Spring Forum, New York City

May 14th-May 16th 2003

Sheraton New York Hotel and Towers
811 7th Avenue at 53rd Street

Schedule

Pre Forum

Tuesday, May 13

9:00-5:00: **METS Board Meeting** - *New York University, Elmer Bobst Library, Room 1135*

Wednesday, May 14

9:00-1:00: **Developers Forum** - *Royal Ballroom A*

10:00-1:00: **E-resources Management meeting (Tim Jewell, chair)** - *Park Suite 5*

Forum

Wednesday, May 14

1:00-2:00: Registration - *Versailles Ballroom*

2:00-3:30: **Plenary 1** - *Versailles Terrace*

Welcome and Introduction: David Seaman, Director, Digital Library Federation.

"Public and Private Initiatives in Copyright Reform." James Boyle, William Neal Reynolds Professor of Law at Duke Law School.

3:30-4:00: Break - *Versailles Ballroom*

4:00-5:30: **Breakout Session 1**- *Versailles Terrace*

Panel discussion -- "Supporting Scholarly Publishing: Opening up the Tools of Electronic Publishing." Chair: Maria Bonn, University of Michigan Library.
As Library support of electronic scholarly publishing has grown, there has been an increased interest in the library community in tools that manage the publication and peer review process. There are a number of commercial products now available but because these tools are proprietary and often expensive they are not necessarily satisfactory options for libraries and academic institutions. This panel will be a group presentation that establishes a context for discussing library interest in this area, then surveys currently available commercial and open source products. This will include some specific highlighting of the capabilities of D-Space and the Open Journal Systems of the Public Knowledge Project (http://pkp.ubc.ca). The panel will conclude with a trial proposal suggesting ways in which the DLF community might collaborate in creating more open and accessible publication management systems and a request for discussion of this proposal. Participants include: Maria Bonn, Scholarly Publishing Office, University of Michigan Library; Marcy Rosenkrantz, Director, Library Systems, Cornell University; Roy Tennant, California Digital Library; John Willinsky, Public Knowledge Project, University of British Columbia; and MacKenzie Smith, Associate Director for Technology, MIT Libraries.

"The Bibliographic Enrichment Advisory Team." David Williamson, Library of Congress

The Bibliographic Enrichment Advisory Team at LC has several projects under way to enhance catalog records with information not traditionally found in the catalog record. These include tables of contents (TOC), publisher descriptions, reading group guides, and sample texts. This information is generated through automated means and cannot easily be put into a MARC field, so the information is instead put on the world wide web and linked to the catalog record through the 856 field. One aspect of the TOC project involves scanning TOCs and making the converted text available. The other TOCs and pieces of information are extracted from ONIX (ONline Information eXchange) XML files provided by several sources. This presentation will describe the projects, how successful they are in terms of how frequently the information is used, and will describe an unexpected side affect-- the information is being indexed by search engines and is bringing users back into the library catalog at a time when many are leaving the library catalog in favor of the search engines. For more information about the projects, please see the BEAT home page at http://www.loc.gov/catdir/beat/.

"OCLC Metadata Switch." Thom Hickey, Chief Scientist; Jean Godby, Research Scientist; Diane Vizine-Goetz, Research Scientist, OCLC

Libraries are creating more metadata for more types of material. They are using different metadata formats. In some cases they may use different metadata systems It would not be unusual for a library to be creating metadata in some or all of the following: MARC, Dublin Core, Encoded Archival Description (EAD); or, for this activity to take place within a library management system, a content management system, a custom-developed database application, or some combination of these. We are also seeing a growing interest in harvesting metadata, pulling metadata from different repositories, fusing it and disclosing it in 'union' services. The mechanics of harvesting are becoming routine and well understood, and this is now introducing the interesting challenge of effectively fusing metadata so that a useful retrieval experience can be offered. This metadata will often not have been created within a framework of consistent practice; approaches to subjects or names will be different for example. The question we have asked ourselves is what type of services would be valuable to libraries in this increasingly diverse environment.

The OCLC Metadata Switch project is a response to this question. The Metadata Switch is an umbrella activity for a set of projects, which are constructing experimental modular services that add value to metadata:

- harvesting metadata
- 'fusion' of metadata from different sources
- schema transformation
- enrichment or augmentation of records with various types of data
- terminology and name authority services

We are testing how to make this functionality available as 'web services', web-based machine-to-machine applications which can be combined in various ways. The presentation will provide an overview of the project and describe the services in more detail.

