

ARCHIVES
AND MUSEUM
INFORMATICS

Cultural Heritage Informatics Quarterly

VOLUME 10 • NUMBER 1 • 1996

Archives and Museum Informatics

Cultural Heritage Informatics Quarterly

Volume 10 • Number 1 • 1996

CONTENTS

EDITORIAL

WorldWideWeb Addresses Achieve the Status of Necessity / 1

ARTICLE

State of Electronic Records Management Worldwide / 3

COLUMNS

- EUROPEAN MUSEUM REPORTS

Videomuseum and Vasari MAGNETS: Museums and Art Galleries New Technologies Study / *Xavier Perrot* / 41

- CD-ROM REVIEW

The Computer Collection of Art on Contemporary Fine Arts published by Russian News Agency ITAR-TASS and ARTINFO MultiMedia / *Ecaterina Geber* / 47

CONFERENCES

Call-for-Papers: DLM Forum (Electronic Records) / 55

Archives and Museum Informatics: Cultural Heritage Informatics Quarterly (ISSN 1042-1467) is published by Archives & Museum Informatics, 5501 Walnut Street, Suite 203, Pittsburgh, Pennsylvania 15232-2311; (412) 683-9775; fax 412-683-7366; e-mail: bpwright@cris.com.

The journal is edited by David Bearman, whose authorship may be presumed for all items not otherwise attributed.

Archives and Museum Informatics carries news, opinions, and reports on information technologies, techniques, and theories relevant to archives and museums.

Submissions of press releases, publications, and software for review, articles, and letters to the editor are welcomed. Copy is preferred double-spaced. Longer articles will be requested in machine-readable form if accepted for publication. Deadlines for contributed articles and press releases are the 15th of March, June, September, and December.

Subscriptions are available on a calendar year basis at US\$90.00 for institutions, US\$50.00 for individuals (paid in advance, by personal check, and delivered to their home address), US\$25.00 for individuals employed with subscribing institutions (mailed to business address), with a surcharge of US\$10.00 outside the United States and Canada. Back issues are also available at a reduced price. All payments must be made in US currency. Credit cards are not accepted.

Archives & Museum Informatics also publishes occasional technical reports available for purchase. Prepaid orders include handling fees. Billed orders are subject to a \$5.00 billing/handling fee. Current titles can be obtained from the above address.

Coalition for Networked Information / 56

Calendar / 68

INBOX

Reports, Books and Articles, Journals and Newsletters,
Ephemera / 69

NEWS / 81

SOFTWARE / 86

STANDARDS

CIDOC Data Modeling Workshop/ *Nick Crofts, Pat Reed* / 89

GILS Advanced Search Facility Project / 91

Standard Notes / 98

CONTRIBUTORS / 101

EDITORIAL

.....

WorldWideWeb Addresses Achieve the Status of Necessity

There was a date, sometime in the past decade, that having a fax address became a necessity of doing business. Incredibly enough, the World Wide Web will achieve the same status sometime in 1996.

The first issue of this journal to carry fax addresses for contacts wasn't until 1989. By 1990, this journal began citing email addresses for some contacts. The first World Wide Web addresses here were published only last year.

What makes all of this remarkable is not that the World Wide Web is so important, but that three major new means of communication have entered our business lives in less than ten years after more than 80 years of only phone and postal service which followed 80 years of postal service alone. It is not unlikely that additional conduits of information will arrive on the scene in the next few years and become as important to business as the existing methods are today. What's more, it is clear that the addition of new conduits does not replace older ones, only supplements them.

So what can we say about the impact of the World Wide Web. First, as the headline suggests, I believe that businesses cannot now live without it. Second, and ironically, we don't know yet what it will be used for. Not only will its actual uses be dependent on the means used to exploit it (software that hasn't yet been developed) but its niche will be determined by its relationship to other communications vehicles. Right now, the World Wide Web

is being used as an electronic billboard, visitor information center, and market sounding board.

So why has it achieved the status of being a necessity? I think it is not simply the cache of being perceived as technically with it. My guess is that it reflects (accurately) that some people will interact with you, order things from you, or find out about you, ONLY if you are on the World Wide Web. So to reach this market, there is no option.

ARTICLE

State of Electronic Records Management Worldwide Spring 1996

The magnitude of recent developments in electronic records management have led me to decide to put together in one place some reflections on numerous major documents that have emerged over the past several months. Outside of the U.S. federal government, I now believe that shifts in attitudes towards what constitutes a reasonable approach to electronic recordkeeping are sufficient to carry us from a fruitful stage of research into a long period of implementation and development. I hope the U.S. federal government will figure out what is going on soon.

In the last issue, I discussed the Australian National Standard on Records Management issued in February 1996 which establishes the basis for understanding records as evidence of transactions and their proper management as a consequence of correctly identifying the business process they document. I also reported on the various projects taken up by state and local archives which are following up on conclusions of the University of Pittsburgh Project on Functional Requirements for Evidence in Recordkeeping. In this issue we can add to those developments the issuance of the Australian "Common Position" statement, my reflections on the completion of the Pittsburgh project, analysis of a report from the Swedish pharmaceuticals firm Astra and the Swedish National Archives on their approach to electronic evidence (SESAM), and a critique of draft guidelines from the states of Queensland and New South Wales, Australia and the U.S. National Archives and Records Administration (NARA).

Together these developments, I believe, set the stage for an implementation phase of electronic records management which marries object-oriented APIs or network transport agents and application systems, even legacy systems, in fully functional recordkeeping regimes. Other implementation projects can explore the marriage of document management systems with recordkeeping, which will prove more difficult because the transaction nexus from which paper flows cannot be automatically linked to the resultant documents. Human analysis at the item level will be necessary to correctly locate the evidential source of documents.

I. The "Common Position" statement of Australian archives

In October 1995, representatives of the national and state archives of Australia, professional associations (Association for Information and Image Management, Records Management Association of Australia, Australian Computer Society, Inc., Australian Council of Library and Information Services, Australian Society of Archivists, Inc.), public sector agencies and large private companies, universities with graduate programs in archives and records administration, and Standards Australia and the National Preservation Office, met to discuss a framework for electronic recordkeeping that could be adopted nationwide. This extraordinary gathering produced a document, issued in its final form in May 1996, which has considerably more specificity than anyone might reasonably have expected. Essentially, it provides a context in which electronic recordkeeping regimes will be established with similar success criteria and outlines key undertakings that will be required to get there.

The premises of the common position statement are quite simple:

Recordkeeping is fundamental to business.

Increasingly business is conducted electronically.

Records must satisfy a business, accountability, and cultural function.

The consequences are by now familiar. Electronic recordkeeping requires systems that "capture and reliably maintain the evidence embodied in records." This demands that "electronic business processes routinely involve the capture of records necessary to document them." The methods of capture must ensure that records "remain accessible and useable for as long as they are needed."

To these ends the common position statement identifies four strategies each of which is grounded in one or more guiding principles. The strategies are:

- Creating electronic records and capturing them into electronic recordkeeping systems,
- Designing, building and using electronic systems that keep records,
- Maintaining and managing electronic records over time, and
- Making electronic records accessible.

The real meat of the statement lies in the principles which underlie these strategies, because it is here that the commonness of the position comes through. Among the basic principles which I believe must underlie any proper regime and which are identified in this statement are:

- Each organization should define its own boundaries for determining when a business communication that needs to be captured as a record crosses into or out of it.
- Each organization should determine, through risk assessment, the degree to which activities require reliable records.

- Each electronic recordkeeping system should be able to segregate records automatically from other kinds of information.
- Each electronic recordkeeping system should be built to satisfy identified recordkeeping requirements, including operational business needs, legal requirements, industry best practice and the expectations of society.
- Electronic records should be maintained in electronic form for as long as they are needed (and)...in such ways that the required evidence of the business process is preserved and accessible, enable them to be understood, and to support their management and use.
- Each organization should specify, capture, maintain and migrate metadata and contextual information should be retained in conjunction with the electronic records.
- Each organization should determine who will maintain and manage its electronic records of continuing value — itself, an archives institution, or another organization with a suitable business and computing environment that is able to maintain the content, context, and structure of the records over time.
- Each organization should use technology to provide electronic access to records irrespective of their location.
- Users should be able to access records as records (with their structure and context intact and adding essential meaning to the content) and as information.

Finally, the common position commits the institutions which were party to it to a legislative, standards setting, training, and monitoring agenda through independent and cooperative activity. More specifically it proposes that each jurisdiction should identify gaps in current legislation, standards, and codes of best practices; that a national standard on electronic recordkeeping be developed

based on the recent national standard on records management; that training needs be identified and addressed; that incentives be developed for adherence and performance measures be developed to monitor compliance; and that “regimes for compliance” be promoted. On the whole, the statement is so impressive that I felt it worthwhile to reprint it here. If similar concrete commitments with explicit guiding principles could be forthcoming from organizations elsewhere, it would be a basis for implementations with a large common market - one of the real requirements for progress. As always, commercial reality will dictate the availability of off-the-shelf solutions.

II. University of Pittsburgh Project on Functional Requirements for Evidence in Recordkeeping

Although (because?) I’m hardly an impartial observer, having served as the principal consultant to the “Pittsburgh Electronic Records Project” as it was often colloquially known, I feel it is essential to report for readers on the methods and results of that study. I think it has occupied a pivotal place in the evolution of ideas found in the plethora of other reports on electronic records management that have emerged in the past several months. Fortunately readers can check out my interpretations by reading the final documentation themselves at <http://www.lis.pitt.edu>.

The Pittsburgh project was not conducted in an ideal, top-down, fashion, but it has nonetheless produced a coherent result that we express in that form. Simply put, we claim that:

- Evidence is a social and legal concept embodied in laws, regulations, and best practices.
- The functional requirements for evidence can be derived from warrant found in the most respected sources identified by lawyers, auditors, business managers, information technology professionals, records managers, and archivists.

- These functional requirements can be expressed as specifications, in prose, or in formal English “production rules.”
- The specifications can be satisfied by observable metadata linked to records at the time of creation and maintained over time.
- Electronic evidence will be trustworthy if the content of transactions can be encapsulated with metadata generated at the time of the transaction and the resulting record is removed from active information management systems to a system which satisfies the requirements for maintaining evidence and providing access.
- Archivists could currently implement such environments, even where legacy application systems operate, but they will be best served if all information systems generate “business acceptable communications” which embody the metadata required for evidence in a layered record structure which can also serve the needs of commerce and of inter-networking.
- The proposed Reference Model for Business Acceptable Communications, and discussions of its implementation requirements, reflects the functional requirements for evidence in their most granular form.

Numerous practical and philosophical issues are raised by these rather bland statements. I would like to discuss some of these further.

We did not begin by recognizing the importance of the widely accepted societal meaning of evidence. Instead we started by inviting expert archivists and records managers to a brain storming session at which these functional requirements were first derived from our collective experience. Only later did we search the literature for justifications. And only after that did we find out

what sources the different professions thought were most authoritative and search for validation of the functional requirements in those sources. To date we have not finished demonstrating that professionals believe the functional requirements more because they have warrant or because they have warrant from the literature of their own professions. This is the subject of Wendy Duff’s dissertation which will be completed this year. Nevertheless, it is clear from both the reception accorded these functional requirements by archivists and records professionals and from the way in which lawyers, auditors, and senior management respond to them that they are an exceptionally powerful way of expressing what archives are.

We did not at first recognize the fundamental difference between the principles governing information systems and record-keeping. Only when we pushed the analysis of functional specifications for evidence did these become clear. Nor did we begin with a clear idea of what tactics might best achieve record-keeping functional requirements; only the analysis of these requirements revealed that the essential problem was how electronic systems are to create records, and thereby pointed to viable strategies for implementation. Here again, the project did not go about defining the technical methods for implementing systems of electronic evidence in a very straightforward way. Even today the “production rules” version of the functional specifications is not sufficiently rigorous and the metadata requirements that were to be derived from those production rules do not follow as logical necessities but rather are pragmatic gestures towards a definition which would command complete acceptance. Nevertheless, what they show is that one can move from the functional requirements to a concrete set of observable metadata and hence that technical means of capturing metadata at the transaction level can satisfy the functional requirements for evidence.

This finding led to the last, and greatest, leap which the project made, and if it is taken up by archivists could also lead to the

greatest gains. Once the project established that evidential electronic records would be those linked inseparably with appropriate metadata, the tactics of standards as a means to achieving evidential recordkeeping became obvious. By expressing archival metadata requirements within a framework (Reference Model) of metadata encapsulated objects which is on its way to becoming the most important foundation structure of electronic commerce, we are paving the way towards universal implementation of the requirements for evidence as a basic architectural element in networks and a fundamental functionality of application program interfaces (API's). Here at last is a strategy for creation, identification, selection, maintenance and access to archival records which directly supports the business needs of organizations and therefore will be implemented and enforced.

But the promotion of archival, or evidentially required, metadata in electronic communications will not take place entirely without our help. I am concerned that North American archivists will not take up the directions suggested by the Pittsburgh project, despite its adoption in the Common Position statement of the Australian Council of Archivists and its influence on the SESAM framework recently sponsored by the Swedish National Archives. What I currently see in North America is an unnecessarily personal split in the profession in which some people seem to feel that acceptance of the Pittsburgh approach is a sign of defeat by alternative analyses such as that out of the University of British Columbia and the Department of Defense. There are quite a few projects taking the University of Pittsburgh study forward, as was evidenced at the February meeting of the project. Hopefully they will begin to introduce new findings and modify our conclusions. It is not a matter of either/or.

III. State and Commonwealth Policies Guidelines in Australia

The Australian National Standard unloosed a slew of state initiatives that in many ways go further than the standard itself in reflecting new attitudes towards electronic recordkeeping and methods for management of electronic records.

One of these is the Queensland Information Policy Board Draft Information Standard 31, "Retention and Disposal of Government Information," dated January 9, 1996. Among the principles articulated in this document are:

Principle 1: Accountability. Each agency is accountable for the retention and management of records of continuing value to ensure that the business, accountability, and cultural needs of government are met.

Principle 5: Record Structure and Content. The essential characteristics of records are content, structure and context. Agencies must ensure that information about recordkeeping systems is recorded and maintained over time. It is necessary to maintain information about the context, or business transaction, in which records are created and managed as well as the structural relationships of the data which comprises the content of the record and the data itself.

Principle 6: Record Accessibility. Agencies are responsible for ensuring that records identified as having continuing value remain accessible to people with appropriate authority, both inside and outside the agency, for the designated retention period.

Definitions in the Guidelines include:

Evidence (which) "refers to the records of a business transaction which can be shown to have been created in the normal course of business activity and which are inviolate and complete. They are evidence of how an office or person conducted their business.

By inviolate it is meant that the records are secure, time-bound, and unalterable. By complete it is meant that records have structure, content, and context.

Role definitions in the guidelines include that records management should:

* ensure that the appropriate framework is in place so that corporate information requirements and recordkeeping obligations are met, and that information of long-term value in whatever format or media is identified and preserved for as long as that value continues.

Responsibility is squarely placed on the chief executive of each agency to maintain proper records.

The guidelines include an analysis of different tactics for long-term retention of electronic records and their advantages and disadvantages as perceived by the authors at the time of publication. Responsibility for assessing these options and maintaining awareness of the shifting benefits and of new options, is placed on the agency. The records managers, presumably, are expected to continue to update guidance of this sort but they cannot substitute for agency-based risk management decisions with responsibility falling on the agency.

The Archives of New South Wales issued Draft Policies on Electronic Recordkeeping for comment early in May. These policies, which build on Australian Standard AS 4390, ISO 9000 and the NSW white paper Documenting the Future, set out seven policies, which can be found on page 14 of this article.

At the same time, the Archives of NSW also issued an exposure draft of a Policy on Electronic Messages as Records. This policy fits within the framework of the broader policy, and carries it to a more detailed and implementable level. I think these policies are

so good, I've reprinted them on page 17 of this article, with permission of the authors.

