPSR's archival collection.

Finally, a variety of questions concerning the researchers and reference service remain unresolved. Because of the lack of descriptive work, the only public indication that ICPSR has acquired the derivative data is a footnote in the

published article. This is clearly in line with the major objective of the program, namely to permit validation of published results, however researchers will have difficulty identifying and locating derivative data sets generated from the same primary data for comparative study. Because of the lack of processing and validation, researchers are also likely to encounter problems in using the data. Researchers seldom have the time or resources to organize the data and documentation to make it readily usable by others. Thus users of derivative data can expect data discrepancies, and incomplete or even erroneous documentation to interfere with their efforts to verify methodologies.

Certainly as the program unfolds, answers will be found to most of these unresolved issues. The answers may result in unforeseen innovations in data archives and data libraries. Derivative data from empirical economic research offers a new type of information: the archival management of it can easily involve the data repositories with other disciplines in the social and natural sciences that generate derivative data. The lack of processing might lead to a re-evaluation of the essential elements in acquisition of data files for archival custody. A similar re-evaluation could also focus on reference services. Thus the experiment may have benefits beyond the immediate goal of providing access to derivative economic data to validate published research results.

□ Thomas E. Brown □

### Standards for MRR Reference Service

For the past two decades, most repositories administering data files for research use have provided their users with technical documentation and a copy of the data on tape for users to manipulate and analyze on computers to which they have access. This approach to reference service arose when the only way to analyze such data was by batch processing on mainframes. Technology now offers us other options.

The goal of reference service for data in electronic form should be to provide no less than minimally acceptable reference service. In the partnership between the researcher and the archivist, it is the researcher who initially defines criteria for evaluating services. The

criteria that are typically cited are that reference services should provide information about or from the records: 1) where it is wanted, 2) when it is wanted, and 3) in the format in which it is wanted. Reference services could, therefore, be evaluated according to these criteria on a continuum from unacceptable to fully acceptable, as represented by the following matrix:

	WHERE	WHEN	HOW
Fully Acceptable			
Min. Acceptable			
Unacceptable			

In reference service for machine-readable records, each element has a range of options:

- WHERE:** at only one location  
 at multiple locations within the institution  
 at selected institutions  
 anywhere the researcher desires
- WHEN:** Received  
 only during certain times of day  
 twenty-four hours a day  
Filled  
 within seconds of receiving the request  
 same day service  
 within ten working days
- FORMAT** exact tape copy of the original record  
 tape copy to user specifications  
 new format as requested  
 on-line access

To the researcher, if a repository cannot provide data when, where, and how it is needed (one element is unacceptable), the whole is unacceptable. Thus the archives must meet at least the minimum acceptable standards in all elements if it is to provide minimally acceptable reference service. An archives may determine that, due to other considerations, it will not meet the researcher defined minimum for one or more elements. In this situation, I believe the archives should not provide reference, even if the other elements meet researcher defined minimally acceptable standards. Rather, it should propose an alternative reference service

---

that could satisfy all three user defined requirements.

Archivists should be aware that what users will consider unacceptable, minimally acceptable or fully acceptable is affected by the original records and the nature of research use. Thus appraisal and accessioning of machine-readable records must include establishing a reference strategy for the record.

□ Thomas E. Brown □

### Fair Use of Electronic Archives

Optical storage technologies offer remarkable opportunities for providing efficient access to text and images. Information stored in digital form, whether primary text or image codes, can be located by computers with extraordinary speed. By relieving scholars from tedium and repetition, the computer increases productivity and allows analysis much closer to the natural "speed of thought".

However, the process of recording information on optical disk necessarily involves copying, typically by scanning (in the case of text) or analog recording (in the case of images for videodisc). While it is widely assumed that copying onto optical media for archival purposes is fair use under Section 107 of the Copyright Act, there is no consensus on how far the fair use principle extends or should extend.

For example, once an archival optical disk is created, how can it be made available to the scholarly community. Can the disk be offered for circulation by mail? -- or is that a violation of the distribution right? An owned copy can normally be displayed publicly at the place where it is located (Section 109(d)), but does that principle apply to a copy created under fair use? [Consider the same issue in a different context: Is a time-shifted television recording permitted under the Betamax case law an "owned copy"? If so, it can be sold under Section 109(a)].

Can the archival copy ever be copied -- in whole or in part? Can a scholar download portions for his or her private use? Can the disk be mounted on a local area network or a bulletin board for access by more than one user? Can it be mounted on Bitnet or other wide area networks, so it can be made available to

researchers in other locations? What if the availability for remote use is publicized?

In short, the more accessible the archive is, the more likely it is that fair use principles cannot be stretched to cover the situation. Securing permission from copyright owners is a prohibitively burdensome task for archival reproduction in the first instance. The more publication paths are contemplated, and the more value to more users, the more impossible it becomes.

There is virtually no case law on the applicability of fair use to archival copying or public display. However, considerable funds can be expended to develop machine-readable archives before it becomes apparent that what makes these archives valuable may also render them useless. Unpublished material, even if it is of very narrow interest, presents special difficulties because the authors may not want even limited publication.

Fair use guidelines have been developed in several other areas. Guidelines for library photocopying have been incorporated into the Copyright Act as Section 108. Guidelines for photocopying and off-air taping for classroom use have been developed by representatives of the education community and copyright owners. The latter guidelines, although lacking the force of law, have been endorsed by the House Committee that deals with copyright.

Some initial assessment is called for to determine the scope of the problem. To what extent is the development of machine readable archives inhibited by the copyright problem? To what extent is the problem being addressed by existing archives. How is it being addressed? Is there sufficient consensus to do anything about the problem? What are possible solutions? The author invites readers to share their reactions and experiences with him.