"RedLightGreen." Merrilee Proffitt, RLG
With generous funding from the Andrew W. Mellon Foundation, RLG began work in late 2001 to begin looking at ways in which it could make the rich information held in the RLG Union Catalog available to a wider audience in a freely available web environment. During the intervening time, we have learned volumes about:

- what undergraduate users want from online information resources
- how data mining software can uncover valuable new information hidden in the RLG Union Catalog
- how to provide access to a wealth of complex information through a simple, easy-to-use interface
- new opportunities for using bibliographic data to help end users find authoritative sources of research information
- what's involved in the complicated process of transforming MARC records to XML
- incorporation of concepts outlined in the Functional Requirements for Bibliographic Records (FRBR), an emerging standard for distinguishing between various editions

In the coming months we expect to learn even more about our key user audience. More than with any other project in the past, RLG is working to reach the end user directly. This is a new area for us, and we will be able to share our findings on user needs as well as what we learn in more technical areas. This presentation will feature an explanation of the history, motivations, "lessons learned" and future directions of the project, a live demo of the pilot system, and outtakes of use studies.

6:00-9:00: Reception. Low Library, Columbia University.

Cab fare from the Sheraton Hotel and Towers to Columbia University is about $15. The campus is closed to traffic so the cab will drop you at the main gates. Enter through the gates and walk to the center of the campus. The Low Rotunda will be to your left. **Subway**: Take the 1 or 9 train at 50th Street and 7th Avenue uptown to the West 116th Street and Broadway stop. When you exit the subway, you will see the main gates of the university. Enter through the gates and walk to the center of the campus. The Low Rotunda will be to your left.

Thursday, May 15

: Continental Breakfast - *Versailles Ballroom*

: **Plenary 2** - *Versailles Terrace*

> **"The DLF Today and the Case for the Distributed Online Digital Library (DODL)." David Seaman, Director, Digital Library Federation**

: Break - *Versailles Ballroom*

: **Breakout Session 3 - Versailles Terrace**

> **Federated Searching: Vendor Panel Discussion. Invited parties include Endeavor, Innovative, Ex libris, Sirsi, and Fretwell Downing. Moderator: Laine Farley, Director of Digital Library Services, California Digital Library**

: **Breakout Session 4** - *Royal Ballroom B*

> **"FEDORA Digital Repository Implementation at UVa--How, Why, and What We're Doing with It." Leslie Johnston, Director, Digital Services Integration, University of Virginia Library.**
> In late 2001, funded by a grant from the Mellon Foundation, the UVa Library began production implementation of the FEDORA system with FEDORA's original developers at Cornell University. In April 2003 the first release of the software will be completed, and UVa will launch its first phase production repository. This talk will not focus on the system architecture, but will instead cover how the system was implemented and why, and what user services are in development based on the system. Among the topics covered will be:
>
> - How the implementation of the FEDORA architecture provides the necessary infrastructure for a repository and its public user services;
> - How the implementation of the system at UVa affected digital production specifications and workflow (file formats, file naming, batch production processes, descriptive metadata practices, etc.);
> - How object models and disseminators were developed to support user services needs (necessary delivery formats, static and on-the-fly file delivery, end user download and printing requirements, etc.);
> - How the delivery interface interacts with the disseminators and objects in the FEDORA repository (indexing and searching, use of

XSLT and JavaScript to format delivery, a Java image delivery applet, etc.); and

- Future authoring and delivery services that will be developed on top of the repository for users (an XML authoring tool, distributed repository storage in affiliated user's "home directories" elsewhere in the University environment) a "shopping cart" for the repository, and the linkages between repository disseminators and commercial tools for end-users).

"The University of California's Collection Management Initiatives: Findings on Use of and Preference for Digital Journals." Gary S. Lawrence, Director, Library Planning and Policy Development, California Digital Library.