In February, the Australian Archives issued its policies on "Using Electronic Mail." These appear to be internal agency policies, although they don't say so, but their concreteness about implementation environments and the examples they use suggest it. If so, they are an excellent model for other agencies. They begin by asserting the responsibility of each employee to document his or her activity and of the Australian Archives to be accountable. They also note some of the advantages and disadvantages of using e-mail in the current implementation context. They then explain how to determine if the e-mail message is a business transaction and, if so, how to manage it properly. Subsequently they provide working guidelines on whether to make convenience paper copies or file paper copies and discuss some existing problems with security in the current e-mail environment at AA. Finally they excerpt a section of policy guidelines on electronic mail and voice mail from a Documentation Standard currently being developed by AA for Commonwealth agencies.

Throughout the guidelines there is a straightforward recognition of the current, rapidly evolving state of affairs, together with a clear understanding of what constitutes a record and what outcomes are desired. While it is not always possible for the authors to tell readers what to do, they are able in this framework to explain some of the issues and problems and leave room for future guidance. I like its modest style, although I think it could have defined a more rigorous set of policies, as was done in New South Wales.

Archives of NSW Policy on Electronic Recordkeeping

1. Electronic recordkeeping should comply with recognised best practice

Agencies should keep and manage their electronic records in compliance with standards and codes of best practice and guidelines to be issued under forthcoming State records legislation and with Australian Standard AS 4390, *Records Management*. Electronic recordkeeping should also comply with recognised authoritative sources influencing recordkeeping that are relevant to an agency's business environment, such as the AS/NZS ISO 9000 series of quality systems Standards.

2. Electronic recordkeeping should be built into business processes and tools

Agencies should ensure that recordkeeping is a routine part of business in the electronic environment by building it into business processes and tools. Agencies should use the opportunities afforded by business process re-engineering exercises, the replacement or upgrading of information systems and quality systems certification.

3. Business conducted by electronic means should be adequately documented to meet identified recordkeeping requirements

Agencies should ensure that they adequately and properly document those aspects of their business that are conducted in the electronic environment by making and keeping electronic records. Agencies should ensure that these records satisfy identified operational and accountability requirements and community expectations.

4. Electronic records should be maintained in electronic form

The Government's objective of increasing the efficiency and effectiveness of government processes through improved IT&T services will be supported by maintaining electronic records in electronic form where it is appropriate to do so. Electronic records are more accessible and generally have greater value than printed versions of them. Some kinds of electronic records, such as compound documents, cannot be maintained in hard copy form without loss of content or meaning.

5. Electronic records should be maintained in reliable recordkeeping systems

Agencies should systematise electronic recordkeeping practices through the design and operation of reliable electronic recordkeeping systems. These may be dedicated recordkeeping systems or business systems and processes which incorporate recordkeeping and, thereby, function as recordkeeping systems. Reliable recordkeeping systems result in authentic and trustworthy records.

This principle requires that agencies establish policies, business rules, formal methodologies and procedures for recordkeeping in the electronic environment. Agencies should document these recordkeeping systems, including changes to them, and assign responsibilities and delegations regarding their development, modification, operation and use.

6. Electronic records should be managed effectively as part of a comprehensive records management program

Agencies should ensure that policies, practices and systems for the management of electronic records are fully integrated into their records management programs. While electronic records should be regarded as a mainstream part of records management, compliance with this principle is likely to require specific and systematic attention.

Where electronic records and records in other forms document the same business activities or matters, agencies should manage these records in such a way that all relevant records can be easily retrieved and the relationships between the records are clear.

7. Maintaining and providing access to electronic records over time is a shared responsibility

Maintaining electronic records and providing official and public access to them over time, including as State archives, involves shared responsibilities between agencies and the Archives Authority of New South Wales. Agencies should co-operate with the Authority to establish arrangements for the long-term management of electronic records and for providing access to them.

Establishing and operating effective electronic record-keeping systems and practices within agencies requires a multidisciplinary approach extending well beyond records management. Agencies should make effective use of the necessary range of expertise, particularly in IT&T areas, available in the agency and elsewhere.

Archives of NSW Draft

Policy on Electronic Messages as Records

1. Electronic messages are records

- 1.1 An electronic message is a form of business communication. The sending of an electronic message is a business transaction and therefore a record. Consequently, such records should be retained for as long as required to meet identified recordkeeping requirements.
- 1.2 Records sent and received by a government employee in the course of official duties are to be treated as *official records*. For agencies subject to the Archives Act 1960 these are *public records* as governed by the Act.
- 1.3 As records of government agencies, electronic messages, like records in other formats, are subject to legislation such as Freedom of Information Act and to legal processes such as discovery and subpoenas. The records may also be required by Royal Commissions, the Courts, auditors and other people or bodies to whom or which they may be subject.

2. Electronic messages are to be captured and maintained as functioning records

- 2.1 In order to function as a record, electronic messages require the preservation of their structure, context and content.

Content refers to the content of the message as received. Context refers to the information documenting the source and destination of the message and gives the message meaning (found for example within header information for electronic mail).

- 2.2 In order to maintain their value as evidence electronic messages as records must be inviolate, that is, they cannot be altered or manipulated.

3. Electronic messages must be captured into an identifiable recordkeeping system

- 3.1 Electronic messages required as evidence of business activity should be captured directly into an electronic recordkeeping system.
- 3.2 If the preferred method of capturing records above cannot be satisfied an agency must satisfy one of the following options:
- Recordkeeping functionality can be built into the electronic messaging system.
 - Electronic mail and voice mail, with appropriate contextual detail, can be printed or transcribed and filed into a paper-based recordkeeping system.
- 3.3 Back-up stores of electronic messages within messaging systems and ad-hoc saving of messages to directories are not to be considered a form of recordkeeping system.

4. Electronic messages must be effectively managed

- 4.1 Effective management practices for electronic messages as records will ensure that an agency's accountability and recordkeeping requirements are met.
- 4.2 Agencies subject to the Archives Act 1960 should dispose of electronic messages as records in accordance with the Act.
- 4.3 Electronic messages are to be managed in accordance with sound data management practices.

5. Electronic messages must be appropriately accessible

- 5.1 Electronic messages must be readily accessible to meet business and accountability requirements.
- 5.2 Electronic messages need to be accessible for as long as they are required to be retained.

6. Electronic messages must be appropriately protected

- 6.1 Arrangements for maintaining electronic messages as records must be in accordance with government-approved policies, for example, regarding privacy issues and commercial interests.

7. Management of electronic messages as records must be supported by internal policies, procedures and guidelines

7.1 Corporate level policies, procedures and guidelines should cover the following:

- ownership of the messaging system and messages received or sent
- definitions of terms
- legislation that directly affects the whole organisation's recordkeeping and legal processes
- conditions for use of the corporate messaging facilities including any private use
- outline of responsibilities for management of the messaging system and electronic messages as records
- capture, management and disposal of electronic messages as records
- security and access requirements
- user guidelines for the management of electronic messages as records.

IV. SESAM: Philosophy and Rules Concerning Electronic Archives and Authenticity, Ulf Andersson, Astra AB, 28 February 1996

Last year, Ulf Andersson of the Information Systems staff at Astra AB, a Swedish pharmaceuticals firm, undertook an extensive and international investigation into how to create and manage records in electronic form. The SESAM report, cooperatively developed with the Astra Legal Affairs and Corporate Security groups and the Swedish National Archives, is the product of that investigation and was the subject of a conference held by the Swedish National Archives late in May.

SESAM is written to be “read and understood, not as a handbook,” but it goes further in the discussion of implementation issues than virtually anything ever written about electronic records management. It is grounded in a foundation of definitions and axioms about management of electronic information that I believe are absolutely sound. While somewhat different from my own, the basis remains that records are evidence of transactions and can be preserved and made available by capturing metadata documenting content, structure and function in encapsulated objects stored in distributed system.

Part one of the report addresses “Strategy.” It opens with definitions in which Andersson explains the background axioms:

A business transaction is an event which occurs in the business process. A business record is the documentation and evidence of a performed business transaction. Business records must be maintained over a certain period of time depending on internal requirements and external demands. If the record is to be maintained for a long period, in other words is to be archives, two areas of responsibility are essential:

1. The information owner must ensure that the stored information is stored within its context to make it possible to understand the record in the future.

2. The archiving operations must ensure that they maintain the record in a way that the authenticity of the record is not lost.

The report goes on to note that “a business record can not exist without a medium to carry it and formats and syntax to express its content and context” and that “explanations about stored information and structures (in this document called metadata) are often incomplete.” In further explanation, it points out that:

A business record is the documentation of a business event. Electronic business records and their management are not isolated to what is entered when using an application. An application has two components:

1. The logic and structures within the application software which enable the users to operate on the business records.

2. The data and information entered . . . by the users.

This section introduces for the first time a potentially interesting new idea: “the causa.” According to Astra, “a causa is a sequence of events forming a logical unit, controlled by rules and limited in time and extent. One or several business records will belong to a causa. A receptor is the documentation of the conduction [sic] of a causa.” The utility of this concept becomes apparent later in the paper as an organizing principle for records retention decisions and for the linking of related business transactions to business process rules within the equivalent of the traditional paper file.

Following the definitions and background section, Andersson explains the user’s vision:

Astra must, within the next five years, be in a position to commence the transition from paper to electronic media for approved and signed business records. A holistic archiving environment, physically and logically separated from the IS and IT production environment, shall be established. The evidential value of an electronic business record is not to be affected by the format for either electronic transferring or archiving.

Considerable attention was paid in the framework to the issues of migration. In particular a tactic decision was made that the “migration of business records shall not include a conversion of the logic of the applications used to manage them” because that logic would be documented in metadata, and that “it must be possible to extract documentation and load it into other systems (i.e., the systems of the various regulatory authorities) without affecting the value of the record as evidence,” by means of encapsulation.

Other tactical decisions, with which I fully agree include that:

An infrastructure shall be established that is physically and logically separated from the production environment for archiving of documentation of a causa.

Formats, logical and physical structures, shall be identified and utilized in such a way that a migration of stored business records is only a migration of formats used for storage, not involving changes of the structure.

Another tactical decision, introduced by Astra, is a requirement for event logging independent of the records. This is consistent with our requirement in the University of Pittsburgh functional

requirements for evidence in recordkeeping (www.lis.pitt.edu/~nhprc) that records must be comprehensive. Astra carries the requirement further in proposing that "it shall be possible to invoke certification (based on) event-logging . . . by an external public body...."

Part two of the report is entitled "Generic Models." Some of the theoretical principles it espouses may seem redundant, but the progress throughout the paper, from greater abstraction to greater specificity, requires a bit of overlap as concepts are carried further. For example, the examination of how the "archiving IS and IT environment, is to be independent of the data production environment" leads to recognition that the "production environment, the logging environment and the archiving environment are, from an organizational point of view, to be totally independent of each other. No one is allowed to have access, outside operational systems, to more than one area." User organizations are responsible for documenting *causa* and creating business records. Archiving organizations are responsible for maintaining authenticity of records and making them available."

In this part of the report, strategies are refined into tactics for implementation. The principles of independence of the three systems environments, for example, lead directly to a discussion of how to structure documentation or "requirements on models of receptor and envelope" which are implementation tactics. What is proposed is that records as evidence must be extracted from the work-flow, must contain an ID related to the event-log, and must reference approvals and required actions that are part of the rules of the *causa*. This documentation resides in a receptor which includes metadata identifying the *causa* and its management (context), discovery and retrieval data, data necessary "to ensure the capability for long term understanding of the *causa* and understanding of the context of the business record." As explained in the report, "in its simplest form, the *causa* is only metadata concerning creation of a business record and the reference which

allows comparison between the event-log and the metadata or content of the *causa* and envelope."

Ultimately, Astra accepts the metadata requirements I advanced in the University of Pittsburgh project on functional requirements for evidence in recordkeeping. In the text and explanations, however, it is evident that Astra pays less attention to issues of what constitutes necessary and sufficient structural metadata than will ultimately be needed for implementation. Their concept of structural metadata, usually called "metadata for interpretation" but sometimes included within "context defining metadata" by the report, takes a too limited view of the requirements for "re-presentation" and linking in the future, even with the addition of the conceptual linking mechanism introduced by the *causa*. In addition, they seem to rely in part on WORM media properties and cross comparisons between hash-like information in event logs and records for ensuring inviolability than on encapsulation type protections. I expect these are differences in tactics only, and may be resolved as implementation proceeds, especially in light of the rapid progress being made in the past several months within the realms of electronic commerce on the use of encapsulated objects for resolving terms and conditions of access to records.

While not explicitly introducing the concept of risk management into their model, the Astra report reflects a recognition of the reality that the strategies and tactics we adopt in every instance are reflections of judgments about risk. One particularly sophisticated discussion of these issues takes place within the context of determining how much metadata is required to ensure that any given record can be considered preserved after a migration. The report distinguishes between preservation for periods of time during which the organizational and procedural structures that created the record still exist and preservation beyond that time.

Other tactical implementation problems faced by the Astra report include the way in which files of different applications will

be linked together in records and the requirements for recalculating “signatures,” by which I presume we are discussing physical document hashes, across migrations.

Eventually, these discussions lead to how to structure the archives, and particularly how to insulate the archives from organizational changes and technical dependencies. The tactical framework adopted is one of modularization of the logical dependencies of records so that “when formats become unsuitable for a selected purpose they must be migrated as one activity within the logical unit.” Of course, this will become quite challenging when other parts of the record, or other linked records, have different dependencies, as Astra acknowledges in the follow on sentence [“The migration must be done under conditions that guarantee the executability of links between business records within a logical unit.”] and in the subsequent discussion of the difference between business records of a causa from a single computer system and from multiple computer systems.

The broad answer, of course, is standards. But Andersson’s presentation of the reasons standards must be introduced, and where and how they are applicable, especially with respect to the discussion of SQL and SGML within the retrieval standards framework as part of the requirements for access, is particularly (and refreshingly) sophisticated.

The final section of this part of the report addresses authenticity. Simply put, “it must be possible to prove if a business record has been changed or not.” The proposed tactical framework is that:

1. Multiple ways to support proving authenticity must exist.
2. Both context and content must be described in such a way that they are easy to explain.

Ultimately:

“All transactions must be stored with sufficient metadata in a way that ensures that they are understandable outside the computer system creating them.”

The third part of the report identifies issues for Astra raised by the generic models. These are useful for understanding the rules and responsibilities flowing from adoption of this framework by any organization. The discussions in this section relate, broadly, to issues of architecture.

The first section addresses standards requirements including the adoption of ISO and CALS frameworks which encompass SGML, HyTime, Step, GKS, IGES, JPEG, TIFF, CD-DA, MPEG, MJPEG, and eventually, Unicode. It adopts the architectural requirement of distinguishing between recordkeeping systems and information systems which is fundamental to the University of Pittsburgh conclusions. Oddly, it asserts some other architectural rules which I cannot understand at all — for example, that “records logically belonging together shall be physically stored together,” and that time intervals need to be established for regular for introduction of new records into archival storage. At first I was taken aback by the claim that “approval of documents with electronic signatures must be made after data and information have been converted to the formats to be used for electronic archiving” but ultimately it became clear that these electronic signatures were RSA algorithms (calculated hashes supporting public/private key access control) and therefore logically would have to be created as the last activity before records were finally put into the authentic store.

In this section the author introduces the fact that the technical basis for process metadata management is the adoption of STEP, the CALS originated standard for documentation of process engineering using the Express meta-language (p.46) and the adoption of the Pittsburgh functional specifications for evidence in record-

keeping (p.47) and discusses nitty gritty architectural issues involved in security, access control, encryption, logging, and current workspace management.