□ Brian Kahin, J.D. □

[Research Affiliate, Research Program on Communications Policy, M.I.T., 254 Concord Ave, Cambridge, MA 02138]

---

## CONFERENCES

### AMERICAN SOCIETY OF INDEXERS, 20th Anniversary meeting, May 13

The ASI meeting was an intensive, day long session, held in conjunction with the On-Line meeting in New York. The opening talk was given by Hans Wellisch on the 'Literature of Indexing', followed by Dorothy Thomas on 'Book Indexing Principles and Standards'. These exposed a weak literature with inadequate and conflicting standards; Ms. Thomas argued that 2 of the 9 criteria advanced by the Chicago Manual of Style (the limit of 3 sub-headings and the deletion of "trivial" items from an index) are plain wrong and noted that the major texts in the field and the NISO standard Z.39.4 all use different terminology for the basic elements of an index. Needless to say, she called for action.

Linda Fetters updated her list of software currently available on PCs for indexing and suggested a typology consisting of two main categories: "Indexer Controlled" and "Software Controlled". Indexer controlled systems include standalone, imbedded and hypertext systems in which the indexer selects terms; those in which the software selects terms include text retrieval, automatic indexing and AI systems. She suggested that these systems can help do the clerical work of indexing, are raising the consciousness of publishers to the subtleties of good indexing, and provide a realm for thesaurus development.

The impact of new distribution media, both online data bases and CD-ROM was a topic addressed by many speakers. Barbara Preschel of the Public Affairs Information Service (PAIS) spoke to it directly. As an axiom, she stated that indexing methods must be dictated by the method of retrieval, therefore, the characteristics of print and CD-ROM were contrasted. Print assumes browsability; this could be provided in the CD or online, but rarely is. Print assumes one view; CD or online may support numerous views. In print indexes, there is a fixed list of headings whose depth and specificity is limited; in CD the depth and specificity and number of headings are open-ended. In print, the index is assumed to be in proximity to the text; online it

usually isn't but in CD's it can be even closer. Print indexes are always pre-coordinated; in online and CD environments they can be post coordinated. Print indexes employ natural language from the texts; in CD's, a controlled vocabulary may be required to provide consistency given the potential size of the database.

Everett Brenner, a pioneer of abstracting and indexing, reviewed the history of the use of computers in indexing from the 1950's to the present and concluded that controlled vocabularies and thesauri have failed as a substitute for natural language but so has full-text retrieval and called for research on how to make retrieval work, rather than making more thesauri. Like Bruce Croft who spoke on automatic indexing later in the program, Brenner saw promise in the automatic indexing techniques he and his colleagues have used at the API and proposed similar analysis of the queries provided by researchers.

James Anderson and Tefko Saracevic of Rutgers University presented research frameworks for analysis of retrieval in which it is evident that the number of variables are such that we can draw few large conclusions other than establishing, as Saracevic and his colleagues do in the *Journal of ASIS*, that it doesn't work well. Anderson did provide some guidelines on file organization and display, using library filing rules as an example of "pernicious display". Saracevic focussed on the results of their study which showed that when multiple searchers retrieve the same items they are more likely to be judged relevant by users, suggesting that if money is no object, we should hire second, third and fourth opinions in our database searches.

Ben Ami Lipetz rounded up the conference with a provocative assessment of the usefulness of indexes. He has concluded, from an analysis of book reviews, that indexes don't matter to even the most critically informed readers, and he suggested that back of the book indexes won't therefore matter to publishers until someone creates a product to cumulate published indexes, which serves the interests of publishers by selling sections of books (on-demand printing) to users who now search only the periodical literature. Perhaps the back of the book index

---

has a future as a free standing information product.

I came away impressed by how little explicit consideration indexers have given to implementation environments in the traditional focus on the terms, the structure of indexes and the techniques of index construction. Barbara Preschel, who came closest, ignored the fact that CD's carry their own retrieval software. Jim Anderson noted that a good system makes its rules apparent to users and that pernicious systems hide them, and Ev Brenner tried to turn attention to the way the index works in practice rather than the vocabularies, but the meeting as a whole never came to grips with it. Probably we assume that we all know how printed indexes work and forget that retrieval is system dependent. □

### **AMERICAN SOCIETY FOR INFORMATION SCIENCE, mid-year meeting, May 15-18**

The ASIS mid-year conference on Artificial Intelligence: Expert Systems and Other Applications demonstrated how routine AI has become in the past few years. Speaker after speaker presented real systems, smarter systems that used AI concepts, but, they stressed, these systems were just plain old programming. Each speaker seemed to be engaged in a systematic demystification of this delicious mystery, in part to explain why their application was not smarter and in part to indicate how AI can be appropriated by any system. As Mitchell Waldrop stated in his plenary address, the most important influences of AI will be hidden in programs that operate as intelligent secretaries, office managers, travel agents and the like.

I was unable to attend the entire meeting, but I left with strong impressions of a few "nice little ideas". Chuck Goldsteins' presentation of user-cordial (user specific) interfaces written with the Borland Turbo Prolog (\$160) tool set that imitated the Grateful Med system developed in expensive mainframe environments with many man-years of programming, not only suggested that such personalized systems can indeed be made and will indeed be with us soon, but underscored the democratization of AI. Goldstein

showed that if a handful of production rules are defined for each database and a separate set for each utility, users can exploit the primitives of a generalized environment to conduct searches on numerous systems and combine them usefully, while maintaining their own personal interaction with their local system.

Everett Brenner described the Iterative Improvement of the Machine Automated Indexing System implemented at the American Petroleum Institute since 1982. Initially implemented to assist indexers by proposing terms for their consideration, it has now displaced them. The system began with only 40% hits and an equal number of noise terms but has been progressively refined and now achieves over 70% good assignments with little noise. It is noteworthy, especially to those who equate term selection with professionalism, the PC AT based system contains painfully few rules.