At the Spring 2001 DLF Forum, the University of California reported on the launch of its Collection Management Initiative, a two-year study supported by the Andrew W. Mellon Foundation. The project is intended to explore how scholars and libraries can best integrate collections of scholarly journals that are published in both print and digital formats. At the Spring 2003 Forum, UC would like to share with DLF members preliminary results from the project. Data include:

- A year-long study of the use of about 300 journal titles. At one UC campus, print volumes of a study title were removed from the shelves; at another, the title was retained on the shelves and use was monitored. Both digital use and print use (including recalls from storage) were recorded at both locations.
- A survey of approximately 7,000 faculty, students and staff across the nine-campus UC system. The CMI User Preference Survey examined possible barriers to using digital journals, types of use suitable for digital journals, elements in the users' environment that might promote or discourage use of digital, and characteristics of journal presentation and content that might be important to digital or print use and preference.

Preliminary findings from these two data sources will be presented and some tentative conclusions discussed.

"Cushman Exposed: Exploiting Controlled Vocabularies to Enhance Browsing and Searching of an Online Photograph Collection." Michelle Dalmau, Interface and Usability Specialist, and Jenn Riley, Digital Media Specialist, Indiana University Digital Library Program.

Indiana University's Digital Library Program is making accessible online approximately 15,000 color slides taken by amateur photographer Charles W. Cushman between 1938 and 1969. The collection is described both by Cushman's own detailed descriptions and controlled vocabularies, including the Thesaurus of Graphical Materials I: Subject Terms [TGM I]

and Thesaurus of Graphical Materials II: Genre and Physical Characteristic Terms [TGM II]. Because natural language and controlled vocabularies work together to provide a complete description of an image, browse and search mechanisms should be able to integrate both types of metadata. While controlled vocabularies are often used in subject cataloging, they are seldom used to enhance searching. By not using the syndetic structure to browse and search the collection, the user's language cannot be mapped to the controlled vocabulary thereby limiting access. In an effort to provide an integrated browse and search interface, we are investigating ways to use the inherent syndetic structure of the controlled vocabularies to provide users with transparent access to authorized terms as well as the ability to expand their original query by either broadening or narrowing current searches. We will present the results of iterative user testing of interface prototypes that explore browsing and searching using both natural language and controlled vocabularies. Studies were designed to understand fundamental user browse and search behaviors as a basis for exposing the controlled vocabularies as integrated, dynamic tools to enhance browsing and searching in an intuitive manner.

12:30-2:00: Lunch - *Versailles Ballroom*

2:00-3:30: **Breakout Session 5** - *Versailles Terrace*

"Designing a library support infrastructure for the use of digital content for teaching." Danuta Nitecki, Associate University Librarian for Public Services & Systems, Yale University Library; Karen Reardon, Manager of Web & Workstation Services, Yale University Library
Digital initiatives have addressed issues of platform design, integration, content organization, and intellectual property rights. The Yale University Library has embarked on a three-year, grant-funded project to explore the impact of digital content on teaching and learning of American studies, which in addition assesses the implications for its library services and support infrastructure. Six working groups composed of library and IT staff are addressing these issues. They are producing feedback to platform vendors [Luna] and library system designers, programming solutions to integrate digital content with courseware, self-service cataloging templates, instructional tool kits, decision trees for respecting copyright and licenses governing content, and assessment tools. In this presentation, the project's managers will review this investigation and its organization, including use of course-specific faculty support "swat" teams. They will critique preliminary insights about this organizational structure, the nature of collaborating with faculty, staff participation on support teams, and the challenges of bringing together the right mix of people to foster use of digital content in teaching. The presentation will discuss techniques introduced to faculty to assess how digital content contributes to achieving

their pedagogical goals. It will conclude with speculations on how future library service support will be sustained as demand for this service continues to grow. The presenters will welcome audience discussion of our experience with this library support infrastructure and its applicability to other settings.

"Ask A Librarian and QuestionPoint: Integrating Collaborative and Digital Reference in the Real World (and in a really big library)." Linda J. White, Digital Project Coordinator, Library of Congress