The report recognizes that:

To get user acceptance the systems must provide an automated collection of the necessary metadata. That means support from the moment the business record is created, finally stored, reused and until it is deleted.

Further:

Monitoring of changes of the IT market must be performed in such a way that the conversions to new formats are made at the appropriate time.

Systems for information retrieval shall be directed against archives and never against operational systems.
(p. 53)

Part four of the report addresses industry and organization-specific implementation problems such as the requirements for evidence in patents, new drug applications (NDAs) and requests for permissions to conduct clinical trials (INDs). Not only are there concrete requirements associated with these events by the numerous national jurisdictions in which Astra does business, but there are significant demands on business records and business process documentation as a consequence of the kind of activity in which Astra engages. The discussion of how rules for opening projects, defining studies, selecting instruments, determining test species, and defining measurement parameters have impacts on the way in which the metadata envelopes are defined in these business processes (causa).

The final part, entitled "Technical Aspects/Technical Implementation" opens with a discussion of STEP and Express as

documentation meta-languages and the relationship these have to content represented in SGML/HyTime. I fully agree with the choices made and the arguments presented, although actually carrying out the markup and metadata capture will prove difficult. Preferences expressed for modularized and slim DTDs are definitely correct, but the premise that the records, once archives, should be accessible from Web-like tools (by which one assumes browsers, parsers, and stateless retrieval engines) is hard to justify and the author makes no effort to explain this choice. Decisions that records will be stored on CD-ROM and DES algorithms for encryption will be used are plausible enough. Lack of discussion of access control (Kerberos?) is less understandable. The discussion of how multiple format elements will co-exist within a record envelope is, I think, the best solution possible today, but it raises lots of issues about de-synchronization during migration which the authors understand.

Throughout the report, excellent graphical representations make the intentions of the author clear and help to overcome the ambiguities created, in part, by the fact that English is not Mr. Anderson's native language and probably was not the original language of the document.

Problems And Issues:

This report is the most complete discussion of the implementation issues surrounding the adoption of item-level electronic evidence management, including in my own writings. It follows in almost all respects a set of principles and logic which I have been promoting for a number of years. My quibbles with it, therefore, are not intended in any way to redirect the approach, but only to strength the implementation decisions and give the whole effort a better chance of succeeding.

The first problem I see lies in the introduction of the concept of causa which is potentially valuable but probably will prove cumbersome. I understand that it has been invented to overcome

the ugly reality that what constitutes a business transaction is subject to rules which relate to business practices and meanings, which, in fact, often contradict the facts of physical steps involved activities or even the information flows observed in work processes. What we find in fact is that some steps of an activity are considered by participants to be independently documentable and to require evidence while other steps are considered subservient and only meaningful as part of the business activity of which they form a part. But the implementation proposed by Astra (see. p.38 for a clear demonstration) requires the creation of a separate information systems entity, outside the organization, to log events and relate them to business processes in order to build a boundary definition of the information entities participating in a transaction. While this creates a one-to-one relationship between the envelopes of evidence (receptors) and the business transactions (causa) it does so only by referring the question of what constitutes a business transaction to a prior stage, the definition of causa. I prefer the much simpler model which considers all actions resulting in communication to have crossed a records boundary and links records with predecessors. As an aside, I believe the proposition that records should be upwardly linked unnecessary if the separate causa documentation environment is actually going to be built.

A second concern arises from the emphasis in the report on WORM-media. This is disturbing to me because I don't think the logical requirements should be satisfied in such a hardware/software-dependent physical way, and because occasional statements like "archiving and migration of log-archives should preferably be done by a certified third party. The physical medium should be transferred to this organization directly after completion," creates a likelihood that media will end up being untended and the logical records will cease to be preserved. Throughout the report, references to CD-ROM and WORM media strike a jarring note since they are short-term manifestations of dense storage media and have very short usable lifetimes. The author knows this, of course,

but the decision to leave such references in the document suggests that others may find comforting what should be a source of worry (that the bits recorded on such media may last a very long time).

Finally, the report suggests structures for a DTD for causa envelopes. These contain types of metadata which map to my Reference Model for Business Acceptable Communications. Each contains a header (Ident" and SearchKey" to Astra), terms and conditions (Security"), context (Project", Study"), structure (Transact", Signatur", Extlinks"). However, I believe the layers of the Reference Model, which was designed to fit into the larger universe of encapsulated envelopes developed for other purposes, provides a much better direction for future development than the essentially random groupings imagined in the Astra DTD. I think it will be critical for Astra, and for the National Archives of Sweden and other governments, to adopt architectures for metadata encapsulated objects that can and will be widely provided within systems architectures and network communications infrastructures in the future. As a consequence, the choice of how the envelopes will be constructed and where the various essentially essential metadata will be found is not a trivial issue and cannot be left to chance. Requirements taken from the world beyond archives, particularly from electronic commerce and wide-area network object management, need to be addressed in the design of the envelope structures.

V. NARA, Records Management Requirements for Electronic Recordkeeping (College Park MD, 1996, DRAFT for comment May 1996)

While in many ways the draft guidelines for electronic recordkeeping issued for comment by NARA in May are NARA's best position on electronic records to date, they fall short of being useful because they lack the coherence that would make good, or implementable, policy. Instead we have a confused intellectual framework and hence a confusing document.

Problems with the draft surface in the Preface which focuses entirely on “preservation” of records in electronic format for “transfer of permanent records” without even passing mention of the critical issue of records creation which ought to precede any concern with preservation. In so far as records creation is even hinted at, it is to exclude from consideration any records “stored in formats that cannot be converted to ASCII without loss of content and some records related or linked to other records.” These are the very records it needs to address. It also excludes most records created by employees in the course of business using application systems where linkage between records is essential.

More serious problems are revealed by the Introduction which makes clear that NARA has not been able to conceptually segregate functionality related to (1) creation and maintenance of evidence; (2) managing information related to business needs of the organization; and (3) archival management of records over time. NARA’s definitions of information systems and recordkeeping systems demonstrate that it is unclear about the distinction between them. This leads to a confusion about systems architecture that makes NARA sound as if it still wants agencies to implement information systems that fully incorporate recordkeeping functionality. One of the findings of the past few years is that information systems are designed to be non-redundant, timely, and manipulable, and recordkeeping systems must be non-manipulable, time-bound, and hence highly redundant. Trying to achieve both ends within a single application system will require expensive redesign of virtually all agency information systems. This is not a cost effective or maintainable architecture. Ultimately NARA will need to accept that agency electronic information systems are designed to facilitate the business of the agency and that recordkeeping systems are distinct repositories for managing the evidence of transactions (records) created during the course of business. Information systems hold convenience copies of data for the purpose of providing information; they cannot be made to hold

records without unacceptable burdens being placed on their design and migration.

NARA also seems unwilling to accept that records are created by transactions even when the transactions are not ones in which the government has an abiding interest. As such it introduces a distinction between “documents” or “documentary materials” and “records” and places the burden for determining which documents are records on the end-user. End-users, of course, neither understand federal records laws nor are impartial with respect to decisions about keeping records based on their content rather than on the business process. One could hardly elect a group less likely to make the proper determinations. The continued effect of this distinction is to make automatic capture of records almost impossible to implement and to reduce the comprehensiveness of recordkeeping. It also reduces the evidential weight that can be assigned to those records which are retained, since they are always tainted by the possibility that they have been selected by users to tell a particular story and are not the full story in themselves.

From time to time, the NARA guidelines seem to suggest that they understand that records retention decisions should be grounded in the analysis of business transactions, but even when discussing business processes, their confusion about what is essential and what is tangential to records management principles is revealed. In a series of irrelevant asides about computer system selection, they advise agency data processing personnel that “adoption of electronic recordkeeping must be based on agency mission requirements and the functional processes that produce records. Before installing such a system, the agency should conduct information flow, process flow, and cost-benefit analyses to determine if improvements or changes in processing of information are needed and desirable.” (p. 3).

So what is it that NARA is trying to say? In the most explicit statement in the volume, they assert that:

“To meet recordkeeping requirements, electronic recordkeeping systems must control or allow the user or system manager to control the creation, identification, storage, accessibility, retrievability, integrity, security, and disposition of the records in the system.” This sounds good, but within the text, many of these terms remain undefined, ambiguously defined, or defined in ways that conflict.

The guidelines themselves are organized around three major stages in the records life-cycle: (1) Creation and Receipt, (2) Maintenance and Use, and (3) Disposition.

Creation and Receipt. The first requirement associated with creation and receipt is “designating when a document is a record” which NARA says depends on determining that it is “final, signed, or otherwise in a form that should not be changed” but then goes on to say that “in some cases interim or final drafts or working materials may also be records...” This kind of guidance doesn’t help anyone. Why can’t NARA just say that documents created or received in the course of business are records? Period. And then leave the decision about how long to keep certain drafts and working papers to the scheduling of transactions? The concept that agencies should provide proper training to all employees so that they can make proper determinations about what is a record sounds like a prescription for reducing federal productivity to zero or creating a totally haphazard body of evidence.

The only other requirement discussed in the document regarding creation and receipt is that records are required to have unique identifiers and that multiple documents forming part of a single record should be linked. Not only is this inadequate, NARA’s presentation of these requirements is utterly confusing. In the “Definitions” section, “document” is defined as “recorded information regardless of physical form or characteristics. Often used interchangeably with record,” leaving no meaningful distinction between two terms used differently in the text and no notion of how one establishes the boundaries of either a document or a

record. What is the operating principle that enables us to establish whether documents are linked, or belong to separately identified records? Is there a missing concept of business transactions?

What does an “identifier” consist of. It sounds like a unique number, but then, almost parenthetically, in discussing record identification information, they say, “the system should be required to record the identity of the record creator or source of the received record, the date of creation and receipt, the level of security classification or other access restriction, if any, file classification designation, indexing information such as subject or thesaurus terms, and records disposition schedule citation if different from file classification designation.” Obviously additional contextual metadata is necessary for records to be uniquely identified. Since NARA appears to recognize a number of other requirements why don’t the guidelines state what they are?

Why does NARA then insist that this data should be “compiled into a data dictionary . . . [that] . . . specifies the length, format, access control” etc. Not only is such data management advice gratuitous, in their own definitions again “data dictionary” is defined entirely as documentation of physical databases — “a list and description of all the files, fields, and variables used in a computer system” without any reference to logical entities, business processes, or even system procedures and routines. These contradictions undermine the credibility of the document.

Somewhere, NARA staff recognized that the set of requirements discussed under the heading “creation and receipt” were too limited, so they inserted a list, intended for a “box” off-set from the narrative. In the box they list 14 requirements, almost none of which are dealt with in the prose. The organization seems ad hoc (none are tied to any higher level principles) and the effect is almost certainly going to be that these more detailed and useful instructions won’t be used because people won’t remember the 14 requirements. But the list contains most of what is really wanted if NARA staff can find a way to drop their current text and

substitute a prose version of the list with some principles to assist in organization.

Maintenance and Use. The second section of the document reveals a complete confusion in NARA as to what a recordkeeping system is or why they want to specify its functional requirements. The first requirement — search capability — is an extended discussion of functionality required by offices in the active use of records, again accompanied by unnecessary advice to agency data processing personnel about types of storage strategies to employ. The second and last requirements — security and backup — are equally unwelcome advice on basic systems administration. A single paragraph addresses the fourth requirement, integrity of records, which contains useful text and meaningful distinctions. It explains, for example, that “once a document is designated as a federal record, as defined by the agency and implemented during records creation or receipt, the system must prevent any alteration of the record” and that changes are new records, but that such changes should be linked for business and contextual reasons. Finally, “the system should effectively distinguish the final version from a working version so that users know immediately whether they are viewing a draft or an approved document.” All of these are not only good points, but they establish concrete and testable benchmarks for recordkeeping systems. For example, the requirements raise crucial questions agencies need to answer about their recordkeeping systems:

- Do they designate documents as federal records during creation or receipt?
- Do they make such designations based on implemented agency criteria?
- Do they prevent alteration of records?
- Do they create new records for any changes made to pre-existing records and link these back to the original?

- Can they identify the end of a business process and thereby distinguish final drafts from working drafts?

Unfortunately the guidelines rarely help the reader to move this far towards implementation. Except for a few asides in the discussion of search capability, the draft guidelines literally fail to address software and hardware dependence or to mention issues associated with privacy, confidentiality, proprietary data, leading to requirements for redaction. Another “list” for a box associated with this section of the guidelines itemizes seven requirements including allowing searches on defined fielded data and supporting multiple-user access (neither of which, clearly, has anything to do with necessary requirements for maintenance or use of records, although either might be a desirable feature of a delivery environment).

Records Disposition. The third section of the guidelines continues to provide unnecessary advice about systems management (“storing inactive electronic records off-line pending their eventual disposition may improve performance...”) while neglecting to give advice on records retention management that could have been helpful to agencies. It might have included a suggestion that agencies record the schedule authorizing an action or the transaction that generated a record, rather than recording a concrete date or time period. Such “hard” data as dates and time periods resist change when the laws or regulations surrounding records retention are revised; too many records with different justifications for retention share the same time period or date.

Systems Documentation and Migration of Data. A final section of the guidelines addresses what in NARA’s framework seems to be miscellany. Arguing that “the quality of system documentation ... is a key factor in ... the ability to maintain the integrity, authenticity, and legal acceptability of the records for their entire life spans,” NARA cites various elements of that documentation without any explanation of why they are of value. It makes no reference to the item-level structural and contextual

metadata mentioned within the section on records identification, but exclusively addresses systems-level documentation. For example, they contend that “documentation should specify how records are added, deleted or replaced in the system...” and that it should “include any agency-compiled thesaurus used for data entry...,” although I can’t see what use either of these will be to ensure evidence, unless the evidence one is seeking is about the operations of the information system as opposed to evidence of the transactions those systems might have recorded.

The advice NARA provides on migration will not prove very helpful to agencies. They will hardly be surprised to discover that “a critical aspect of implementing system changes is ensuring that existing electronic records remain accessible over time,” but the claim that “this can be accomplished by migrating the records to the new version or the new system” certainly begs all the important questions of what constitutes the essential properties of the record which will make the migration and how the agencies ought to control and document the transformation process. The issue of how to ensure preservation of complete electronic records across hardware and software generations is too important to be left to an exhortation, bald assertion, or wishful thinking. Different types of dependencies need to be addressed and strategies for overcoming each need to be examined in anything purporting to be “an instructional guide.”

The final section of the draft guidelines left me dismayed. The definitions are so inadequate, and so unrelated to the context of the use of these terms within this document, as to leave the reader completely at a loss for making any distinctions within the text. The definitions of electronic information systems are virtually indistinguishable from recordkeeping systems. The definition of metadata specifically avoids mention of content, structure, and context metadata that is widely accepted by electronic recordkeepers elsewhere. The terms “end user” and “version,” which have operational meanings in the guidelines as used are given industry

standard definitions out of a glossary somewhere that is utterly useless in the context of these policy guidelines.

If NARA intends its guidelines to be taken seriously by others involved in electronic recordkeeping around the world, it needs to completely rethink its position. If it merely wants federal government employees to be able to implement its position, it will nonetheless need to completely rewrite this draft. The best thing one can say about this draft is that it was a good thing NARA decided to issue it for comment before issuing it.

Conclusions

These developments reported in this short review are, of course, not all that is going on worldwide in electronic recordkeeping. In the same time period, I’ve seen internal and consultant reports prepared for the National Archives of Canada, drafts of recommendations from the ICA Electronic Records Committee, and numerous other working papers. I believe that when taken as a group, however, these reports strongly suggest that building electronic records management guidelines from axioms and first principles of records management, based on transactions and the concepts of content, structure, and context, does enable archivists to move towards implementation frameworks that show considerable progress and to articulate policies that are internally coherent.

Not all the policies, nor all the implementation frameworks, fully agree with each other. Nor will they all prove equally successful. But they are extensible, and their implications can be defined. Specifications can be derived and definitions can be made into building blocks for action.