Martin Dillon reported on work underway at OCLC to weigh title terms according to their scatter across LC classifications in order to provide more strategic retrieval when using keywords from titles. The results suggest that phrase extraction and matching against such a structure can vastly improve full text retrieval.

I chaired a session in which Toni Petersen related the development of the Art and Architecture Thesaurus to the problems of semantic analysis and Pat Molholt represented the statements that can be made using the controlled AAT vocabulary in natural language phrases. The concept that the multi-faceted AAT is a language and that its hierarchies are the grammar of a disciplinary discourse leads to exciting implementation ideas. Pat Molholt, for example, analyzed the types of relations other than "whole-part" or "is-a" that are represented in natural language statements mapped to the AAT and considered the opportunities we have to operate on these extended relationship links. □

Readers will note that the editor attends many conferences. He will be attending all or part of the conferences starred in the list on the next page, however, he welcomes any readers to propose themselves as reporters for any of these meetings and/or to suggest other meetings of interest to readership of this newsletter.

---

## CONFERENCE SCHEDULES

July 20-23, 1988; National Association of Government Archivists and Records Administrators, Annapolis, MD \*  
[Steve Cooper, Maryland State Archives, 350 Rowe Blvd., Annapolis MD 21401]

Sept. 14-17, American Association for State and Local History, Annual meeting, Rochester NY  
[AASLH, 172 Second St. Suite 102, Nashville, TN 37201]

September 21-24, 1988; International Conference on Terminology for Museums \*  
[sponsored jointly by the Museum Documentation Association and the Getty Art History Information Program, Cambridge, England; Contact: MDA, Building 0, 347 Cherry Hinton Rd., Cambridge CB1 4DH ENGLAND]

Sept. 29-Oct 2, Society of American Archivists, Annual Meeting, Atlanta \*  
[SAA, 600 Federal St., Suite 504, Chicago, IL 60605]

October 3-6, 1988, Association of Records Managers and Administrators, Baltimore  
[Contact: ARMA, 4200 Somerset, Suite 215, Prairie Village, KS 66208]

October 2-6, 1988; Library and Information Technology Association, Boston \*  
[LITA, 50 E. Huron St., Chicago, IL 60611]

October 4-7, 1988; International Council on Archives Meeting of Experts on Descriptive Standards, Ottawa, Canada \*

October 23-27; American Society for Information Science, Annual Conference, Atlanta GA \*  
[ASIS, 1424 16th St., NW, Suite 404, Washington, DC 20036]

October 26-28, Museum Computer Network Annual Meeting, Santa Monica CA \*  
[MCN Conference, 5600 Northumberland St., Pittsburgh, PA 15217]

## IN-BOX

### REPORTS

**Developing Descriptive Standards: A Call to Action**; Jean Dryden & Kent Haworth for the Planning Committee on Descriptive Standards, Bureau of Canadian Archivists. Occasional Paper #1, Ottawa, Bureau of Canadian Archivists, 1987

As the title implies, this report is a call to action, pointing to benefits to be gained from standards and suggesting how and why the Canadian professional community should proceed.

**Information Architecture Progress Report**; Rae Thompson, Smithsonian Institution, Office of Information Resource Management, May 1988 67p. plus lengthy appendixes.

Although this is an interim draft of a working document in an on-going project, this progress report on the information architecture of a complex of museums and on the process employed by users and systems analysts at the Smithsonian for defining an information architecture is exceptionally valuable for anyone involved in modelling information systems for cultural repositories. By definition, information architectures are not directly transferable to other institutions, and even the process of defining them will look different in other settings, but the progress report has much to teach nevertheless.

**Public Record Office Optical Disk Project: Interim Evaluation Report**, London, Public Record Office, May 1988

In March 1987 the Public Record Office of the U.K. launched a pilot project to explore the feasibility of storing its data archive on optical disk due to the costs of controlled environments and recopying of magnetic tapes. The objectives of the pilot are to: 1) test the suitability of optical disk as an archival medium; 2) explore the implication of public access to computer records; 3) establish a standard of documentation for computer records transferred to the PRO; 4) assess the costs of establishing and running an optical disk archive, and 5) share their findings. The project was confined to traditionally

---

structured digital data, and did not include free text, sound or image.

The interim report for November 1987 to April 1988 represents the earliest results of the project (following training). The PRO is encouraged by costs and ease of use. Still to be conducted are tests of the medium longevity, including one using optical disks that have been stored both vertically and horizontally on regular shelves in the paper storage stacks since December 1987. The only serious issues encountered to date are that the PRO was initially getting only 65% of the promised 1GB storage per disk. Their vendor, Data General, has since rectified the problems and the PRO is getting 95% of the promised utilization. Further reports can be expected.

**Reaching Out: A Review of Education & Public Programs in New York City Cultural Institutions**, Barbara Kalvert, Deputy Director for Cultural Institutions, February 1988, New York, New York City Dept. of Cultural Institutions, 1988, 64pp.

Handsomely illustrated, well presented, useful recommendations for coordination and strengthening of the diverse cultural institutions in a city. A model of how community planning efforts can be presented to win broad support.

**Technology & U.S. Government Information Policies: Catalysts for New Partnerships**, Association of Research Libraries, Washington DC, ARL, 1988; \$5.00 argues that technology has left policy behind and that this has severe implications for cultural repositories that have traditionally provided access to government information, including libraries & archives.

**Videodiscs in Museums: A Project and Resource Directory**, Roberta Binder, Falls Church VA, Future Systems Inc., January 1988 25p. plus appendixes listing projects & resources; \$49

This spiral bound report extends the directory printed in the April 1987 issue of Videodisc Monitor, providing longer entries for each project reported and a short introduction to the technology and its uses.