New patrons, new technology, and new sources of information all converge to create challenges for today's reference librarian. While the fundamental nature of librarianship remains the same (connecting people with the information they need) there are new tools available and new skill sets required to migrate the delivery of information and services online. Learn how the Library of Congress has been able to bring its reference service online using QuestionPoint. QuestionPoint is an OCLC/LC developed online reference management tool and a global network of librarians exchanging reference queries and building a global knowledge-base. QuestionPoint participation is based profiles. Each member builds a profile. Profiled information includes subject (based on LC class), geographic subject, formats, education levels served, open hours, weekly quota, languages served, etc. When a librarian wishes to send a question to the global reference network the librarian "codes" the question in regard to the above mentioned elements. The question is sent to the network where a two-step process is employed. The first step is exclusionary and all profiles not supporting the defined subject/geographic areas are eliminated, as are those who have met their weekly quota, are not open during the life of the question, or don't serve the education level. Step two utilizes a best-fit matching algorithm wherein all remaining members are scored based on a series of weights and values. Question metadata meets profile metadata and a match is made. Exploration is ongoing on ways to incorporate linguistic analysis into this process. The technology is beginning to mature and is able to incorporate established taxonomies like LC class and Dewey making the coding and the profiling process more intuitive. Emerging technology and the possibilities for global collaboration have resulted in an active migration of LC reference services to the web to meet our patrons where they are. This session will highlight the process of this migration and the creation of a world-class public-private partnership and a global network of librarians willing to share their expertise. We'll start with a clear vision, and then outline the steps, the partnerships, the challenges, the successes, the software, the future, and the lessons we've learned in moving the mountain of reference into the online arena.

"Information Communities." Christine Ruotolo, Lead Information Community Coordinator, University of Virginia Library.

This presentation will describe the development of Information Communities at the University of Virginia Library. Information

Communities represent a user-driven, subject-oriented approach to organizing library content, and are a central piece of Virginia's Library of Tomorrow plan to integrate its digital initiatives with traditional collections and services. Information Communities will bring colleagues from related areas of study together to discover resources, build collections (texts, images, videos, audio, and maps), access tools, share news, and collaborate in research and development. Resources available through an Information Community will include digital texts and collections; peer-reviewed and annotated bibliographies; online tools for research and teaching; and communication services such as email lists, discussion forums, electronic conferencing, and directories for individuals and projects. The development of new communities will necessarily evolve around Virginia's own library and faculty priorities. However the model is designed to foster broad collaboration across institutions. In fact, the first prototype community, the Tibetan and Himalayan Information Community, already features an international roster of users and contributors. After explaining the Information Communities model, this presentation will describe the core features and services that will be common to all communities, and discuss how these are currently being developed at Virginia. It will then demonstrate the American Studies Information Community, which we hope will establish the model for collection, tool, and interface development.

2:00-3:30: **Breakout Session 6** - *Royal Ballroom B*

"A Registry for Digital Format Representation Information." Stephen L. Abrams, Digital Library Program Manager, Harvard University Library; MacKenzie Smith, Associate Director for Technology, MIT Libraries.
The concept of data format permeates all technical areas of digital repositories. Policy and processing decisions regarding ingest, storage, access, and preservation are frequently, if not uniformly, conditioned on a per-format basis. The current IANA media type (MIME) registry does not capture format-specific information at an appropriate level of granularity, or in sufficient level of detail, for many digital repository activities. An international effort is underway to establish a sustainable format registry that will maintain persistent, unambiguous bindings between public identifiers for digital formats and representation information about those formats. For the purposes of the registry, a format is defined as a fixed octet-serialized encoding of an information model. Format representation information will capture the significant syntactic and semantic properties of the format, with particular relevance towards the operational needs of digital repositories, including object format identification, character-ization, ingest validation, migration, emulation, and other preservation activities.

"Strategies for Implementing Preservation Metadata in Digital Archiving Systems." Rebecca Guenther, Library of Congress

Report of a new working group that OCLC and RLG is jointly organizing to address strategies for implementing preservation metadata in digital archiving systems. Under co-chairs Rebecca Guenther and Priscilla Caplan (Florida Center for Library Automation) the group commences its activities in April 2003, and this presentation will include a high-level overview of the group, who is participating, its planned objectives, and a summary of the workplan the group has established.

"PDF/A: A New Digital Preservation Format." Bill LeFurgy, Digital Initiatives Project Manager, Library of Congress

Current digital formats present problems for digital preservation. Most formats now in use are proprietary, based on hidden specifications, change rapidly, or do not support required significant properties. This complicates all aspects of life for digital repositories, from issuing guidance to producers, to pre-ingest/ingest, to sustaining, and to making content available. To address these problems, a committee of government, industry, and academic representatives has started work on an International Standards Organization (ISO) specification for a basic subset of the Adobe Portable Document Format. Known as PDF/A, this specification will govern creation of documents that are self-contained, technologically stable, and have the basic properties that users need. Expectations are that PDF/A will primarily be used for text-based documents, although the specification will permit a mixture of text, raster images, and vector graphics. The PDF/A committee is also looking to include the Adobe Extensible Metadata Platform (XMP) in the standard. Based on the W3C RDF/XML recommendation, XMP offers rich possibilities for capturing and using metadata. The presenter is a member of the PDF/A committee and is interested in getting comments and suggestions.