The next stages in our approach to electronic recordkeeping is to test architectural and systemic innovations which alone have the ability to scale up to manage the huge numbers of electronic transactions that will otherwise overwhelm us. Criteria for success in these implementations will be the amount of added effort that

archivists have to make as opposed to the contribution made by computing standards, telecommunications and networking software, agency business process management, and standard work practices. The extent to which archivists can adopt approaches that insinuate themselves into these other domains of practice will simplify the archival role and will make success more likely.

EUROPEAN MUSEUM REPORTS

.....

Videomuseum and Vasari **MAGNETS: Museums and Art Galleries New Technologies Study**

Xavier Perrot

In order to take stock of the situation concerning the use of new technologies in European museums, the European Commission (Directorate-General XIII) has launched a study on the needs and ongoing projects in this sphere entitled MAGNETS. Videomuseum and Vasari Enterprises have been entrusted with the task of conducting this study which falls into three parts: a Museum Survey, a Technical and Economic Review, and Synthesis and Results.

Vasari Enterprises (UK) was established in 1993, with a view to promoting the digitization of the visual arts. Vasari is directed by James Hemsley, the well-known organizer of the EVA conferences in London, Paris, Berlin, and Florence. The company participates in research projects such as VAN EYCK and also provides a strategic consultancy service to the cultural sector.

Videomuseum (France) is a consortium of 40 institutions managing collections of modern and contemporary art, including national, regional, and city museums; foundations; and FNAC and FRAC (the national and regional French bodies in charge of buying contemporary art for public collections). The consortium develops methods and tools using new technologies for managing and documenting the collections in order to better inventory and disseminate its patrimony. We met with Jean-Francois Depel-senaire, head of Videomuseum.

XP: When did the idea of creating Videomuseum take place?

Jean-Francois DEPELSENAIRE: It has been a continuous process since its inception in 1984. At that time, the cultural heritage inventory databases from the Ministry of Culture were all based on the MISTRAL model, derived from the Honeywell-Bull or the Basis technologies. They were basically mono-file systems, not at all flexible when it came to changing information as required for managing a museum collection. In 1984 the FNAC changed this approach, analyzing its user requirements and asking how to meet them through use of multi file databases run on microcomputers. This was the early beginning of the Videomuseum idea. In 1985, "Anciens & Nouveaux," a major exhibition at the Grand Palais, presented the most recent acquisitions of the French public art collections. The actual pieces were accompanied by a complete inventory that visitors could access via six micro-computers. It took four months of intensive work to compile the FNAC software, develop extra components (acquisition), and document and review the 16,000 notices it contained. In 1986 it became obvious that images should be added to the database, and that the scope of the inventory of 20th century art collections should be broadened to a national one. Over the next three years, the project was reevaluated and restructured. The Videomuseum non-for-profit association was set up by the end of 1990 and begin its activities in 1991.

XP: What is Videomuseum today?

Jean-Francois DEPELSENAIRE: It is a project aimed at a computerized permanent and systematic inventory of the 20th century art works from all museums and public collections. Videomuseum is an consortium with a six-person permanent team; a documentation database (accessible to art professionals) inventorying about 100,000 art works with notice texts and illustration images; and cataloging methods and software that provide both local collection management and research capabilities on all

inventoried collections. The consortium is supervised by a Board of Directors ("Conseil d'Administration") representing all the partners and chaired by Bernard Ceysson, director of the Musee d'Art Moderne de Saint-Etienne. A Scientific Committee has been set up, chaired by Germain Viatte, director of the Musee National d'Art Moderne and of the Centre de Creation Industrielle du Centre Georges Pompidou. The Videomuseum Users Club controls the cataloging methods and computer tools.

XP: What about the "Multimedia Dictionary of Modern and Contemporary Art" that was introduced recently at the MILIA'96?

Jean-Francois DEPELSENAIRE: It is part of the Videomuseum mission to explore new technologies and new ways of promoting and using information about the collections. The CD-ROM you mentioned is one means toward that end. In this copublishing venture (Editions Hazan, Reunion des Musees Nationaux, Thames & Hudson, AKAL), Videomuseum had the leading role for the design and production management (IDP was the executive producer and "Tout Pour Plaire" did the graphic design). One-third of the images in the CD-ROM also come from the Videomuseum database. This dictionary has been selected as a European project and benefits from the support of the IMPACT 2 program [cf: European Policies Toward Interactive Multimedia for Museums: Part II, vol. 9, #1, 1995] as well as from the Centre National de la Cinematographie and the French Ministry of Industry. Videomuseum is not going to become an electronic publisher. The point is that our database is a resource of structured information on the collections. We are interested in exploring the best ways of taking advantage of this asset. We are working as well on projects such as CD-ROMs of individual collections, a WWW server, kiosks for museum goers, and new kinds of institutional distribution (schools, libraries, other museums). We have also already used the database to produce paper catalogs. The knowl-

edge we gain with a first inquiring experience in each of these areas is shared with all the members of the consortium.

XP: What is your specific policy toward the World Wide Web?

Jean-Francois DEPELSENAIRE: Online diffusion of information is not exclusive. Depending on the kind of use a database might receive, it makes sense to have local mirrors (schools, etc). For a huge database, the Web is just a complementary way of access. Should this access be 100 percent free? Without control? As we deal with 20th century art, we have to manage the author rights, especially since most of the artists are alive! On the technical side, it would be difficult today to have all the high definition images on the net. They could be stored on local CDV at specific users' locations such as museums, universities, and schools. However, we are currently working on a prototype that will soon render accessible part of the Videomuseum database on the Web.

XP: What evolution are you expecting for Videomuseum?

Jean-Francois DEPELSENAIRE: As part of an ongoing effort, we are working with representatives of the cultural field to increase the number of French institutions taking part in the project so that the database will be as complete as possible on a national scale. Moreover, we are planning to extend the idea of Videomuseum internationally. A network of major institutions in Europe is already working closely with us, but the project is still open. We should soon submit a specific proposal to the European Commission which could support the project. Of course, this approach implies the definition of a shared strategy that enables international, cross-platform, and multi-language documentation and information exchange. What is at stake is the gathering of a worldwide electronic catalog on contemporary art.

The international consortium that will work on this demanding initiative will benefit from the experience of Videomuseum. Jean-Francois Depelsenaire and his team have a valuable understanding of methodologies from their five years of success in managing this kind of project in France. On the other hand, the international extension of the project will gain an up-to-date insight of the effective use of technology in museums abroad from Videomuseum's commitment to the MAGNETS study. The museum survey will be coordinated and, for the most part, conducted by Jean-Francois Depelsenaire and Xavier Perrot. MAGNETS will survey of museum officials in order to provide a description of the general policies in force in each European country. With regard to the actual analysis of requirements and experiences, the study cannot of course expect to draw up an exhaustive inventory of all the applications and projects carried out in European museums.

The project will focus on the current policies in two groups of museums, "Modern and Contemporary Art, Fine Arts" and "Archaeology and History," and will be conducted through a score of meetings and personalized interviews with museum officials. Videomuseum will assume responsibility for the former and Vasari for the latter group. In addition to the projects and applications operating in the target museums, other reference information about the principal software, data banks, applications, programs, Internet sites, conferences, and publications will be compiled and analyzed. In particular, the various actions and programmes initiated by the European Union will be reviewed and details provided of developments in the G-7 "Information Society" programme. The international perspective will be also addressed through encounters with museum officials from Eastern Europe and a review of the latest developments in North America accomplished during attendance at the AAM conference and specific meetings with museums in New York, Minneapolis, and Montreal.

Vasari will lead the technical and economic review, addressing twelve or so sensitive topics involving the development of new technologies, with special emphasis on areas of particular practical interest for museums, e.g., copyright and royalties, online/offline markets, industrial developments, an exhaustive study of uses, public/private economy, digitization, supports, networks, technical bases, data standards, applicable software, and multi-lingualism. A synthesis of standardization initiatives and studies of cataloging methods will cast light on the role played by the various organizations.

The results of the MAGNETS study will emphasize the most significant trends and developments in the field. The conclusions reached should facilitate the political decision-making process at the European Community level. The study will be submitted in September 1996 and the public results published on the Internet by Videomuseum in October (<http://www.videomuseum.fr>).

For further information on MAGNETS or Videomuseum contact Catherine Mueller at: Videomuseum, Centre G. Pompidou, 6 rue Beaubourg, 75004 Paris, France. Phone (+33) 1 42 71 61 39; fax (33) 1 42 71 62 90; magnets@videomuseum.fr

CD-ROM REVIEW

.....

The Computer Collection of Art on Contemporary Fine Arts published by Russian News Agency ITAR-TASS and ARTINFO MultiMedia

Ecaterina Geber

In any multimedia presentation, a tension arises from coming back on one of the multiple possible paths from a multimediated world to a sequential display. This illuminates how difficult it is to reduce messages encountered in multi-dimensional material to a linear presentation. The two CD-ROM volumes reviewed here were published under one title, Contemporary Fine Arts, advertised by their creator, Dr. Sergey I. Senkin, General Manager of ARTINFO, as "... a real opportunity to access the achievements of contemporary fine arts worldwide." They represent a challenge for the interested public. To the question, "Why contemporary fine arts?," ARTINFO quickly responds, "Because we decided to begin with the future: Russia's living artists."

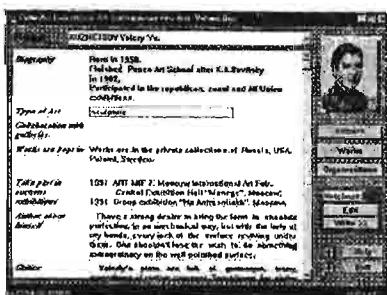
The two Contemporary Fine Arts disks are part of a series called A Computer Collection of Art, published by ARTINFO Multimedia Publishing House in conjunction with the Russian news Agency ITAR-TASS (Central House of Artists, 10 Krymsky val, Moscow, Russia 117049, Tel: +95-290.3725, fax: +95-290.3377, attention: Dr. Sergey Senkin). The series is designed to distribute information on artists, works, galleries, museums, and events, each representing a variety of schools and genres, and to facilitate contacts between the artists and all the interested readers of the CDs. The disks are produced in both English and Russian versions.

The first volume, published in 1994, presents the creative activity of 154 artists from Russia, Kazakhstan, Ukraine, Georgia,

and Azerbaijan, and includes more than 1,000 works of painting, graphics, sculpture, and decorative art.

The second CD-ROM, released in 1995, adds a selection of 100 artists, primarily from Russia and Moldavia, with samplings also of artists from Israel and France. In addition to the 1,000 works, it includes mini-art shows combining images and electro-acoustic music to enrich the resource with the multimediated flavor of the new technology.

The colorful opening window of the Contemporary Fine Arts disk, volume I [Fig.1], provides the user with three possible choices: AUTHORS, WORKS and ORGANIZATIONS (in addition to EXIT and ABOUT DISK). Each one leads to a screen organized having a recurrent pattern which provides a comfortable familiarity while browsing. There is a scroll bar with an alphabetic vertical choice list on the upper side of the screen from which the user can select a name. Three groups of functional buttons are placed on the right.



The first represents the pre-established horizontal relations among the main entities of the resource called AUTHORS, WORKS, ORGANIZATIONS. [Fig.2] The second group represents a very useful interaction tool, in the form of an editable notebook, that allows the user to take notes or add his or her own knowledge concerning the investigated subject. The third one contains the HELP and EXIT functions. Finally, on the lower right side, there is a counterpart of the selection tool, a bi-directional browser, marked by arrows. The information sheet, the core of the information scene, occupies

the greatest part of the window space. The window can be shrunk, moved, or closed.

The catalog-like information on AUTHORS, accompanied by a small color or black-and-white photo of the artist, is divided into variable length fields or chapters, labeled as follows: Biography, Type of arts, Collaboration with galleries, Works are kept in, Take part in exhibitions and auctions, Author about himself, Critics, Bibliography, The author's manager. It takes only a short while for the user to find out that the formal label names or headings, in the English translation, are similar to those in other resources or media.

Browsing through the list of artists' names and the connected information raises the obvious question concerning the selection criteria. Why have these particular artists been selected for this CD-ROM and what do they have in common? According to our knowledge, there are about 20,000 contemporary artists in the Russian Federation who deserve, more or less, a wider public's attention. It is also well known that in Eastern Europe, exhibition history has a different connotation than in the Western world. Exhibitions, in this part of the world, were organized by central bodies, which imposed a certain orientation, taste, and selection rules. Ninety percent of the acquisitions were made by the state. Maybe it would have been clearer to have two separate fields or chapters, one for exhibitions, one for auctions. In this way, had the resource been focused on information and documentation, the exhibition part would have been more developed. On the contrary, had the resource proved to be of a more commercial interest, then the auction part would have delivered more information concerning this area.

In the AUTHORS space, interaction with the user is achieved only in two points. In the first, the artist — under the heading of Author about himself — addresses the user directly (e.g., KIZE-WALTER Georgy D., painter, confesses to his audience "... Balancing on facets of different layers of civilization, at points

where irony, spirituality and kitsch came together, and sometimes combining the incompatible, I seek to bring home to the viewer the meanings which are engendered in this way ...”, ORLOV Igor M., painter, explains: “... My faith is in the integrity of the world, the inter-connection of all, even the smallest particles in the universe. Man is a part of the universe. Space is flowing into me just as I am flowing into space. I hope that this sense of interconnection is reflected in my works. I make efforts to see links between each tiny particle and the universe. One can see a drop of water in the ocean and the ocean in a drop of water”) Then, through the double-click on term from The type of art field, the CD-ROM developers show that they, finally, considered a possibility of classifying and organizing the artists and their works, enabling the user to select or group the artists that conform to the type of the objects or genre: animation, art ceramics, art glass, assemblage, book designing, collage, design, graphics, installation, monumental art, mosaic, painting, performance, photography, sculpture, stage design, tapestry.

The WORKS section exhibits a selection of one to nine color or black and white thumbnail images of the works (the editors do not explain when and why they have included only one work or when and why they made room for nine). Unfortunately, the



displays have borders and heavy shading. [Fig. 3] The text information is totally eliminated and replaced by visual objects. There is no hint on selection criteria, ordering principle, or scale. The

user cannot find any explanations why and who decided upon the selection.

Browsing through the works, one can find everything, starting with naive art and socialist realism to post modernism, a.s.o. It reminds us, to a certain extent, of the Official Salons, we, in Eastern Europe would organize where works are put together mainly to conform to an external pattern. The eclectic variety might be a powerful experience, but again, there is a need for an interactive answer to explain this type of approach or to provide the user with a tool to accomplish this task on his own. Under these circumstances, what the user can do is to click on one of the works and open a new window which displays a magnified image of the work supplemented by a (sometimes too) short text containing the basic elements of a catalog description: author, title, date, dimensions, material. In some cases a detail (FRAGMENT) and MORE INFORMATION functions are implemented.