## JOURNAL ARTICLES & BOOKS

Abell-Seddon, Brian; **Museum Catalogues: A Foundation for Computer Processing**, London, Clive Bingley, 1988, 224p. (ISBN 0-85157-429-7)

Brian Abell-Seddon presents a coherent case for building museum systems on museum data in this important book. The proposed Reference Framework for Organizing Records in Museums (REFORM) is a program for the incremental improvement and migration of manual catalogues to automated forms which turns out to be the architecture of his commercially available museum information system, REFORMATION; but his analytic perspective has validity beyond any specific implementation. The book as a whole makes an eloquent case for system analysis and data modelling, but without ever employing any computer jargon. In its place, Abell-Seddon liberally illustrates his points using actual data from a wide variety of museums, demonstrating concretely how museums can make their information more consistent and useful by developing and applying conventions and terminology controls. He also clearly describes how they can exploit the information they already have for scholarship, collections management, and administration. Museum curators and archivists will find his discussions and illustrations of authority control, vocabulary development and description convention germane, even when the examples don't directly apply to their situation.

Bakewell, Elizabeth, William O. Beeman & Carol McMichael Reese; **Object, Image, Inquiry: The Art Historian at Work**, Report on a Collaborative Study by the Getty Art History Information Program and the Institute for Research in Information & Scholarship, Brown University; Marilyn Schmitt, General Editor, Santa Monica, Getty Trust, 1988 (ISBN 0-89236-135-2)

This long awaited report on how art historians work and what needs they have for images, doesn't answer either question. We learn a bit about how some art historians pursue research, but without much of a conceptual framework with which to make sense of the data. And we are

---

left without guidance about how image retrieval and manipulation systems might be designed to support art historical use of images.

Blackaby, James, Patricia Greeno, and the Nomenclature Committee; **The Revised Nomenclature for Museum Cataloging: A Revised and Expanded Version of Robert G. Chenhall's System for Classifying Man-Made Objects**, Nashville, TN, AASLH Press, 1988

The revised Nomenclature consists of 34 pages of introductory material on what it is, how to use it and the sources for each section, along with a 228 page classified list, a 242 page alphabetical list, and a nine page bibliography. It is a hefty book in part due to use of typewritten lists.

The classified listing consists of ten categories with up to two levels of inverted terms beneath them. Scope notes are only provided for the two highest categories. Terms are "generic object terms used for indexing purposes" (p.11-3), not common names or manufacturer names for objects. The authors warn that "users of Nomenclature will encounter some difficulties. Despite continuing efforts to add to the list, not all objects have object terms. Some objects are placed ambiguously in the system. Some objects in museum collections are parts of things or sets of things. Some things are difficult to classify without referring to their varietal names. Some things have been modified from their original purposes".

What they do not say is that the classification continues to have greater ideosyncrasy because it has not been constructed according to strict rules of thesaurus construction. Object terms are usually binomial. The term itself is a noun, but it may be modified by a range of terms that obey no obvious grammatical rules. MACHINE BATCHER/JOGGER; MACHINE, BEATING; MACHINE-BRICK; all follow closely in the alphabetical listing! Nor does the same qualifier necessarily operate the same way: BOX, NEEDLEWORK and CASE, NEEDLEWORK are "Textileworking Tools and Equipment" under (category 4), while PICTURE, NEEDLEWORK appears under "Communication Artifacts: Art" (category 8) and needlework itself, though a common usage, doesn't even get a cross reference. BOX, BREAD holds bread and BOX, CIGARETTE

holds cigarettes but what of BOX, SHADOW (no relation to shadow boxing)? BOX, LUNCH is the container, not a box lunch which isn't here and BOX, WOOD is a "Temperature Control Device" for which you should use WOODBIN! How is it that NET, FLY is an animal husbandry tool while NET, MOSQUITO is bedding and general NET, PROTECTIVE is clothing -- headwear?

All of which is just to say that this system can be used to classify objects, but it is not as accessible an access language as it should be. Hopefully this will be addressed by: 1) adding varietal and manufacturers terms as lead-in vocabulary, 2) providing the file in a machine readable, searchable form, and 3) illustrating the terms with scope notes and pictures. This should engage the history museum community for the next decade or so. Therefore, the AASLH is selling Nomenclature as a subscription service including quarterly updates and a technical report on converting from "Old Nomenclature" to "The Revised Nomenclature".

Cook, Michael ed., **Approaches to Problems in Records Management II: Computer Generated Records**, Society of Archivists, Computer Applications Committee/Records Management Group, 1987 60p.

These proceedings of a conference held in Liverpool in September 1986 include papers on various aspects of the identification, appraisal, accessioning, and use of machine readable and computer generated records including an initial report of a CCTA experiment with optical disks, an outline of the PRO optical disk study and some assessment of the legal situation of computer generated evidence in British case law.

In a challenging introduction, S.C. Newton, the conference chairman, observes that "perhaps the real issue is not computer generated records, but the records which are not generated by the computer, i.e. those which remain forever in the heart of the machine." He expands upon this subsequently, noting that "if someone gives you magnetic data filed without a DBMS or a query language, it is the equivalent of giving you a box of words - how do you know what sentences they are supposed to make?; indeed, how do you know what sentences they did make when in operational use?"

---

Dearstyne, Bruce; **The Management of Local Government Records: A Guide for Local Officials**, Nashville TN, AASLH, 1988, 146p. (ISBN 0-910050-91-0)

Buy your local government official a copy of this book! Bruce Dearstyne addresses him or her personally from the first page in this readable, practical, and pithy guide. Like the rest of the book, the 10% devoted to automated systems for managing records and the management of automated records can be counted on to distill the best advice from a range of sources. A nice overview indeed.

Saracevic, Tefko et.al.; "A Study of Information Seeking Behavior", Journal of the American Society for Information Science, vol.39, 1988 p.161-216

This three part article on an extensive study of what makes for effective information retrieval is the most important basic research to be published in this critical area for years and should be required reading for anyone involved in providing information services.