3:30-4:00: Break - *Versailles Ballroom*

4:00: **Birds of a Feather Sessions**

    Electronic Resource Management Initiative - *Versailles Terrace*

    DSPACE - *Park Suite 1*

    METS (including docWORKS/METAe) - *Royal Ballroom B*

    Digital Camera Operators Group - *Park Suite 2*

Friday, May 16

"Virtual Remote Control of Web Resources: A Risk Management Approach for Research Libraries."
Nancy Y. McGovern, William R. Kehoe, Richard Entlich, and Peter Botticelli, Cornell University Library.
The rapid growth of the World Wide Web provides extensive new resources of interest to research libraries. There are numerous Web sites and pages that are neither owned nor controlled by libraries, but may provide essential enhancements to existing resources or form the nucleus for new collections. Research libraries will have a vested interest in ensuring the longevity of these sources as elements of their collections. Our approach adapts the risk management model to the monitoring and evaluation of changes to Web resources over time. It will be used to predict probable loss based on the presence or absence of key indicators that may enable or inhibit longevity. Adapting or creating Web management tools provides the basis for the resulting preservation risk management program. This paper discusses our proposed tool-based Virtual Remote Control approach, results of the research to date, and future directions.

"Factors influencing the selection of a digital asset management system."
Deborah Holmes-Wong, Project Manager, Digital Information Management, University of Southern California Library.
USC recently completed selecting a vendor for its Collection Information System (CIS). The CIS will be the infrastructure that stores and delivers our digital assets. During the process the selection team winnowed the list from over sixty potential vendors of digital asset management systems, content management systems, systems integrators, library systems vendors, and open sources solutions to one finalist. This paper will examine how internal and external factors, vendors' philosophy, and system capabilities were taken into account in the evaluation process and influenced the outcome. It will outline the systems capabilities found in commercial and open source software in each category and discuss why they were important in the context of building digital collections and how these capabilities might also be repurposed in the future for use within learning management systems and institutional repositories.

"Rights management for licensing and digitizing resources." Tim Jewell, Collection Management Services, University of Washington Libraries; Joyce L. Ogburn, Associate Director of the Libraries, University of Washington

An overview of rights management focusing on the topics of metadata and standards for electronic resource management and for agreements for digitizing projects.

9:00-10:30: **Breakout Session 8** - *Royal Ballroom B*

"Robotic retrieval and scanning." Sayeed Choudhury, Johns Hopkins University Library.

The Digital Knowledge Center (DKC) at Johns Hopkins University is developing a framework for the large-scale digitization and processing of cultural heritage materials. The framework comprises a hardware system, CAPM, and a software system, Gamera. CAPM features a robotic system that allows the retrieval and scanning of materials located in remote facilities. CAPM will allow users full remote control for on-demand browsing, and batch scanning capabilities to create digital copies for preservation and subsequent access. The first phase of the CAPM Project resulted in the development of a prototype retrieval robot, and an economic framework for evaluating digital library services. Gamera is a toolkit for the creation of domain-specific document analysis systems by document experts. It grew out of the development of an adaptive optical music recognition (AOMR) system that automatically recognizes musical symbols and lyrics from images of sheet music. Gamera extends this functionality by providing a framework to recognize and process symbols, page layout, and other elements within images of cultural heritage materials. Our current work focuses on recognition of early modern English and medieval French. The National Science Foundation, Institute of Museum and Library Services, the Mellon Foundation, Minolta Corporation, and private donors have provided the funding for CAPM and Gamera.

"The Preservation Productivity Paradox in the Modern Digital Library." Catherine Aster, Stanford University Libraries, and Stuart K. Snydman, Digital Library Projects Manager, Stanford University Libraries.