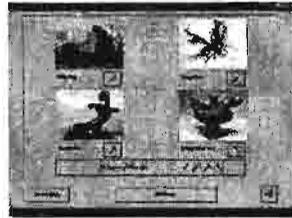
The ORGANIZATIONS entity connects the artists to the galleries offering text information on ten galleries from the Russian Federation. [Fig. 4]



The second CD-ROM, published in 1995, contains some technical and aesthetic innovation. From the first screen, the growing convergence of entertainment, media and communication technologies is obvious. Moving images, combined with electro-acoustic music written for this CD-ROM by one of the important Russian contemporary composers, Eduard Artmeyer, guide you into the resource's world. [Fig. 5]

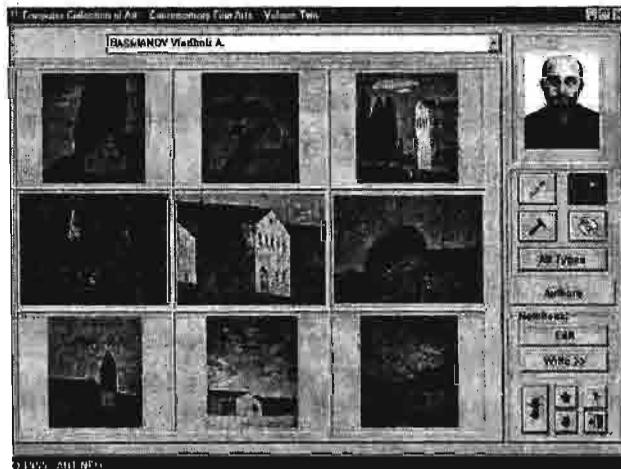


Problems with classifying the artwork as well as access and navigation rules have been overcome in the second volume of the Contemporary Fine Arts disks. In the opening screen the user can define the domain of art he/she is interested in: All Types of Art, Painting, Graphics, Sculpture, or Decorative Arts [Fig. 6]. The functions are icon driven: a click takes you directly to the works of the first artist of the selected domain. The artwork becomes more important and is placed in the path. The catalog-raisonnee-like first volume acquires added cultural and cognitive value, although the user loses some of the navigation freedom.



The images are placed on the screen in a simpler and more direct way, to be chosen by user. The importance of the slide quality becomes very obvious which, again, invokes information about the photographer, the source of the image and other related information. The screen is organized in the same way as in the first volume: the pull down choice list with the artist names, the buttons and the content information. A new functionality is added: a bi-directional continuous browser through the works which is a very useful tool for presentation or lectures. But still, the order is fixed, no rearrangement of the material is possible [Fig. 7].

Rearrangement is an essential functionality in a multimedia environment, enabling the user to define his own views or interest fields, providing



him with a creative and interactive mechanism. Imagine the number of possible combinations and connections among 1,000 works ! This is what might become the vital force of a the new media and its appeal, a new way to invite users to come back again and again, or to put it in more technical words, its reusability potential.

Once a work is selected, catalog information is available. A detailed viewing of the work of art is offered. The user is given the opportunity to view the authors' works sequentially or in full-screen mode. A magnifying glass zooms into a selected fragment of the work which enables the user to study the details or to look for the signature information.

The WORKS entity can be extended with information on the artist by choosing AUTHORS. Information is delivered and organized similarly in both volumes, except that the label names are slightly changed in volume 2. "Biography" becomes "Biographical Information," "Works are kept in" becomes "Collections Where Works are Held," etc.

The Organization entity is merged into the other information sources. Contact information is still present, reinforcing the main purpose of the resource: to use new technologies to distribute information about artists' works, representing a variety of schools and genres.

Before leaving the Computer Collection of Art, the user notices another button called ArtShow. This is Artinfo's surprise: a series of mini-art shows: Presentiment, DIXI, Mirage, W. Morris Dream, constructed by composers who were invited to pick out paintings they liked and to compose music for them. For example, W. Morris Dream is an art show translating into pixel the oil strokes of David Sherenberg's sun-scattered beaches and green fields in "A road outside Moscow," "Sea," and "A highway through the Forest"; the refined calls of Serghei Chernyshev's birds, briars, and thistles in "Morning Twitter"; Vladimir Basmanov "Night Clouds"; and

Natalia Muradova's light-and-shadow games in "Wings for Flight" and "Construction Site Light." The accompanying music is by virtuoso saxophone player and famous composer Alexei Kozlov - one of the most brilliant jazz players of the 1960s.

What the user would like to have, in the coming volume or a different series, is a method enabling him/her to go deeper into the multimedia world and to experience, to build by himself an art show to choose music and images and combine them. What about a virtual exhibition or contest on the outcome and a facility to choose images and pieces of music and build up art shows by himself?

There are discussions going on all over the world about the way a CD-ROM or a multimediated product should interact with the its users or readers or players or whatever we may call ourselves. Concerning the Computer Collection of Art on Contemporary Fine Arts published by ITAR-TASS and ARTINFO, one thing is sure: its content has a powerful impact and initiates a global dialogue — among artists, critics, galleries — distributing information, promoting artists and encouraging art events through more innovative channels.

CONFERENCES

.....

COALITION FOR NETWORKED INFORMATION

Spring 1996 Meeting

[Editor's note: This report borrows liberally from the official CNI report but contains my own observations as well.]

The Spring 1996 meeting of CNI was held on March 25-26 in Washington, D.C. Paul Evan Peters, Executive Director, Coalition for Networked Information, opened the meeting noting that it marked CNI's sixth birthday. He asked what CNI set out to do and what it has achieved. He answered that CNI's initial strategy was to demonstrate best practices and generally encourage innovation and transformation using networks and networked information, while its newer strategy is to focus on resolving specific barriers that are perceived to be reducing the impact of new information in scholarship. Peters issued an invitation to attendees to address how CNI operates as well as the content of its activities and then announced the meeting theme — "the Networked Information User," a community that is obviously growing in size and diversity. He emphasized the need for strategies to reach elements of the population, for methods to assess competing claims about the net, and to determine how to anticipate change.

Four plenary speakers then addressed the theme for the meeting. Kenneth C. Green, who has produced an annual survey of campus computing since 1990, described underlying factors that affect use of technology in higher education: changing student demography, competition for state money, aging faculty, questions about quality from authorities, and infrastructure (e.g., buildings and technology). Green noted that our aspirations for

technology in the 20th century have led to great expectations but that our ability to deliver remains quite small in relation to expectations. Students coming to campuses often have better technology in their homes than they find in laboratories on campus.

Green's research shows that campuses invest in information technology because of:

- market expectations and competitive position
- expectations for curriculum enhancement, and
- labor market preparation (e.g., preparing students for an information economy)

Technology can be used to address in various ways the three components of the instructional mission of higher education institutions:

- Content (e.g., structure of syllabus and transfer of content from teacher and library to student)
- Context (e.g., learning environment, campus, resources, socialization, time and place)
- Certification (e.g., course sequencing, program, degree, skills, licensing, and outcomes)

As the higher education market becomes segmented between adults and traditional (adolescent) students, the emphasis on the importance of each of the three components will shift. In addition, Green observed that there is a tension between faculty on the one hand and provosts and presidents on the other regarding the meaning of "productivity." Faculty generally think that technology makes them more productive in the sense of an improvement of quality. Presidents and provosts in general are focused on improvements in costs.

Green believes that the campus technology infrastructure drives innovation — if it's not out there, faculty rarely will push for more. Infrastructure encompasses campus networks, libraries, bookstores, and copy centers, off-campus network access, smart cards and metermaids, and user support. He closed by describing "Five Not So Easy Issues" that higher education must address:

- making copyright work
- differential access to the network
- sound planning for amortization of equipment
- infrastructure and support services, and
- classroom and instructional integration of technology.

Karen Hitchcock, Interim President, University of Albany, State University of New York, described the higher education cultural transformations that are needed in relation to new technologies. She noted that innovations in curriculum and pedagogy have not been embraced at the core of the enterprise. The use of networks and networked information resources has had profound effects on research and student services, but progress in transforming the learning environment is not very apparent or widespread. She observed that while it is assuredly true that the university must develop infrastructure strategies, human resources investments such as a faculty development program will be essential to assure that the promises of the infrastructure are to be realized. As students become co-discoverers with faculty, technology will help create a new infrastructure that will facilitate exploration and in which students can become part of a research community. Hitchcock closed her remarks with a number of recommendations for change. She stated that major changes are needed in the institutional culture, particularly in regards to the role of faculty, the relationship of faculty to information specialists, and the use of information resources. Despite the many unknowns, we must allocate resources to support the changing learning environment

on campus which involves new faculty roles and new uses of information resources and technology.

John Quarterman, Editor of Matrix Information and Directory Services, gave a rambling, overlong, and totally flaky talk in which he demonstrated his capture of data illustrating Internet use during snapshots taken every four hours. He posts this information on the Internet at URL: <http://www.mids.org>. Unfortunately his depiction of the network population provided very little of value to those trying to understand its subtle composition.

The final plenary was given by Gary Puckrein, President of the American Visions Society and a leader in African-American use of the Internet. Quoting 18th century thinkers on the issue of whether the “people” perceive the government to be legitimate and worthy of defense, he asserted that a feeling of ownership and a sense of belonging will be critical if African-Americans are to be enfranchised on the Net. His interesting advocacy talk basically argued that the Clinton administration’s faith in the private sector and the end of monopoly control over telecommunications will not result in access by African-Americans or other low income people. Unfortunately Puckrein did not seem to me to bring any new strategies to the table, although his reminder that the connection between economic prosperity and social stability is important and must be addressed.

Following the plenary, I attended a breakout session on the Columbia University Online Books Evaluation Project, presented by Carol Mandel, David Millman, and Mary Summerfield, Columbia. It was particularly interesting as an example of how difficult it is to segregate the deep infrastructure costs of electronic campus information services from the product costs. This issue has also confronted the Museum Educational Site Licensing Project in its efforts to understand the real cost of providing access to museum data on college campuses.

Subsequently I heard Nancy Kaplan and Stuart Woford of the University of Baltimore explain what they felt was a problem for scholars of linking their network distributed texts to texts by others that have not yet been written but which will reference their text after it has circulated. To solve this, they developed a software tool, “Citescape,” which is a combination spider and WWW server software to store and display all the URLs containing links to individual nodes and aggregate nodes for a target text. The “out cites” function creates links to the document that include “link-citations” to the aggregate of the target. The “in cites” function links to those documents that include “link-citations” to an individual node. My mind went numb thinking of the network-wide implications of the proliferation of agents like these, but then I decided they were no worse than, or really different from, other kinds of gatherers.

In the final breakout of the afternoon, David Green, the newly appointed Executive Director of the National Initiative for a Networked Cultural Heritage introduced NINCH and Charles Henry (Vassar College) discussed the American Arts and Letters Network. Both these initiatives have been described in previous issues of Archives and Museum Informatics.

Tuesday morning opened with two plenary speakers who provided somewhat contrasting views of the networked intellectual property landscape. Robert Weber, Senior Vice President, Electronic Publishing Resources, Inc., asked whether rights management is an opportunity for higher education. He noted that technologies are coming to market which allow charging and monitoring — these allow electronic value chains; constitute a secure system for electronic commerce; and permit customers and providers to negotiate an appropriate degree of privacy and confidentiality. They will support familiar pricing models such as the subscription, but will also facilitate transaction pricing such as “pay per use.” Weber described this capability as an opportunity for producers and consumers to have the best of both worlds:

transaction pricing and subscription pricing. The availability of actual data on usage, which these systems will provide, will assist librarians and publishers in understanding usage and coming to terms with appropriate pricing strategies. Advanced rights management capabilities will offer persistent rights protection for all kinds of digital properties making copyright compliance easy, which is an important incentive to creators and publishers.

Weber described technical strategies that are being developed for rights management systems. He described a system which employs encapsulated objects (containers) with metadata levels to control access (parameters as an expiration date or amount of time that an item can be used), use (limiting the ways in which an item can be used), and prices (with variations by type of user). He noted that several companies, including his own, will offer these technologies in the next few months and he expects that they will evolve rapidly over the next two to five years. Since the talk, of course, IBM InfoMarket has gone commercial with its product; EPR's is expected in the fourth quarter.

To give the new concept flesh, Weber sketched a scenario of how it might work. A content creator writes an article, marks it in HTML, and uses free encapsulation/encryption software to place it in a "container." He specifies usage rules (e.g., how much someone must pay when he/she reads the article), stipulates conditions (e.g., the user may not print the article), and specifies distribution rules (e.g., the publisher may or may not add a markup). The creator could let the object out onto the network at this point, or provide it to a publisher who could add other usage rules, combine the article with other value (e.g., a photograph), specify distribution rules, and place the modified "container" on his/her Website. A user looks at a publisher's Website, selects the illustrated article and downloads the article without any restriction on recopying it to others. However, before that user, or any other user, can read the article, they need to open it, and at this point an applet contacts the containers' clearing agency, checks the terms

and conditions, bills the user's budget or credit card, and sends the container the key for use. The user then opens and views the item. The financial clearinghouse disaggregates the transaction and sends money due to the creator(s) and publisher.

In closing, Weber asked rhetorically "What happens to copyright law?" He expects that copyright law per se will become less important and will be used mostly as a last resort to go after those who infringe. The focus will be on developing agreements between providers and users on the use of content. Another critical change he anticipates is that the permissioning process will evolve from one of explicit permissions to a system of pre-approved permissions, which will be detailed in an online database. He stated that he feels that "fair use" will decline in significance. In answer to the question posed in the title of his talk, how higher education will be affected, he stated that libraries have been concerned about the fact that much knowledge produced by faculty is given to publishers and then sold back to higher education. These new systems, he suggested, create an opportunity for creative rethinking of pricing structures. These technologies can reinforce the opportunity of self-publishing communities in higher education.

I found Weber's presentation exciting on numerous counts. First, persistent control makes copyright and contract compliance easy, but it is an important incentive not just to publishers and creators (although it is that) but also a critical element in the archival equation. If you don't give anyone the rights to alter an object, it becomes "inviolable." Archivists might also note that the new marketing paradigm which emerges, sometimes called "superdistribution," makes it possible to place the contents of an archive into a "market" and reap benefits from its increased use, thus enabling archives to invest in access! Secondly, I was pleased to find that Weber's independently developed "containers" were instantiations of my "Reference Model for Business Acceptable Communications" and included provisions for the handle layer,

terms and conditions layer and use history layer, in addition to content. For archives this means that only the metadata requirements of the structural and contextual layers of the BAC would need to be introduced into these architectures to ensure that records with evidential value were created and could be maintained across time. The likelihood that a widespread network software infrastructure will be in place to support metadata-encapsulated object-based methods of electronic records management is the best news on the electronic recordkeeping front in a decade.

Weber was followed by Peter Jaszi, Professor of Law, Washington College of Law, American University, and a key organizer of the Digital Futures Coalition. Responding to Weber, Jaszi felt the need to make a claim for the continuing relevance of copyright in “structuring markets and information.” He described the underlying principles of the Statute of Anne, a British law of 1710 on which much of our modern copyright system is based, and which balances the rights of authors and publishers with the rights of society. While publishers gained an easily enforceable publisher’s rights in the statute, they also were subjected to significant limitations, such as the limited duration of copyright and the requirement for deposit of copies in library collections for the purpose of public access. Over the years, additional rights have been secured for the interests of society, including the originality requirement, the “first sale” doctrine, and “fair use.”

Jaszi then discussed the current legislative proposals that have emerged from activities of the Information Infrastructure Task Force (IITF) of the Clinton Administration. Jaszi stated his view that the narrative of the intellectual property white paper which purports to describe current copyright law as applied to the networked environment slants its interpretation towards the interests of information proprietors and that while Congress is debating new copyright legislation (so slowly it now seems unlikely to pass in this session of Congress) the real action has moved to Geneva. Jaszi warned that whatever the fate of the domestic legislation, the

U.S. may be party to one or more international agreements before year end which will compel certain obligations domestically.

Jaszi described three key elements of the current legislation and these international initiatives as:

- (1) the articulation of a transmission right, within the “distribution right” clause, which stipulates that digital transmission of a protected work without permission is potentially infringing copyright;
- (2) the addition of anti-circumvention provisions, to be a new chapter 12 of the copyright act, which prohibit the manufacture and distribution of technologies designed to protect works in digital form (the problem according to Jaszi is that some of these technologies may have significant benefits too); and
- (3) the protection of copyright management information protection (Jaszi’s problem being both that the task of defining the content to be protected is delegated to the Copyright Office rather than being enacted into law and that the language is too broad).