Stone, Nidia; "Infotind: A practical tool for managing information", Information Management Review, v.3 #4, 1988, p.39-46

Discusses an IRM directory developed in New Jersey to identify records systems in state government. The 4100 entries include disposition requirements.

Weber, Lisa, "Educating Archivists for Automation", Library Trends, Winter 1988, p.501-518

The former automation officer of the SAA provides a comprehensive overview of what has been going on in archival automation for the past several years and how it is affecting archival education and the skills required to be a professional archivist.

## NEWSLETTERS

CAD Information CDA (ISSN 0709-4620) published bi-annually for the Commission on Archival Development by the National Archives of Canada, Bernard Weilbrenner editor, 395 Wellington St., Ottawa, K1A 0N3, Canada; free.

Government Publications News (Bernan Associates, 4611 Assembly Drive, Lanham, MD 20706-4391) reports that the long-awaited OTA study of electronic publishing and dissemination of government documents will be distributed through GPO in late summer or early fall.

Heritage Education Quarterly, The Preservation Library & Resource Center, 429 South Main St., Madison, GA 30650 \$12 p.a.

This one year old newsletter on cultural outreach programming is a boon to museum and archives educators.

Humanistiske Data, vo.1.1, 1988 p.4-47, contains the proceedings of a conference on optical media including short reports, some in English and (unfortunately for me) some in Norwegian, which provide a tour of numerous exciting projects in the humanities using videodisc and optical digital disks.

Library Systems Newsletter, vol. 8, #3 &4, contains the Annual Survey of Automated Library System Vendors, in two parts: turnkey vendors and software only vendors. Updates developments in each system over the past year.

Museum Association Security Committee Newsletter, vol.7 #1, Spring 1988 [available to members of the Committee; AAM dues plus \$10.] reports that the National Fire Protection Association's **Practices** #910 "Fire Protection in Museums and Museum Collections" and #911 "Fire Protection in Libraries and Archives", will be revised next year. Contact Jack Watts, Fire Safety Institute in Manchester VT, 802-462-2663 if you want to be involved or can contribute ideas.

Museum Management Institute Newsletter, vol.1 #2 (American Federation of Arts, 270 Sutter St., San Francisco, CA 94108) contains a script for some role playing that has great possibilities for developing some better approaches to personnel management for any administrator. Strongly recommended - but follow directions and don't read both scripts!

---

Optical Memory News, Issue 53, November 1987, published the text of the Texas law on admissibility of Optical data in courts of law. In a recent issue of INFORM (April 1988 p.43-44) the law is reviewed by Douglas Allen, CRM, and found to be seriously flawed. Others promoting similar legislation should note his critique.

Publishing and New Media Technology Newsletter (ISSN: 0885-6214) bi-monthly from Pergamon Press, £130 p.a.

This expensive, stapled, 12 page, newsletter of "company news, market trends, statistics, product innovations, technological developments and media competition" consists almost exclusively of press releases by "new media" firms, purchase offers and conference schedules.

Registrar, vol.5 #2 (AAM Registrar Committee) is devoted to storage environments & planning requirements. A useful treatment.

## EPHEMERA

The Association of Research Libraries [1527 New Hampshire Ave., NW, Washington Dc 20036] will issue the report of its Task Force on Government Information in Electronic Format this summer. Drafts of the report, Technology & U.S. Government Information Policies: Catalysts for New Partnership, have been circulating since October 1987 and have now been endorsed by members. The report defines a framework for government policy and private sector roles that contrasts sharply with that adopted by the Reagan administration, and it develops the implications of these policies with respect to changing technologies of information creation and dissemination in ways that are extremely important to archivists and librarians. The report also presents a taxonomy to categorize the types of governmental information in electronic formats that should advance our discussion of their handling.

Mellon Bank Corporation, Discover Total Resources: A Guide for Nonprofits, free from Community Affairs Division, Mellon Bank

Corporation, One Mellon Bank Center, Pittsburgh, PA 15258 or call Sylvia Clark, (412) 234-3275.

A self-study checklist to assure that a non-profit is doing everything possible to acquire and use resources. Full of good ideas.

University of Toronto Archives, Principles and Guidelines for the Description of Records in the University of Toronto Archives, January 1988

Kent Haworth, Archivist of the University of Toronto, and a leading member of the Canadian working groups on descriptive standards, has written an internal manual for description, based on Steve Henson's manual for cataloging, Archives, Personal Papers & Manuscripts, and the rules for description of graphics records proposed by Elizabeth Betz Parker. Haworth, and Heather MacNeil, who shares responsibility for the U.T. guidelines, have a sophisticated appreciation of the informative value of provenance and the importance of integrating archival descriptive tools that makes this an exceptionally valuable manual. Hopefully these views will inform the emerging Canadian standards, providing a clear path for the acceptance of mutual standards throughout North America.

University of Liverpool, Archival Description Project, Archival Description: A MAD User Guide. Test Draft 1.1, 25 September 1987

Margaret Proctor and Michael Cook offer us the first draft of a users guide to the Manual for Archival Description. Intended as a standard for the U.K., MAD takes "levels" of records as a central principle. The users manual is, however, more philosophical than explicit about what information should be incorporated into description of each level. Although examples are provided, they have few discrete and identifiable data elements, and consist largely of relatively loosely structured text. The User Guide is reminiscent of David Gracy's guidelines for construction of registers and inventories, except for advancing what seems like a somewhat rigid numbering scheme for levels. On balance, the manual does not address questions of source of information, access points and vocabulary control that have come to the fore in the North American dialog about description.

## CD's: a publication note

It is now evident that CD's are coming to archives, not just as media generated by their parent institutions, but as reference resources for archivists themselves.