The values of preservation and high-volume productivity appear to be at odds in the modern digital library. On the one hand digital library programs are faced with burgeoning demand for online access to library collections. On the other hand, digital preservation and conservation professionals are concerned about creating quality, high-resolution digital representations that can be preserved for centuries while at the same time taking deliberate care to protect precious originals from the stresses of the scanning process. A collaboration between two departments at the Stanford University Libraries provides an occasion to consider the

challenges and opportunities of reconciling these seemingly disparate approaches to the digital library. The Stanford Library's Digital Library and Media Preservation units have joined together to build a new digitization lab that aims to produce preservation quality representations of bound materials in a high-volume production scanning environment. The lab intends to satisfy the growing demand for access to digital content, while at the same time ensuring the long-term preservation of both the original documents and their digital surrogates. At the center of this collaboration is a newly developed robotic book scanner, which turns pages and scans books at upwards of 900 pages an hour with little human intervention. Using their experiences in conceiving this new lab, the presenters will discuss the practical and ideological challenges in bringing preservation principles to a production scanning environment. They will illustrate the possibilities created by emerging technologies in digital photography, robotics, networking and software, and will discuss how creative thinking and a spirit of compromise can change the future of the digital library.

10:30-11:00: Break - *Versailles Ballroom*

11:00-12:30: **Breakout Session 9** - *Versailles Terrace*

"Enhancing Interoperability between Digital Libraries and Educational Technology via XML Crosswalks." Raymond Yee and David Greenbaum, Interactive University Project, UC Berkeley.
Lack of interoperability among software systems and repositories from different domains is a major barrier to the exchange of digital content. The UC Berkeley Interactive University Project (IU) has been exploring how semantic interoperability in the following four domains can be enhanced through the use of XSLT-based crosswalks between key XML specifications: 1) digital libraries and repositories (METS); 2) educational technologies (SCORM, IMS specifications); 3) web syndication and portal technologies (RSS); and 4) desktop applications and structured content authoring tools. (e.g., Microsoft Office 11). The outcome of this work includes crosswalks among METS, SCORM/IMS-CP, and RSS and an exploratory architecture and software prototypes for the deployment of crosswalks. The IU will relate the crosswalks to other approaches to semantic interoperability.

"Schema-Driven XML Editor for Metadata Capture." Stephen L. Abrams, Digital Library Program Manager, Harvard University Library
All significant phases of digital repository operation are dependent upon the metadata for the stored digital objects. The preferred form for interchange, and, increasingly, storage, of this metadata is XML. While some metadata may be extractable from the objects themselves, the majority of the descriptive, administrative, and technical metadata must be

collected manually, typically by catalogers who are domain experts with little understanding of XML. In order to streamline this process we have developed a schema-driven XML editing application that can be configured to support arbitrary XML schemas. The editor presents a series of input forms based on the schema structure, and prompts for values of the defined elements and attributes, with automatic verification of schema constraints on optionality, cardinality, type, value, etc. The editor is a Java application portable to any platform. In its preliminary use capturing complex metadata for a biomedical imaging project, the editor has proved to be successful in reducing the burden of metadata capture on collection catalogers.

"The OAI Static Repository: a file-based approach to exposing metadata via the OAI-PMH." Herbert Van de Sompel, Los Alamos National Laboratory, Research Library
Although the OAI-PMH specification is focused on making it straightforward for data providers to expose metadata, practice shows that in certain situations deployment of OAI-PMH conformant repository software remains problematic. The presentation will report on research aimed at devising solutions to further lower the barrier to make metadata collections harvestable. Specifically, it will detail the OAI Static Repository approach, in which a data provider makes a metadata collection available as an XML file with a specific format--an OAI Static Repository. That XML file is made OAI-PMH harvestable through the intermediation of software--an OAI Static Repository Gateway--operated by a third party.

11:00-12:30: **Breakout Session 10** - *Royal Ballroom B*

"ARTstor: A Digital Library for the History of Art." Tony Gill, Director of Metadata, ARTstor; Max Marmor, Director of Collection Development, ARTstor; Linda Tadic, Director of Operations, ARTstor; Bill Ying, Chief Technology Officer, ARTstor.
ARTstor is a digital library initiative initiated by The Andrew W. Mellon Foundation. It's goal is to develop a digital library for the history of art, working with the broader community of museums, publishers, photo archives, scholars, and visual resources collections. In our presentation, we will offer an overview of ARTstor's origins, goals and progress to date. Specifically, the presentation will address

- The origins of ARTstor
- Relationship to JSTOR
- ARTstor goals
- ARTstor collection development strategy
- ARTstor's "charter collections"
- ARTstor's production strategy
- ARTstor's metadata approach

- ARTstor's technical infrastructure
- Testing plans
- Rollout plans

"The Architecture and Features of the Beinecke Digital Library." Brian M. Kupiec, Systems Officer for the Beinecke Rare Book Library, Yale University.