In addition, Jaszi was concerned by the absence of reference to some matters in the new legislation, including:

- (1) The fate of first sale doctrine is not discussed as it was in the white paper.

- (2) The marginalization of fair use goes unaddressed. In particular, Jaszi thinks that fair use has had an economic benefit in that transaction costs have been too high to support licensing in the past. This is probably true, as is the argument that opportunistic uses are good for education, but neither argument to my mind stands up in the new technological context. More powerfully, Jaszi argues that fair use has also had a cultural use in protecting use of works in ways that the creators might not approve of, including transformative uses, but as far as I know, no one has suggested that the new copyright provisions would impair satirical and other critical uses.
- (3) There is no reference to service provider liability which is currently holding up the act in the House. Of course this is true, but Jaszi agrees that having service providers enforce copyright would have a negative impact, so it's unclear why the reference needs to be made.
- (4) Finally, there is nothing in the act providing an educational exemption on transmission rights. If there were, of course, and it was not simply an extension of section 110 that protects distance education now, it would be a new right, which Jaszi says he does not favor.

Jaszi stated that the proposed legislation gives us a vision of a market where information is "locked down" and information owners can relax controls to provide access to those customers they wish to serve and in which every transaction in the system can be observed and monitored. He acknowledged that this vision of the new information environment is OK if we trust information producers and distributors, but if not there is a lot of monitoring and privacy conflict in the rules for this information market. He asked, "Are we comfortable with giving up on the traditional balance?" and urged further discourse around the topic of what kind of system users and creators wish to enjoy.

Needless to say, I found Jaszi's temperate and measured defense of traditional copyright a nicer act than the more strident ones I often hear from academics who believe that fair use is a right rather than a defense against infringement, but the bottom line is the same. I continue to believe that copyright can either adjust or become irrelevant. Technologies such as those described by Weber could hasten the arrival of irrelevancy if the user community does not accept that the electronic environment requires a rebalancing of the equations to protect property more.

I followed the plenary with a briefing by Jeff Crigler, an IBM Vice President, on its new InfoMarket service (see site at www.infomarket.ibm.com). InfoMarket is a commercial implementation of the kind of architecture described by Robert Weber (who was a consultant in its development). It can be seen as a content distribution utility. Recognizing that publishers need to get material found and to get paid by users, it serves as a way that they can disseminate their content broadly and receive financial rewards through a clearinghouse. IBM has signed on search engines like PLS, Verity, Netscape, Folio, OpenText, Notes, and Fulcrum, as well as a larger number of major content vendors. They see themselves as a facility, like a stock exchange, not an online provider. Ultimately this means they will need to make the "containerizing" capability available to anyone. Their plans are to do so this summer. In addition they favor open standards in this area, and are currently using very open concepts such as ZIP to ship their "cryptalopes" because ZIP is widespread, freely available technology

The most interesting aspect of the discussion with Crigler was his speculations on the implications of the new distribution paradigm for business models. When he promised that IBM would work with publishers to determine appropriate pricing models, I couldn't help but think how exciting the re-invention of lots of publishing would be in this environment.

The next breakout I attended fit neatly into these thoughts. It was chaired by Robert Ubell who has been heading an effort by CNI to develop cost measures of the current publishing and information consumption market. Ubell and Associates presented their most recent findings, based on multi-day focus group meetings with publishers, buyers, and intermediaries, to an audience in which everyone fit one of those categories. There was spirited discussion of the categories and of possible new categories of costs which will no doubt be reflected in future drafts.

I was unable to attend several other project briefings of interest to readers, but reports are available at www.cni.org:

- (1) CIMI. "Consortium for the Computer Interchange of Museum Information (CIMI) Cultural Heritage Information Online (CHIO) Project: Update on Z39.50 Application Profile for Cultural Heritage and SGML DTD for Museum Exhibition Catalogs," by John Perkins, Computer Interchange of Museum Information; Ray Denenberg, Library of Congress; Robin Dowden, National Gallery of Art; and Steve Dietz, National Museum of American Art
- (2) MESL. "MESL Project Description," by Jennifer Trant, Getty Art History Information Program; Steve Dietz, National Museum of American Art; Gregory Welsh, American University; David Millman, Columbia University; and Howard Besser, University of Michigan.
- (3) "Planning for Digital Archives," by Ronald Larsen, University of Maryland at College Park; and Peter Hirtle, National Archives and Records Administration.

CALL FOR PAPERS:

DLM FORUM (Electronic Records) 19th-20th December 1996; Brussels, Belgium

A multidisciplinary Forum on Electronic Records will be held 19-20 December 1996 at the Borschette Centre in Brussels, Belgium. The Forum will be hosted by the European Union Member States and the European Commission (Secretariat General, DG XII Science, Research and Development). Participating specialists and executives are coming from the Member States and abroad. Actors involved in the information Flow and Electronic Records Management of their departments/services/organizations will bring expertise and debate on possibilities for wider cooperation in the framework of the European Union concerning management, storage, conservation, and retrieval of Electronic Records.

The objective of the forum is to prepare a set of guidelines for best practice on Electronic Records and to investigate possibilities for wider cooperation between the various parties involved with Electronic Records, between Member States and at the Community level.

The Forum will consist of plenary sessions and parallel working party meetings. The main topics include:

- Information Flow/Work Flow
- Life-cycle of Electronic Records
- Relationship between creators, users, and custodians of information

Terms and concepts

- Cooperation Europe-wide

Please submit your papers to:

Mrs. Christina Beckers, European Commission SG 1-AH (Office : SDME 5/72), 200, rue de la Loi B-1049 Brussels BELGIUM. Tel: +32-2-296 08 2, +32-2-296 26 47; Fax: +32-2-296 10 95; Email: dim-forum@sg.cec.be

CALENDAR

- **June 17-22** Boston, MA Ed-Media & Ed-Telecom 96: World Conferences on Educational Multimedia and Hypermedia and Educational Telecommunications [804-973-3987; fax 804-978-7449; AACE@virginia.edu]
- **July 17-20** Washington, DC National Association of Government Archivists and Records Administrators [518-463-8644; fax 518-463-8586]
- **July 22-26** London, England Electronic Visual Arts: EVA London [VASARI +44 1252 812 252; fax +44 1252 815 702; jameshemseley@cix.compulink.co.uk]
- **August 21-23** Arlington, VA Interactive Multimedia 96 [800-457-6812; fax 540-349-3169; salt_conference_info@lti.org]
- **August 26-31** San Diego, CA San Diego Society of American Archivists [312-922-0140; fax 312-347-1452; info@saa.mhs.com-puter.com]
- **Oct 30-Nov 2** Ottawa, Canada Canada Museum Computer Network [301-585-4413; fax 301-495-0810; mdevine@cni.org]

INBOX

.....

Reports

- * Canada, Information Highway Advisory Council of Canada, Final Report. Connection, Community, Content: The Challenge of the Information Highway (Ottawa, September 1995). Also <http://info.ic.gc.ca/info-highway/ih.html>.

The Canadian Advisory Council on the NII reported earlier than its American counterpart (see vol. 9, pp. 446-448), but struck many of the same notes. I regret I didn't have both reports in hand last issue to compare them then. Not surprisingly, the Canadian report emphasizes the importance of protection from cultural imperialism and the need to develop distinctly Canadian content more than its counterpart in the US. On the other hand, openness and promotion of new endeavors was the leit-motif of most of the recommendations and except for

some nuance (more harmonization and balancing) the two reports are complementary.

- * Kenneth C. Green, The Campus Computing Project, Campus Computing 1995: The Sixth National Survey of Desktop Computing in Higher Education (Claremont, CA: March 1996). 32 pp.

There are a few surprises and lots of confirmation of trends in this year's (Fall 1995) survey of campus computing. Overall use of information technology on campus grew dramatically: e-mail leapt from 8 to 20 percent penetration, classroom computing from 16 to 24 percent, and the use of multimedia doubled from 4 to 8 percent. Data suggests that over 7 million students and faculty regularly used the Internet and WWW. Despite this growth however, two countervailing trends are equally evident: universities are not investing a greater proportion of their resources in instruc-

tional materials development and few universities have developed long-term equipment replacement budgets to cope with the ever-growing costs of infrastructure support.

Among the statistics I found surprising were those indicating that public universities are more than three times as likely to charge students a computer use fee than private ones (47.5%:14.3%), that the ratio of students to computers in universities was roughly 3:1; that over 60 percent of faculty have computers; and that more than 20 percent of universities still do not have a written policy on software copyright/piracy (although many have one under development). Approximately 60 percent of the universities reported faculty support for instructional courseware development, while less than one-third supported development of research software. Oddly, 63 percent of public universities had a royalty sharing program for faculty developed courseware while only 27 percent of the private universities had such policies. Very, very few institutions had a for-

mal plan for using the Internet and WWW in instruction or distance education (although the majority has such plans for self-advertisement). Despite strong pressures to reduce costs, over 90 percent of institutions reported having decided NOT to outsource academic computing operations or academic computing itself.

Some of the confusion surrounding campus computing comes through in the assessment by personnel responsible for it of the most important strategic issues over the next two to three years in which they ranked "clarifying goals and campus plans for technology resources" and "expanding computer networking across the campus" as most important, followed by "developing budget mechanisms to replace aging equipment on a routine basis." It went almost without saying that network issues were ranked as "more important this year than last year" by 78.1 percent of respondents, with 21.3 percent estimating them as the same and only .06

percent indicating they were less important."

Technically, I found it interesting both that 87 percent of public and 95 percent of private universities had a fiber optic backbone and that none reported using ISDN. Of the academic and administrative units, it was noteworthy that the fine and performing arts had more access to the backbone than engineering, but substantially less than humanities or social sciences. In universities, however (as opposed to four- or two-year colleges), the penetration was approaching 100 percent. In general, computing staff rated the technology skills imparted to students above the preparation given to faculty for instruction, research, or information seeking. The Internet and World Wide Web were almost universally seen as an important source of content and instructional resources over the next 2-3 years.

* Getty Art History Information Program, Research Agenda for Networked Cul-

tural Heritage (Getty Trust, Santa Monica, 1996) 80p.

This is the report of a series of online and off-line discussions growing out of the perceived need to define just what it is about humanities and arts computing that requires special research funding and support. The project was first conceived in the summer of 1994 after Getty staff and I completed a crash research effort to find arguments that would sway the NII Task Force and make the case for culture on the Internet. One of the issues that kept arising was whether arts and humanities computing actually had distinct research needs that would not be satisfied by progress in scientific computing. In the spring of 1995, we invited experts to nominate fields of computing expertise that were especially germane to the digitization of culture and subsequently invited the nominees to describe the state of the art in their domains and identify the most significant areas for future research. This report contains their papers and the discussion which their papers

engendered in a series of hosted Internet-based conferences over six months in 1995. As the author of one special interest paper and the overview, and the host of the conferences, my view of the contents is biased, but I would recommend reading the report for the nuance it gives to the question of "what exactly is humanities computing?"

* Charles R. McClure and Cynthia L. Lopata. *Assessing the Academic Networked Environment: Strategies and Options* (Washington, DC: Coalition for Networked Information, February 1996). \$15.00. Appendices include: Linda H. Fleit, "Self-Assessment for Campus Information Technology Services," CAUSE Professional Paper Series #12, and the CAUSE/EDUCOM "Evaluation Guidelines for Institutional Information Resources."

This report is a relatively comprehensive effort to define how to document the campus

information network infrastructure. The network infrastructure, as defined by the authors, encompasses the information and media services, products, hardware, software, resources, and support in place in the academic institution. They set out to frame not only the basic questions about this networked environment (Who uses what, how much, and in the context of what activity? What does it cost?), but also to explore how to tease out the much more difficult qualitative issues (How has access to these resources changed academia? what benefits and drawbacks are there?). One of the useful things about the report is the range of experiences it attempts to portray. Within institutions, the different experience of the networked environment for different users is examined through measures from network benchmarking, focus groups, critical incident analysis, user activity logs, interviews, group processes, site visits, observation, and scenario development. The next test will be whether any institution (or more importantly institutions) can afford to

conduct such an aggressive set of evaluations. If not, presumably the next reports should help us determine which methods of evaluation are most effective for what purposes.

While the primary consumers of this report will probably be university administrators and sociologists of the new electronic culture, I would suggest that senior managers in all institutions in our society would be well served by a careful reading because their institutions will enter the deeply networked environment soon. Understanding the kinds of measures and different types of measuring tools proposed by the authors will help them tune their evaluation plans and could result in very significant savings over having to invent such measures themselves.

* Robert Ubell Associates, *Cost Centers and Measures in the Networked Information Value Chain*. Expert Panel Sessions: Publishers, Intermediaries, and Buyers (loosebound, CNI Meeting, March 25, 1996). 71 pp.

This draft report of a project to define cost elements in electronic information dissemination summarizes the findings of a series of three-day expert panels which met in the summer of 1995. Experts were asked to identify the cost elements independently of whether they would be charged for or what implications they had for quality or content. In the process of debating what their costs were, and how they were incurred in the system, experts developed models of the system and defined many terms. Publishers generally felt that present and future modes of operation will co-exist in the future. Intermediaries saw themselves in niche markets with a great appetite for digital content. Buyers were most aware of the implications of different ways of acquiring data on their staff and on the organization of their collections.

The report concludes phase 1 of the project. In the next phase, reviews of the document will be solicited to validate and clarify the models and definitions. In July, a report on

value-added functions will be circulated to reviewers who will be asked to suggest ways to assess these costs and their importance. In the final phase of the project, actual costs will be associated with steps in the process to help design "a system of measurement capable of identifying how costs and functions will change over time." A report on cost-centers and measures will be distributed in October.

Books and Articles

C.J. Armstrong and R.R. Fenton, eds. **World Databases in the Humanities** (London: Bowker Saur, 1996).

At 1060 pages and \$250 this may be one of the last of the great dinosaurs — a comprehensive reference compendia for libraries. Given its odd coverage (I failed to find any of the three known image database and multimedia publishing services), its format, and its rapidly dated material, we can expect this volume to go with the dinosaurs eventually. Why publishers haven't learned that publishing databases (espe-

cially databases of databases) in paper formats is silly, I'll never understand.

Jean Dryden. "Archival Description of Electronic Records: An Examination of Current Practices," *Archivaria* 40 (Fall 1995):99-108.

Alas, the picture of rank confusion drawn by Jean Dryden is accurate. Hopefully the final version of the RAD Chapter 9 and my "Metadata Requirements for Evidence," published since she wrote her piece, will contribute in part to focusing the image.

John Feather, Graham Matthews, and Paul Eden. **Preservation Management: Policies and Practices in British Libraries** (Brookfield, VT: Gower, 1996).

I asked to review this work to determine the impact of the latest studies in digitization on preservation practices in libraries. Since the authors address the topic in one page, use literature published between 1992 and 1994 (reporting work completed only as late as 1992), there were warnings, promises,

and little impact. Elsewhere, digitization and access are discussed somewhat more favorably but the literature is largely from the 1980s and does not take into account developments on the Internet, to say nothing of the explosion of the WWW. Sometimes the book as a medium is just too slow.

Chris Hurley. "Ambient Functions — Abandoned Children to Zoos," *Archivaria* 40 (Fall 1995):21-39.

Hurley once again locates an important issue, faces it directly, and arrives at an unusual conclusion. In what are now becoming a remarkable series of discussions with himself on knowledge representation of archives, he has invented yet another, very useful, concept in "ambiance" and given it a location and a utility. No one thinking seriously about information as a means of control in archives can ignore this.

Jane Lusaka, Susannah Cassidy O'Donnell and John Strand. "Whose 800 lb. Gorilla Is It?: Corbis Corporation Pur-

sues Museums," *Museums News* 75:3, p.34-.

The authors seem to think the rhetorical question has two answers, but despite the nice face they put on the creature, there can be little doubt that Corbis is not in the best interests of museums, though it may come to vend access to some digital images from some of them. More promising are investigations currently underway in AAMD and AAM to provide a cooperative licensing venue for digital rights owned by and for museums.