The Library of Congress Cataloging Distribution Service is selling Library of Congress Subject Headings on CD-ROM (complete through March 1988) for \$300, including the original, two fully cumulated updates and manual. The CD will play on any standard CD drive with a PC using DOS 3.1+. Library of Congress Name Authorities will follow soon.

Chadwyck Healey Inc. intends to market the indexes to the National Inventory of Documentary Sources on CD. No publication data has yet been set.

AIRS Inc. has announced the publication of the complete NUCMC Index (1958-1988) on CD-ROM. To be available in September for \$595 the CD is being offered at a pre-publication price of \$495 by July 15. Annual updates will be offered for \$225 to original purchasers. The product will feature Boolean searching, saved queries, hypertext browsing, and the printing of reports to disc or printer.

A large number of archives are associated with universities and state libraries that are engaged in major retrospective conversion/CD-ROM publication projects. Anyone in the position would be well advised to read a consistently useful series of articles on CD-ROM Public Access Catalogs (PACs) that has been appearing in *Library Hi Tech*. The first article, by Linda Helgerson, was entitled "Acquiring a CD-ROM Public Access Catalog System" (vol. 5, #3, p.49-75). The second article, by Linda Bills and Linda Helgerson, on "Database Creation and Maintenance" examined how the different vendors prepare the databases for CD recording (vol.6 #1, p.67-86. The third article, also by Bills and Helgerson, is entitled "User Interfaces for CD-ROM PACs". It is a superb critical review and comparison of the commercially available systems, important not only to potential CD-ROM users, but to anyone involved in user interface design or in the use of systems (vol.6 #2, p.73-115). Read these reviews.

**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 not otherwise attributed. Subscription to the **Archival Informatics Newsletter** (ISSN 0892-2179) is available for \$24.00 per year, pre-paid, to U.S. addresses; \$30.00 per year, pre-paid to foreign addresses; and \$40.00 per year billed, worldwide.

A subscription to both the Newsletter and its companion quarterly publication, **Archival Informatics Technical Report** (ISSN 0894-0266) is available for \$160.00 p.a. in the U.S.; \$180.00 abroad; and will be billed at no additional charge.

Individual technical reports are available at \$45 each, prepaid; \$50 billed. Titles currently available include:

- Optical Media for Archives & Museums
- Collecting Software: A New Challenge for Archives and Museums
- Functional Requirements of Collections Management Systems
- Acquiring & Implementing Automated Systems
- Directory of Software for Archives & Museums

Forthcoming issues include:

- Smithsonian Seminar on Authority Control in Archives (August 1988)
- Archival Appraisal of Online Information Systems (Nov.88)
- Functional Requirements for Exhibits Management (Jan 89)

---

## PROJECTS & PROPOSALS

### **Fiorello H. LaGuardia Photo Project**

A recent blitz of publicity unleashed by the Archives of the City University of New York regarding its computer system for small archives or museum photo collections led me to ask them for some documentation. Dr. Richard K. Lieberman, project director, provided me with a copy of the final report he submitted to the NEH, documenting the system that was developed to catalog 2800 photographs of New York City during Mayor LaGuardia's administration. His press release states that "in the absence of an accepted standard for the computer cataloging of photo collections under 50,000 items, the Archives has designed its own information retrieval system. The system allows researchers to seek items by site, persons in, or forty-five topics. . . ."

The system is an extremely basic, single file, catalog. Each record in the system consists of eleven fields: Caption, year, id #, subject categories (from the 45 terms available), subject headings (from eight broad categories), donor, site, negative flag, copyright holder, participants and location. Users, we are told, may review the subject categories and subject headings, as well as all sites and participants from printed lists prior to using the system. Searches, based on character string occurrence only, produce either a count or a serial display but cannot provide both.

Apparently the system lacks value table or authority validation, provides for repeating occurrences of fields but not repeating field groups, and requires all data to be keyed at the item level even when the same donor record in a multi-record database could have served as the source for a number of item records. The examples of records given in the report itself illustrate a variety of undesirable practices such as citing LaGuardia himself when he appears in a photograph as FHL and employing an exceptionally idiosyncratic list of 45 "categories" that include terms such as Agriculture, Candid, Funeral, Leisure, Out-Of-State, and UNRRA! [Archives, LaGuardia Community College, CUNY, 31-10 Thomson Ave., Long Is. City, NY 11101]

### **Further on Form from North Carolina**

Barbara Cain, Computing Coordinator of the North Carolina State Archives, shared with me a memo she drafted to the internal computer committee at NCSA in which she makes a number of observations that will be useful to other archivists. I am guilty of extracting a few passages out of context from her lengthy report, so any lack of coherence should be laid at my door:

"Some of our difficulties with form in the past have resulted from confusion between form and format. (This mistake was made in the 1985 RILIN list of Form Terms we studied, compiled by Hickerson and Engst). A map is a map whether it is an original, a manuscript copy, a xerox, a photostat, a photograph or on microfilm. And any one looking for maps will want the search to be across the board. It is important, therefore, that MAP be used with all maps, in whatever format they appear. The same would be true of diaries, whether the format be a bound volume, a looseleaf notebook, typed transcriptions, xerox, or microfilm. The information on format, then, goes in the DESCRIPTORS field in FAIDS. It is important to remember that the physical format we have in hand (microfilm, volume, photographic negative, electrostatic copy, etc.) will be found in DESCRIPTORS, not FORM. . . ."