The Beinecke Library has been delivering digital surrogates via the web for over three years. What started out as a searchable database of some 12,000 photonegative images has grown into four cross-searchable databases, containing a total of almost 50,000 images, along with several spin off's containing several thousand more images. The current Beinecke Digital Library model has matured into a modular, easily duplicated database and software framework for storage, discovery and presentation of digital surrogates. Further, this platform allows for easy (three SQL scripts) exportation of image data into Luna's Insight V3.x database. From its beginnings as an IBM based package running on UNIX (The Digital Library Collection Treasury), to its current incarnation as an Active Server Pages application running on Windows 2000, the Beinecke Digital Library has continued to evolve. It has grown with a fuller feature set while also morphing into a more generic and modularized product. Because of this, it has seen use in the classroom and has served as the model for several non-Beinecke digital surrogate collections. This presentation will describe, discuss and demonstrate the Beinecke DL model, the current features, the collections that it hosts, and the strategies that inform future developments.

"Electronic Records Archives Program." Gregory Bluher, ERA Task Leader, Systems Engineering Division, National Archives and Records Administration (NARA)

A. Overall purpose and significance of the session: The National Archives and Records Administration (NARA) plans to build an archives that will preserve U.S. Government records of continuing value and make them available electronically to anyone, at anytime, in any place, for as long as need. In order to achieve it's goals, NARA had to establish a program management office to strategically manage the building and implementation of this future system. Program management at the office level is new to NARA and the archival community in general and has introduced some interesting concepts in the area of managing electronic records. **B. Content description**: The mission of the National Archives and Records Administration is to ensure, for the citizen and the public servant, for the President and for the Congress and the courts, ready access to essential evidence that documents the rights of citizens, the actions of Federal officials, and the national experience. Increasingly, records are created and maintained in electronic formats. To continue to fulfill its mission, the National Archives is responding effectively to the challenge posed by the diversity, complexity, and enormous volume of electronic

records being created today and the rapidly changing nature of the systems that are used to create them. Electronic Records Archives (ERA) is envisioned to be a comprehensive, systematic, and dynamic means for preserving any kind of electronic record, free from dependence on any specific hardware or software. ERA, when operational, will make it easy for the National Archives customers to find records they want, and easy for the National Archives to deliver those records in formats suited to customers' needs. The ERA Program Management Office (PMO) is responsible for managing the building and implementation of the future Electronic Records Archives. In order to achieve its goals NARA had to take a strategic direction that is not traditional to Archives. This direction included hiring both a government and contract staff whose composition range from system engineers, research directors, and quality assurance officers to risk management officer, as well as archivists, and implementing the basic principles of a program management strategy. The program management strategy introduced NARA to concepts such as Change Management, Communications Management, and Risk Management. This session will discuss our work within the ERA PMO. We will discuss the overall program direction and the keys to success. Additionally, we will discuss our work with planning and managing change, and the challenges of transitioning to the archives of the future. We will also discuss our work with communicating the ERA program, our encounters with different audiences, and the challenge of translating the technical requirements into the archival. Finally, we will discuss our work with implementing an aggressive risk management strategy and the keys to mitigating risk of such a significant undertaking.

12:30-1:00: **Closing remarks** - *Versailles Terrace*

David Seaman

DLF Forum Fellowships for Librarians new to the Profession

Congratulations to the winners of the Spring 2003 Forum Fellowships:

Brandon Burke, Graduate Research Assistant, Digital Library Services, University of Texas-Austin.

Carole George, Human Factors Researcher, Digital Library Initiatives, Carnegie Mellon University.

James Sanborn, Assistant Head / Metadata Architect, Digital Library Initiatives Department, North Carolina State University.

Ryan Max Steinberg, Web Services Programmer, Digital Library Services, University of Michigan.