Rick Prelinger. "Archives: Winners or Losers in the Digital Revolution?," *AMIA Newsletter* 32 (Spring 1996):19-22.

The co-chair of the Access Committee of the Association of Motion Image Archivists has intriguing observations to make about the interest in archives as sources of content that has suddenly taken hold of the commercial sector and the new position that archives find themselves in between technology and the market.

Glyn Sutcliffe. **Slide Collection Management in Libraries and Information Units** (Brookfield, VT: Gower, 1995).

Despite its over 200 pages of text, this book is a bare introduction to the most basic aspects of how-to-do-it slide librarianship without any analysis of the methods employed, the larger context of information retrieval into which they fit, or the economic and technical consequences of these methods. The sections which address databases for information retrieval and optical disks are too superficial to be useful and full of dated (and rapidly aging) kinds of information rather than useful guidance on the actual technical parameters of image database management.

Journals and Newsletters

The April issue of *D-Lib Magazine* is now available at <http://www.dlib.org>; the UK Office for Library and Information Networking maintains a mirror site for *D-Lib Maga-*

zine at <http://uk-oln.bath.ac.uk/dlib/magazine.html>. This month, it features a series of stories and briefings on current digital library issues and research, including electronic commerce, legacy data and systems, and work at the Library of Congress. D-Lib has also developed a mailing list, which we are using to announce the issue each month. If you wish to sign up for the list, send a message to dlib@cnri.reston.va.us.

D-Lib is sponsored by DARPA on behalf of the Information Infrastructure Technology and Applications (IITA) Working Group of the High Performance Computing and Communications (HPCC) program, William Y. Arms, chair.

For further information, contact Amy Friedlander, Editor, D-Lib Magazine, Corporation for National Research Initiatives, 1895 Preston White Drive, Suite 100, Reston, VA 22091. 703-620-8990; fax 703-758-5913; reba@cnri.reston.va.us.

Document Management (March/April 1996) contains a section (pp. 17-32) on workflow-based systems which makes it evident that workflow solutions have entered the mainstream. All the discussions archivists have been having about labeling transactions within a business process rather than opening software applications are now supported by a range of tools. Developments in workflow standards and offerings of different firms are discussed.

Multimedia Monitor on CD-ROM [Future Systems, P.O. Box 26, Falls Church VA 22040-0026; 703-241-1799; fax 703-532-0529] is \$99.95 (\$49.95 for subscribers) and covers all issues from September 1983 to December 1995. Mine are on the shelf, but still I'm tempted to buy the CD as the best industry reference source ever.

Online Publishing Report [O'Reilly & Associates Inc., 10-1 Morris St. Sebastopol, CA 95472; opr@ora.com; www.ora.com/newsletters/opr/; \$295 annually.

This monthly 8-page newsletter contains relatively technical discussions of developments in SGML, HTML, Unicode, Z39.50 URL's, WAIS, and other such topics of interest to online publishers. The reports seem current, but are anonymous so it's difficult to assess their likely authority. The cost, obviously, is high enough to dissuade casual interest.

Visual Resources, Volume 11 (3-4) is a special issue edited by Murtha Baca and Patricia Harping, devoted to the "Art Information Task Force: Categories for Description of Works of Art." The volume is divided into two parts: a series of analytical articles and the categories themselves. After an uncredited introduction to the Art Information Task Force and its work, Mary Case leads off with a discussion of the Categories for Description of Works of Art which sets the art historical "point-of-view" in context. Suzanne Folds McCullagh discusses the value of the Categories in making nuanced descriptions of works of art. I explore the relations be-

between entities in the Categories as a first step towards bringing our data models, including the CIDOC data model, into conformity with the Categories. (This spring the CIDOC Data Modeling working group took a step in this direction in deciding to recast its model in an object-oriented form that would emphasize relationships). Greg Tschann, in assessing the value of the Categories for systems development, reveals many of the weaknesses of the AITF point-of-view, including its failure to account for collections management data, the decision to articulate "categories" rather than data elements, and the inattention to retrieval issues. But Tschann argues for the benefits of the "Meta-Categories" developed to cluster the AITF categories; he believes these could be the basis for successful data interchange. In an essay that skims a variety of implementation issues, Dustin Wees illustrates the kinds of workaday issues that the Categories will not resolve, just as Brendan Cassidy, in his more extensive article on iconography details the kinds of intel-

lectual problems with description of art that no standard can address. The final article in the volume, before the section which reproduces the Categories themselves, is the edited transcript of a 1994 College Art Association conference session on the AITF which contains testimonials by numerous art scholars to the efficacy of the Categories in their work and Jennifer Trant's (useful) characterization of the Categories as a "bridging document" allowing scholars, systems developers, and information providers to speak in common terms.

Jim Wheeler. "Videotape Formats," *AMIA Newsletter* 31 (Winter 1996):24-26.

This brief technical note describes in a clear way the differences between analog and digital video formats and composite and component recording methods. It also identifies the various standards that have prevailed in all four kinds of recordings since the birth of video in 1956. A useful reference resource.

EPHEMERA

Isabella Stewart Gardner Museum. Rights and Reproductions Survey (November 1995). 44 pp.

During 1995, in preparation for revising their own Rights and Reproductions policies, the Isabella Stewart Gardner Museum conducted a survey of 126 museums rights and reproductions fees and policies. They received 109 responses from 47 states and summarized their findings in a report to the participants which they have since made available to others. At the present time this is the best picture we have of the situation of R&R in American museums and I believe it has some very interesting policy and planning implications for the profession.

The findings that museums distinguish between commercial, non-profit, and scholarly users, charging the highest fees to the former and the lowest to the latter, is no surprise. The degree of difference charged to different categories of users however is rarely very high and

averages less than 60 percent. Median charges for all categories of sales and rental seem unlikely to fully recover the costs of administering a rights and reproductions service. At the present time, without cooperation between museums to systematize their practices, the number of categories of types of uses employed by institutions and the number of different scales used to establish price are so great as to defy any reasonable effort to support one stop shopping through an intermediary. [For further information contact: Patrick McMahon, Registrar, ISG, 2 Palace Road, Boston MA 02115; 617-566-1401; fax 617-566-7653].

National Preservation Office, National Library of Australia. "Draft Statement of Principles on the Preservation of and Long-term Access to Australian Digital Objects."

This statement, developed at a workshop on Preservation of Digital Objects, December 6, 1995, and influenced by the U.S. Task Force on Archiving of Digital Information enumer-

ates seven principles to which I would assign the following shorthand descriptors:

- (1) Cooperation of all players
- (2) Creators' responsibility for on-going accessibility
- (3) Collective responsibility for custody
- (4) Retention for period of continuing value
- (5) Balancing rights of creators, providers and users
- (6) Adoption of best practices and standards
- (7) Appropriate laws and policy

It is not yet clear what one is supposed to do with such principles.

Australian Archives, "Using Electronic Mail" (February 1996). 12 pp.

This short brochure, designed to guide staff uses of e-mail within the Australian Archives, contains useful advice and concepts, but the most interesting part is the excerpt

on "Electronic Mail and Voice Mail" from the AA Documentation Standard being developed for all Commonwealth Agencies. It states that e-mail and voice mail are forms of communication and if used in course of business should create records as evidence. That some records are of ephemeral value and needn't be recorded in recordkeeping systems, but a business decision needs to be made about what records should be captured and how long they should be retained. Capturing records needs to be an automated systems function insofar as possible. E-mail and voice mail policies should make it clear that these travel through corporate channels and are not private.

NEWS

.....

Museum Associations Explore Rights Collectives

Both the American Association of Museums (AAM) and the Association of Art Museum Directors (AAMD) have taken steps in the last few months to authorize investigation of the possibility of establishing museum-owned and operated licensing collectives to provide one-stop shopping for images and interpretive texts, especially digital content, from museums. These feasibility studies are designed to answer questions about the practicality of such collectives as well as to build a case for museum participation in them. [For further information on AAM, contact Pat Williams, Vice President, Policy and Programs, 202-289-1818, and on AAMD contact Maxwell Anderson, Liaison for Information Technology, 416-979-6660.

National Digital Library Program

In February, the U.S. Copyright Office received its first digital application and deposit using the Copyright Office Electronic Registration, Recordation and Deposit (CORDS) system. The workings of the system were described in the NDLP newsletter #7 (April).

In March the NDLP released five new collections to the Library Internet site and launched an "On-Line Learning Page" for students and educators. In Mid-April, Ameritech Corporation gave the Library \$2 million to establish a competitive grant program through which libraries across the U.S. can apply to digitize their unique Americana collections. At the end of April, the Library of Congress announced that Reuters Foundation had provided a grant of \$1 million to digitize the papers of George Washington

and Thomas Jefferson and make them available online.

NARA-Nixon Agreement on Tapes

The Archivist of the United States, John W. Carlin, was pleased to announce an agreement designed to lift legal obstacles to the release of "Watergate"-related Nixon White House tape recordings this spring because it got this pesky issue off the agenda, but the public is likely to be much less happy with the result. Under the agreement tapes related to what the law calls "abuses of governmental power" during the Presidency of the late Richard M. Nixon could become publicly available within the current year. The agreement further provides a process through which additional Nixon tapes may be released, but it is neither an automatic nor an expeditious process. Nevertheless, the agreement which NARA reached with three other parties to litigation over the Nixon tapes — Public Citizen, History Professor Stanley Kutler, and the estate of Richard Nixon — resolves

litigation over release of the tapes that has been holding up any releases since 1993. With one exception, the agreement settles a lawsuit brought by Professor Kutler and Public Citizen to accelerate the release by the National Archives and Records Administration of approximately 3,700 hours of White House tapes recorded during the presidency of Richard Nixon. The agreement also allows the lifting of an injunction, granted by the court to former President Nixon, prohibiting tape releases by NARA.

The case was filed against NARA in March 1992 because only 63 hours of the White House tapes had been made available to the public since 1974, when Congress passed a law requiring review and release of Watergate materials and Nixon presidential historical materials. Former President Nixon intervened, arguing that NARA should give priority to returning all private conversations on the tapes to him. In August 1993, former President Nixon obtained a court order forbidding any further releases

until NARA had completed its review of all of the tape recordings and returned all private or personal materials. As a result, there have been no subsequent releases of any of the remaining tapes, including any of the 201 hours that NARA has concluded demonstrate Watergate "abuses of governmental power." Under the terms of the settlement, the parties have agreed to a schedule for processing of the remaining hours of the tapes and a process for objections by the Nixon estate and others whose voices are heard or whose names are mentioned. Regarding the abuses of governmental power materials, the Nixon estate has agreed to waive objections to specific items, and to allow the release to go forward unless it objects by October 1, 1996, and a panel of archivists from three Presidential libraries then concludes that the designation is "clearly inconsistent with the definition of the term 'abuses of governmental power'," as defined in applicable law and regulations. Unless the Nixon estate finds problems and the special panel agrees, or there are objections from other parties, the tapes

should be available as early as November 1996, followed by release no later than April 1998 of the Cabinet Room tapes, which cover about 278 hours before possible deletions for national security and other matters listed in the Presidential Recordings and Materials Preservation Act of 1974. The agreement also provides for release of the remaining 2,338 hours in five segments over the next several years. With each release of tapes, NARA will also make public the corresponding portion of its detailed, updated 27,000-page tape log and other finding aids that will help researchers and reporters locate conversations of particular interest. Unlike the earlier releases, for which transcripts had been made by the Watergate Special Prosecutor's Office, there are no transcripts. Former President Nixon had previously insisted that transcripts were required prior to any releases, but that demand has now been withdrawn by his estate.

Under the agreement, only one issue will remain in litigation: the question of the Archi-

vist's retention and maintenance of the original tape recordings in their entirety, including those segments deemed to be private or personal to former President Nixon, along with the master preservation copy. The Government believes that it is complying with the Presidential Recordings and Materials Preservation Act by retaining the original tapes and a preservation copy. Under existing law it cannot open to the public material on those tapes deemed personal and returnable. But to ensure protection of the Nixon estate's rights to keep that material private, the estate's representatives take the position, with which Public Citizen and Professor Kutler agree, that NARA is not legally entitled to retain the personal and returnable portions of the originals, nor a preservation copy, once reviews are completed and public material is released. All three parties have agreed to submit this issue to court resolution.

Researchers should soon be able to listen to lots of Nixon tapes, but I wouldn't bet that

the estate won't continue to dig in as much as it can and that the dubious privacy claims won't continue to restrict access to recordings made (after all) when Nixon was in the Oval Office, presumably acting as President.

RLG Joins Forces with Getty AHIP

On March 29, 1996, the Research Libraries Group, Inc. (RLG), and the Getty Art History Information Program (AHIP) announced a partnership that combines RLG's network infrastructure with AHIP's art research databases to foster broader information access and contribution by the international cultural heritage community.

AHIP initiatives have produced the Bibliography of the History of Art (BHA), the world's most comprehensive art-historical bibliography database; the Provenance Index, several files on the history of collecting and the provenance of individual works of art; and databases addressing the need for coordinated vo-

cabulary and information for cultural heritage documentation, including the Union List of Artist Names (ULAN) and Thesaurus of Geographic Names (TGN). Concretely, the partnership means that over the next year, RLG-AHIP working groups will plan and implement three key pilot projects under the new agreement:

- * *BHA*: Add the Bibliography of the History of Art to the set of RLG CitaDel (citations access and delivery) files—including information from BHA's predecessors, RILA (International Repertory of the Literature of Art) and the French RAA (Repertoire d'Art et d'Archeologie). Devise a method for contribution to BHA based on RLG models, involving RLG members—the Art and Architecture Group.
- * *Provenance*: Develop a plan for adding the Provenance Index files to RLG's RLIN (Research Libraries Information Network) and develop mechanisms to support contributions of

data by European members of AHIP's Provenance Documentation Collaborative.

- * *Vocabulary*: Resolve issues of vocabulary and authority file coordination across AHIP and RLG, with input from prospective users and contributors to a future vocabulary resource on RLIN.

[for further information see <http://www.rlg.org> and/or <http://www.ahip.getty.edu>]

Free Access to GPO Access

Since December 1995 all Internet users (as well as dial-in users) have been able to receive full text access to the Congressional Record, Federal Register, congressional bills, and many other government documents on the day of publication through GPO Access. [Try it: www.access.gpo.gov/su_docs - or dial in to 1-202-512-1661, type swais and login as guest; or telnet to swais.access.gpo.gov and login as guest]

SOFTWARE NOTES

Kodak Photo-CD/WWW Viewer

Eastman Kodak Co. and the National Center for Supercomputing Applications (NCSA) have released technology to enable WWW users to zoom into details of PhotoCD images transmitted over the Internet. To try it, check

www.ncsa.uiuc.edu/SDG/Software/WinMosaic/HomePage.html and download a version of the Photo-CD Aware NCSA Mosaic 2.1 browser for Windows.

Kodak also announced that to discourage unauthorized use of copyright images it is developing technology that will limit access to higher resolution Photo-CD data and allow content providers to watermark their images.

Business Process Driven Archiving

The Workflow Automation Company Inc. [240 Riviera Dr., Unit 1, Markham Ontario Canada L3R 5M1; 905-479-2270;

fax 905-479-2385], demonstrated the applicability of their generic workflow management front end tools for John McDonald (National Archives of Canada) and me the other day. The test John had set to them was to convert the icons of his desktop from software products to work tasks and attach to each task the transaction-based retention rules and appropriate distributions so that he could conduct business, open applications needed, send reports, be assured of their appropriate distribution and retention, simply by clicking on the proper task icon. In a quick and dirty mockup, Workflow Automation Company showed it could fit the bill.