Elsewhere in her memo, Barbara states: "Our use or non-use of subdivisions of forms has been very subjective so far, depending on the importance each archivist attaches to the more specific variety of form, and depending on the context in which it is found. We now have in FAIDS both Certificates and Election Certificates; Resolutions and (in the General Assembly records) Simple Resolutions and Joint Resolutions; Photographs and Aerial Photographs; Deeds and Sheriff's Deeds; Maps and Profile Maps. We have Bonds now and may wish to add Marriage Bonds. Otherwise, I would think we would enter Bonds plus Apprentices, Ordinaries, Officials, etc. as subjects. Do we add Death Certificates, Birth Certificates, and Marriage Certificates as separate forms? A case could be made either way. . . . It is important to remember that the FORM is not designed to define everything we have, but to assist with retrieval. For example, it may not be necessary, and may

---

even be disruptive, to distinguish joint resolutions from simple resolutions, since one would normally be looking for a resolution relating to a certain person or subject, not specifically for a joint resolution. . .".

Now that Barbara has launched this discussion inside North Carolina, I hope that others will become involved, either directly with her or through these pages.

### Function & the RLIN 7 States Project

The first grant funded project involving seven state archives in cooperative exploration of the use of the Research Libraries Information Network (RLIN) has been completed. A proposal to expand the community of users to 17 states and the National Archives and Records Administration is pending before the NHPRC and, if funded, could be underway by December 1.

A particularly detailed evaluation of the use of functional access in the first phase of the project is being conducted by Kathleen Roe (New York State Archives) and Alden Monroe (Alabama Dept. of Archives & History) this summer. Roe and Monroe were awarded NEH/Mellon Fellowships at the Bentley Historical Library to assess by case study, and test searches the utility of the functions vocabulary and prepare guidelines based on this analysis that will improve its effectiveness.

### Clearinghouse on Art Documentation Proposes to go online on RLIN

Pat Barnett, Director of the Clearinghouse on Art Documentation and Computerization [Thomas J. Watson Library, Metropolitan Museum of Art, Fifth Ave. @ 82nd. St., New York, NY 10028] has submitted a proposal to the Council on Library Resources to enable the Clearinghouse, which has previously been available only by mail and phone, to make its resources known through the Research Libraries Information Network.

The Clearinghouse is also involved in a cooperative project with the Art & Architecture Thesaurus to maintain vocabularies relating to computerized documentation.

## SOFTWARE BRIEFS

### Crowninshield Software Incorporated

[98 Crowninshield Rd., Brookline, MA 02146; 617-232-1488] is distributing demonstration diskettes of its optical publishing software system, **MediaBase**. MediaBase allows users to build local multi-media (including motion images and sound) databases, index them and output them in High Sierra format on 9 track tape for direct transfer to CD-ROM. The demonstration diskette doesn't show much, but the literature looks promising and I will be pursue it.

**Cuadra Associates** [11835 W. Olympic Blvd., Suite 305, Los Angeles, CA 90064; 213-478-0066] celebrated its 10th anniversary this Spring. STAR NEWS, vol.1 #1 made its debut with an announcement of its **STAR®/PC** product which supports six simultaneous users. Two new users groups are organizing - in Los Angeles and Washington DC.

**Electronic Text Corporation** [5600 North University Avenue, Provo, UT 84604], publishers of CD-ROM versions of Shakespeare, the Bible, Goethe, and such American classics as Emerson, Twain, Melville, London, and James, is publicizing its offerings in News etc. Vol.1 #1 reports on the **WordCruncher** software it uses, in conjunction with WordPerfect™.

**Master Software Corporation** [8604 Allisonville Rd. Suite 309, Indianapolis, IN 46250; 317-842-7020] reports in Master Files (vol.3#1) on new features of **Fund-Master** 6.0, due for release at the end of the summer. Included are spousal donation accounts, the maintenance of "letter" histories to warn you before you send the same form letter twice to the same donor, and automatic calculation of membership categories based on donation level are among the new twists. The issue also describes **MasterPiece**, their new collections management system, a plan for regional users group meetings, and the recently released **Fund-Master** 5.0 which includes numerous desktop publishing features.

---

**Michigan Microcomputer Cataloging System** [University of Michigan Software, Intellectual Properties Office, 225 W. Engineering Bldg., Ann Arbor, MI 48109-1092; 313-936-0435] is being advertised as a slide cataloging system for PC/XT/AT systems running DOS 2.0+ using compiled FoxBASE+. For \$1000 per site the system provides 16 fields up to 110 characters in length and stores variable length records and prints slide labels or shelf card lists.

**Time-Space Systems Inc.** is working with the Metropolitan Toronto Archives and Records Centre on an on-line catalog system with a unique cartographic/chronological interface to be navigated by mouse. Company President Richard Hill sees the interface as being uniquely able to provide naive users with immediate feedback about any aspect of Toronto area history as well as linking to machine readable data concerning the city and its demography [for more information call 416-699-2003].

**Image OnLine Inc.** [P.O. Box 1124, Topanga, CA 90290; 213-455-3883] announces the **Electronic Library**, a merger of imaging systems and library bibliographic systems to provide libraries of print, audio, and still images, on-line. Image OnLine is a new firm, but its principals have many years of experience in libraries and solid experience serving as an integrator.

**Blackbaud MicroSystems Inc.** [160 East Main St., Huntington, NY 11743; 516-385-1420] has introduced **Planned Giving II**, an upgraded version of their earlier planned giving software package that features comparison, on one screen, of remainder values for up to six gifts or variations on a gift, calculations of cash flows for various types of Unitrusts, data for determination of the gift and estate tax implication of Inter-Vivos Charitable Lead Trusts, and calculations to determine if gifts will trigger Alternative Minimum Tax and the maximum gift that will not trigger AMT. Numerous other features make the package a 'must-see' for planned giving professionals.

A membership addition to the **Raiser's Edge** Membership and Development software was also being shown at the AAM. It is described in a large packet of documentation and reports available from Blackbaud. Unfortunately, I didn't have time to examine the system personally.

## NEW SOFTWARE DIRECTORIES

American Society of Indexers; **A Guide to Indexing Software**, 2nd edition, Linda Fetters, compiler, 10p. from ASI, \$7.50 members, \$10.00 non-members

Listing of nine indexing packages with a structured one page description of each. The editorial judgements expressed in this "compilation" are relatively pronounced and may be useful guidance.