IBM Launches Cryptolopes

The IBM InfoMarket Service announced last fall became a reality this spring. The concept is to enable pub-

lishers to provide information in packages that can be freely copied from their servers or client's machines but which require permission from a terms and conditions resolver in order to be read. The result is to overcome concerns about copyright protection, allow for effective collection processes when large numbers of small transactions are being conducted, and exploit the natural sales force of existing customers by encouraging them to copy interesting articles, images, music, or whatever to their friends and colleagues.

The initial content providers are among the largest information services in the country, including American Business Information, Associated Press, Disclosure, ESPN, Gale Research, Information Access Corporation, West Publishing, and dozens of similar firms. But IBM expects to make the software for constructing "cryptalopes" — encrypted, encapsulated envelopes that lie at the heart of this new social and economic model for distributing content — widely available. Such software will probably be free, because IBM hopes to make money not

from the sale of the encapsulation systems but from readers seeking to open the envelopes who will need to get a "key" from IBM and make payments to the content creator from which IBM will first deduct its royalty and collections fee. [For more information, check out www.infomkt.ibm.com/]

Archives Database to Web Pages

AIIMS Inc. has joined Cuadra Associates in the list of software firms providing direct database to Web page generators for archival data. AIIMS WebWiz is available as an add-on for \$900 and with a new purchase for an additional \$1500. [AIIMS, 1349 Cross Creek Way, Tallahassee FL 32301, 904-878-3096; fax 904-877-1771; msdsales@supernet.net]

STANDARDS

DCI and Gallery Systems Merge

Digital Collections, Inc.
[1301 Marina Village Parkway,
Alameda CA 94501; 510-814-
7200; fax 510-814-6100;
sales@digital-collections.com],
and Gallery Systems, Inc., the
makers of The Gallery System
and The Museum System, have
announced a merger in which the
New York based Gallery Sys-
tems will take over EmbARK
and the museum marketplace
while DCI will continue to de-
velop products for other image-
oriented vertical markets.

CIDOC Data Modeling Working Group

Nick Crofts and Pat Reed, Co-Chairs

At the ICOM meeting in Stavanger, Norway, the Data Model Working Group and the Data Terminology Working Group merged and became the Data Standards Working Group. Each of the former groups presented a product: the CIDOC Relational Data Model of the Data Model Working Group and the Information Categories Guidelines of the Data Terminology Working Group. Project plans of the Data Standards Working Group included mapping three data standards to the CIDOC Relational Data Model: Archaeological Sites, Ethnology, and Information Categories Guidelines. The working group plans also to conduct a workshop in Nairobi, details for which are not complete at this time.

The Data Standards Working Group held an interim meeting in Heraklion, Crete, Greece, March 6 - 8 to discuss object-oriented (O-O) methods and techniques for extending and enriching the CIDOC Relational Data Model. Following presentations, demonstrations, and much discussion, the group agreed to the following:

(1) Future CIDOC data models will be developed using O-O methods, techniques, and formats. While this effort is under way, the current CIDOC Relational Data Model will be available, although it will not be modified or enhanced. The O-O model will retain all the related information contained in the existing relational model, and it will provide flexibility and extensibility not possible for a relational approach. Information that is implicit in the relational model will be explicitly represented in the O-O model. Specifically, the O-O model can provide variable levels of

information, from a general overview to precise detail, enabling the representation, use, and retrieval of information from multiple points of view simultaneously.

The scope of the current relational model concentrates on the documentation of museum objects and their provenance. The O-O model will enable more specific collections management and research information to be included.

Communication, interchange of information, and public access to museum information will be enhanced.

(2) The group intends the O-O model to be a “Conceptual Reference Model” (CRM), which can be enriched by complementary, domain-specific models. The CRM is seen as a key product which will provide the intellectual and conceptual framework for defining and integrating formal compatible subsets and extensions. The role of the group is to provide a forum for the consolidation, validation, and integration of these CRM components.

(3) A first draft of the CRM will be created by a subgroup which will analyze the current relational model and convert it to an O-O format. The subgroup will report to the larger group on problems it encountered, identify the necessary transformation rules, and make recommendations for enhancing the CRM. This first version should be completed and distributed to working group members in time for the September CIDOC meeting in Nairobi.

Future work involves testing and extending the CRM by applying it to specific applications, projects, and problem areas. This work will be entrusted to subgroups with specific interests and expertise who will report to the full working group.

To complement the O-O model, the group intends to provide support material to include presentation documents, guides, transformation rules, and other information needed to understand and

use the new model. The O-O model requires a different mind-set for looking at museum data: although it is more flexible and extensible, it also is more complex.

The Working Group looks forward to sharing this work with the CIDOC community in September in Nairobi.

[For more information contact: Nick Crofts, Ville de Genève, Direction des Systèmes d'Information, (022)70 08 86; fax:(022) 733 19 1; e-mail: nicholas.crofts@ville-ge.ch; or Pat Reed, Smithsonian Institution, OIT, A&I 2310, MRC 33, Washington, DC 20560; (202)357-059; fax:(202)786-2687; e-mail: preed@sivm.si.edu]

GILS Advanced Search Facility Project

As anyone who has tried to use GILS has discovered, the “locator service” is likely to fail if its search facilities aren't dramatically improved. Of course much of the problem lies in the quality of the data presently being provided by government agencies, which in turn reflects the incoherence of government policy about what is supposed to be in GILS and how it is supposed to be used. Nevertheless, efforts are underway to improve the technical facility. One of these, an RFP for the “Advanced Search Facility” is designed to provide an efficient means of gathering and indexing federal information on the Internet. It will support the ability to construct many different types of customizable indexes suited to each information collection. The system will provide a flexible scheme for sharing gathered information among indexes that need it. The system will deploy automatic indexing of the content of sites rather than the collection of all pages into a single index. The aggregated index will also include the contents of the Government Information Locator Service compliant locators (GILS Application Profile as specified in FIPS PUB 192) at each of the Federal agency sites.

The following text is taken from the draft RFP which documents three categories of requirements: I. Search Engine; II. User Interface; and III. Implementation issues.

I. Search Engine — The search engine will have multiple components or “services”

A. A gatherer service that extracts documents from network servers. The gatherer service component of the search engine will:

1. be able to automatically access a Federal information provider from across the network using the File Transfer Protocol (FTP), Gopher, Z39.50, or HTTP protocols as well as native system protocols.
2. allow interoperation via international standard protocols with systems that do not run the search engine system software.
3. be run at the information provider site. If a site does not support the gatherer then a gatherer will be run remotely.
4. allow for the gathered information to be shared by many brokers.
5. determine if a Persistent Uniform Locator (PURL) or a Handle exists for each gathered document.
6. assign PURLs to documents that do not have them and add that information to the system PURL resolver.
7. allow for gatherer administrators to limit and tailor data collection by location, data type and other criteria.
8. allow for the addition of gatherer information from batch updates.

B. A broker service will collect information from many gatherers, to build an index of widely distributed information. The broker service component of the search engine will:

1. cascade indexed views by retrieving information from other brokers.
2. communicate via Z39.50 at a minimum.
3. identify documents that have GILS locator records and make that information available to other parts of the system.

C. An indexer service will parse gathered documents and identify index terms representing the content of documents.

1. The system must support a variety of indexing techniques including:
 - a. simple word-based indexing
 - b. indexing based on part-of-speech tagging and phrase identification
 - c. indexing by domain-dependent features such as company names, dates, locations, etc.
 - d. fielded tag-value pairs
2. The indexer will create GILS compliant locator records for any documents that do not currently have GILS compliant locator records.
 - a. The generated locator records will be as complete as possible from information provided by the broker, indexer, and any domain-dependent information available.
 - b. PURLs will be used as identifiers in the GILS records.
 - c. GILS records will be exportable via Z39.50 and in a human editable form.

- d. GILS records can be imported into the system.
- D. A query processing service involves a series of steps to identify the important concepts and structure describing a user's information need.
1. The system must represent and use complex structured descriptions in a probabilistic framework.
 2. The system must integrate natural language, Boolean, and proximity queries, including field-based retrieval.
 3. The system must represent and use spatial information in indexes and queries.
 4. A part-of-speech tagger is to be used to identify candidate search phrases.
 5. Domain-dependent features are recognized and meta-terms inserted into the query representation.
 6. The relative importance of query concepts is also estimated, and relationships between concepts are suggested based on simple grammar rules.
 7. The system must expand the query using relationships between concepts found by either using manually specified domain knowledge in the form of a simple thesaurus, domain knowledge in the form of GILS compliant locator records or by corpus analysis.
 8. The system must provide relevance feedback. Relevance ranking must be able to use GILS compliant metadata to determine ranking by GILS attribute under user control.
 9. The system must support learning techniques for query modification.
 10. Document indexing and query processing are customizable.

11. The system must identify "best passage" phrases matching query terms.
 12. Boolean queries can be evaluated using normal Boolean logic or probabilistically. Both "hard" and "soft" Boolean operators must be provided.
 13. Accept queries via HTTP or Z39.50.
 14. Expose indexes via Z39.50 for automated search.
 15. The server must be able to accept and respond to spatial queries. A spatial query is one based on an enclosing rectangle of latitude and longitude on the earth's surface. A spatial query must also be able take time (date) into account in the indexing of documents and ranking as requested.
 16. Ranking based on similarity must take spatial coverage into account when that information is available and appropriate for the user's query.
 17. The server must be able to accept and respond to a special form of spatial queries which is based on officially named locations. This query type should support the ability to automatically construct the latitude and longitude rectangle from a distance from the known point.
- E. A PURL resolver to translate and maintain document names in a system independent fashion.
- F. A document server.
1. Return identified PURLs via HTTP. PURLs once returned by the document server will be resolved in subsequent queries to this or another PURL resolver.
 2. Return GILS compliant locator records via Z39.50
- G. The system including Z39.50 support must be extensible so that additional domain application profiles may be used as they are developed.

II. User Interface

- A. Allow alternative forms-based query specifications: natural language and fielded-boolean. Both of these alternatives must exploit all features of the search engine, including relevance-ranking feedback and best-passage highlighting.
- B. The locator records shall be person-editable so that additional information can be added to the locator records in order to manually weight the ranking process by system administrators. This will allow versioning and chain of authority in GILS to be used in ranking search results.
- C. An "original" document must appear first in a ranked list if it can be identified as such. It and all of its versions must appear before documents referencing it. The ranking order must allow for newest version first ranking of "original" documents.
- D. Appropriate user interface screens for accessing all of the system features from a platform independent client located somewhere on the internet. The basic user access must be available through HTML 2.0 compatible browsers. The interface screens should be based on commercial WWW browser technologies which may include JAVA (tm) applets for advanced user features specifically identified during development by mutual agreement.
- E. "Best passages" determined by the indexer must be highlighted.

III. Implementation issues

- A. Prototype software be installed and tested at two sites on the internet. Two sites are needed to test and demonstrate the distributed indexing and search capabilities of the prototype system.

- B. The government will provide access to two or more Z39.50 V2 compatible GILS compliant servers to test the ability of the system to search, index, and retrieve documents referenced in a GILS record on the network.
- C. The government will provide access to one or more Z39.50 V2 compatible servers to test the ability of the system to search, index, and retrieve documents referenced in a GILS record that contains latitude and longitude information and/or place name information.
- D. The government will provide access to two or more HTTP servers connected to the internet to test the ability of the system to search, index, and retrieve documents from these sources.

STANDARDS NOTES

Z39.50 Profile for Access to Digital Collections

Draft Seven (and final) of the Z39.50 Profile for Access to Digital Collections is available for review at <http://lcweb.loc.gov/z3950/agency>. Much of the document is available as html. PostScript and WordPerfect versions are provided via the above listed web page or directly

<ftp://ftp.loc.gov/pub/z3950/profiles/collections.ps>

<ftp://ftp.loc.gov/pub/z3950/profiles/collections.wp>

Comments are welcome through June 1. Subsequently, following resolution of comments, the profile will be proposed for adoption. Background is included within the document.

To summarize progress to date, in August 1995, the Library of Congress (LC) convened a team of representatives from several institutions to develop a Z39.50 profile for access to digi-

tal libraries. Early in its development, the profile was renamed "Z39.50 Profile for Access to Digital Collections" as its scope was narrowed to navigation of digital collections. Other groups were initiating independent efforts to develop profiles aimed at specific types of objects and collections. The intention was to coordinate these efforts; these latter, "companion" profiles would be developed as compatible extensions or subsets of the Collections profile. Example are the "CIMI profile" (for museum objects) and the "Z39.50 Profile for Access to Digital Libraries" being developed by LC.

Set On-line Payment Protocol

In February 1996, Visa and Mastercard finally announced a preliminary specification for credit card transaction interoperability.

The Secure Electronic Transaction (SET) specification will be tested over the next several months and is expected to be in use by consumers by fourth quarter. One venue for such implementation was almost immediately announced by Oracle Corporation, VeriFone, Inc., and Wells Fargo which announced that they will implement SET in end-to-end transaction solutions for the Internet. Netscape Communications Corporation and First Data Corp. have launched a separate Internet payment system using FDC's processing services. None of these announcements in themselves represent a striking change, it is just that the reality of full scale Internet commerce is now scheduled for 1996 arrival.

New Standards Publications

The National Information Standards Organization has published Bill Moen's readable *A Guide to the ANSI/NISO Z39.50 Protocol: Information Retrieval in the Information Infrastructure* (Bethesda, MD: NISO Press, 1995). 12 pp.

\$15.00 [301-567-9522; fax 301-567-9553]

The 1995 edition of the **Library of Congress** publication, "Thesaurus for Graphic Materials I" (TGM I), is now available on the WWW. The thesaurus contains over 5,500 subject terms and 4,300 cross-references compiled and used by the Prints and Photographs Division of the Library of Congress and by libraries, archives, and museums elsewhere. The WWW version of TGM I will be updated regularly. Check it out at: lcweb.loc.gov/rr/print/tgm1/.

The Powerhouse Museum, Sydney, Australia, has published its two volume "Collection Thesaurus" consisting of over 5,000 entries for object names in an hierarchical list (vol. 1, 62 pp.) and an alphabetical list with full entries (vol. 2, 292 pp.). These are available for AU \$65.00, plus postage, most conveniently paid by Amex, Visa, or Mastercard, from: Powerhouse Museum, Registration Department, PO

Box K346, Haymarket 2000,
Sydney, Australia; fax+612-
217-0158.

The University of Utrecht,
Computers & Humanities Pro-
gram has published the CD-
ROM version of ICONCLASS
with a browser for Windows. For
those who would use ICON-
CLASS I suppose this is very
nice. I still find it unintelligible.
You can ask to try it [Achter de
Dom 22/24, 3512 JP Utrecht, the
Netherlands; fax +31-30-253-
9221; e-mail ICON-
CLASS@LET.RUU.NL]

CONTRIBUTORS

.....

David Bearman has been President and Senior Consultant at Archives & Museum Informatics in Pittsburgh, Pennsylvania, since 1986. Previously, he served as Deputy Director of the Smithsonian Institution Office of Information Resource Management and Director of the National Information Systems Task Force of the Society of American Archivists. He co-chaired ICHIM '95 held in conjunction with the Museum Computer Network in San Diego in October 1995.

Ecaterina Geber is from Eastern Europe. She works at ARTEXPO Foundation in Bucharest, Romania, an independent, non profit, non governmental organization, dedicated to the promotion and development of visual arts, printing crafts and modern communication technologies, supporting, through joint projects in the country and abroad, programs to enhance the impact of visual arts.

Xavier Perrot is an information scientist specializing in interactive multimedia and hypermedia for museums. Perrot lectures at the Hypermedia Department at the University of Paris, where he has recently earned his Ph.D., and is a research fellow of the Studio for Creative Inquiry at Carnegie Mellon University. He also serves as an independent consultant on interactive multimedia content production and systems design and as a regular columnist for *Archives and Museum Informatics*.