[Ed. note: Among the IBM PC indexing packages not listed in the ASI directory are: **IN>SORT** [Kensa Software, P.O. Box 49, Cathedral Station, New York, NY 10025 - \$79.95 post paid]; **WINDEX** [Watch City Software, 24 Harris St., Waltham, MA 02154 - \$109.95]; **MACREX** [Bayside Indexing Service, 265 Arlington Ave, Kensington, CA 94707, \$325].

Archives & Museum Informatics [5600 Northumberland St., Pittsburgh, PA 15217] has published its **Directory of Software for Archives & Museums**, 1st edition, 1988. The directory contains two page structured data comparing 35 packages along more than 100 dimensions and twenty pages of tables comparing the systems according to an additional 300 characteristics. Indexes by system name, hardware, operating system, vendor, application and utilities are provided for easy reference. The report is available for \$45. prepaid, \$50 billed, from Archives & Museum Informatics.

National Archives & Records Administration, Records Administration Information Center is distributing **Records Management Software Packages**, 1987". 10p., free. This latest NARA-RAIC Information Package list 9 software packages for records management along with a one page prose description of each. A good place to start for addresses.

---

## STANDARDS

### Framework for International Exchange of Museum Information

At the AAM Meeting, Peter Homulos, Director of the Canadian Heritage Information Network (CHIN) and Acting President of the Museum Computer Network (MCN) presented a paper on International Exchanges of Museum Information that has significant implications for museum data standardization efforts.

He laid out a framework, applicable to any information exchange, that builds from **agreements to standards to networks.**

Recapitulating the steps taken by CHIN over the past several years, Homulos detailed the terms of an International Agreement on Museum Information Exchange, presented by the Canadian government as a draft international convention. The Draft convention addresses ethical, legal, economic, and technical issues involved in such exchanges. The technical issues were, by far, the easiest to resolve; they boil down to a standard format for information interchange. Ethical issues such as agreements to reveal the problems known to exist in the data, the uses of data by non-signatories, exchange of research data, and non-compliance are more complex. Legal issues, such as national regulations, data ownership, duration and limitations on data use, and copyright involve national sovereignty and can only be accommodated within the framework of a general agreement when coupled with a mechanism for actual data interchange that presumes all such interchanges reflect specific bilateral contracts between institutions involved.

Moving to standards, Homulos noted that the activity taking place at a national level, such as has been sponsored by the MDA in England, CHIN in Canada, the MCN and AASLH in the United States, and MARDOC in Holland, is now being dealt with at the international level by the Committee on Documentation (CIDOC) of the International Council of Museums (ICOM) through its database survey, documentation centre, terminology and terminology control working groups.

Finally, Homulos noted, if we have agreements and standards, we can exchange data, and he pointed to the success of CIN, the Conservation Information Network, mounted on computers owned by CHIN and accessed by almost 150 member institutions from around the world. The CIN, which is less than two years old, was able to utilize CHIN as a utility because it built upon existing standards and existing physical connection methods. Observing the success of both the Conservation Information Network and CHIN itself in the years since its establishment in 1972 as the National Inventory Project, Homulos suggested that further exchanges of information between museums were only awaiting agreement from a few potential participants, and they too could take off on the foundation of an existing network and standards.

### Photographic Thesaurus Progress

Diane Vogt O'Connor, author of the draft photographic thesaurus circulated last spring by the Smithsonian Archives Photo Survey Project (see vol.1 # 2), reports that she has joined forces with Richard Pearce-Moses, Photo Curator at the University of Arizona, to extend the thesaurus and convert it to a strict ANSI thesaurus structure. They will be adding more tradenames, vernacular terms, generic terms and analytic terms for processes and structuring the data. Users of the draft who have comments, suggestions, or additions are welcome to contribute them. [Smithsonian Archives, A&I 2135, Washington DC 20560]

### Art & Architecture Thesaurus

The Art and Architecture Thesaurus has issued User Update #14 (April 1988).

The AAT is making final arrangements to distribute its terminology in machine-readable form and to give software vendors an opportunity to incorporate the vocabularies into their systems. Project Director Toni Petersen reports that this distribution will take place as soon as reasonable terms can be worked out to protect the integrity of the AAT.

---

# SMITHSONIAN SEMINAR ON AUTHORITY CONTROL IN ARCHIVES

**Edited by Avra Michelson**

**Archival Informatics Technical Report**

vol.2, #1, Spring 1988

In October 1987, the Smithsonian Institution Bibliographic Information System Archives User Group held a one day seminar on Authority Control to which they invited archivists and museum curators from throughout the Smithsonian. Invited speakers included:

Jackie Dooley, Special Collections Librarian, University of California, San Diego  
Tom Garnett, Systems Administrator Smithsonian Institution Libraries  
Marion Matters, Senior Records Analyst, Minnesota Historical Society  
Lisa Weber, Automation Program Officer, Society of American Archivists  
Richard Szary, Systems Administrator, OIRM, Smithsonian Institution  
David Bearman, Consultant, Archives & Museum Informatics

Papers introduced the concept of authority control, discussed its application in library systems, examined its use in an historical society both in manual and automated implementations, reviewed the position of the profession vis-a vis authority control, identified the functional requirements of systems of authority control required by archives and museums, and suggested directions for future applications of authority control. Audience questions and an extended group discussion at the conclusion of the day were transcribed and edited for this volume.

The volume is introduced by an original paper, by Avra Michelson, editor of the Proceedings, that presents the state of standards in the archival profession and identifies the place of terminology control within the framework of archival description standards.

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 series, including also four issues of the Archival Informatics Newsletter, are US\$160. p.a. for domestic addresses; US\$180 for foreign addresses
--