

Coalition for Networked Information Spring 2012 Membership Meeting

April 2-3, 2012
Baltimore, MD

**Tag This Meeting
CNI2012spring**

**Twitter Hashtag
#cni12s**

For the latest news & information:

Announcements: www.cni.org/cni-announce/

News Feed: www.cni.org/news/

Podcast: www.cni.org/cni-conversations/

Twitter: [@cni_org](https://twitter.com/cni_org)

**CNI Spring 2012 Membership Meeting
SCHEDULE-AT-A-GLANCE**

MONDAY, APRIL 2

8:30 a.m.

Executive Roundtable (*Loch Raven*)
prior registration only

11:00 a.m.

Registration Opens
(*Chesapeake Gallery*)

11:30 a.m.

Orientation for First-Time Attendees
(*Harborview Ballroom*)

12:15 p.m.

Opening Break
(*Chesapeake Gallery*)

1:15 p.m.

OPENING PLENARY SESSION (*Chesapeake BR*)
James J. Duderstadt (U. of Michigan)
*Reinventing the Research University to Serve a
Changing World*

2:30 p.m.

Break
(*Chesapeake Gallery*)

CNI Spring 2012 Membership Meeting SCHEDULE-AT-A-GLANCE

MONDAY, APRIL 2

2:45 p.m.

PROJECT BRIEFINGS

National Status of Data Mgmt ResourceSync	<i>Chesapeake BR Potomac</i>
Interoperating Media Repositories	<i>Loch Raven I</i>
retroReveal.org	<i>Loch Raven II</i>
CDL & Public Knowledge Project	<i>Camden</i>
Federal Depository Library	<i>Severn I</i>

3:45 p.m.

Break (*Chesapeake Gallery*)

4:15 p.m.

PROJECT BRIEFINGS

Archiving User-Contributed Content	<i>Chesapeake BR</i>
NSF Open-Access Repository	<i>Potomac</i>
NYU Implementation of Sakai's OAE	<i>Loch Raven I</i>
Lib's/IT Experience w/ Consultants	<i>Loch Raven II</i>
JSTOR & New Econ Models/Mkts	<i>Camden</i>
Competing Priorities, Dig Collections	<i>Severn I</i>

5:00 p.m.

Break (*Chesapeake Gallery*)

5:15 p.m.

PROJECT BRIEFINGS

Networks & Active Learner Paradox	<i>Chesapeake BR</i>
Update on Activities of BRDI	<i>Potomac</i>
Google Scholar Indexing IRs	<i>Loch Raven I</i>
Creating Dig Preservation Network	<i>Loch Raven II</i>
Advances in Discovery	<i>Camden</i>
Portal for Geospatial Resources	<i>Severn I</i>

6:00 p.m.

Reception (*Harborview Ballroom*)

CNI Spring 2012 Membership Meeting SCHEDULE-AT-A-GLANCE

TUESDAY, APRIL 3

7:30 a.m.

Breakfast (*Harborview Ballroom*)

9:00 a.m.

PROJECT BRIEFINGS

Building Grateful Dead Archive	<i>Chesapeake BR</i>
SciBd on Scientific Data Mgt	<i>Potomac</i>
Dig Public Lib of America	<i>Loch Raven I</i>
Transforming Research Support	<i>Loch Raven II</i>
NSF EarthCube Program	<i>Camden</i>
Open Source CM for Dig Archives	<i>Severn I</i>

10:00 a.m.

Break (*Chesapeake Gallery*)

10:30 a.m.

PROJECT BRIEFINGS

Linked Data for Libraries	<i>Chesapeake BR</i>
A Study in SCARLET	<i>Potomac</i>
Electronic Journals Ownership	<i>Loch Raven I</i>
Curation for Born-Dig Newspapers	<i>Loch Raven II</i>
Climate Data for Our Future	<i>Camden</i>
Fembot	<i>Severn I</i>
Sustaining Open Source Projects	<i>Severn II & III</i>

11:45 a.m.

Lunch (*Harborview Ballroom*)

1:00 p.m.

PROJECT BRIEFINGS

DMPTool: Phase 2	<i>Potomac</i>
Federated Identity	<i>Loch Raven I</i>
Pay Once, Preservation Forever	<i>Loch Raven II</i>
Too Big to Know	<i>Camden</i>
Ranking in the Scholarly Domain	<i>Severn I</i>
US End-of-Term Web Archive	<i>Severn II & III</i>

2:00 p.m.

Break (*Chesapeake Gallery*)

2:15 p.m.

CLOSING PLENARY SESSION (*Chesapeake BR*)

Phillip D. Long (U. Queensland)

Key Trends in Teaching & Learning

3:30 p.m.

Meeting Adjourns

Monday, April 2

1:15-2:30 p.m.

Chesapeake Ballroom

OPENING PLENARY SESSION

Reinventing the Research University to Serve a Changing World

James J. Duderstadt

President Emeritus

University Professor of Science and Technology

University of Michigan

The seemingly incompatible imperatives of a changing world — massification (extending college degree attainment), league table rankings (achieving world-class research capacity and quality), exponentiating technologies (cyberinfrastructure, open learning resources, social networking), and shifting public priorities (viewing education as less a public good than a private benefit) — are all posing formidable challenges to higher education. While these challenges are driving many institutional changes at the margin (increasing enrollments, expanding use of part-time faculty, and rising tuition levels), recent studies at the international, national, regional, and institutional level suggest that not only is a more fundamental restructuring of higher education necessary, but new paradigms of learning, scholarship, and engagement may be required that will radically change the public purpose, mission, and character of the research university itself.

About the speaker:

James J. Duderstadt is President Emeritus and University Professor of Science and Engineering at the University of Michigan. A graduate of Yale University and the California Institute of Technology, Dr. Duderstadt's teaching, research, and service activities include nuclear science and engineering, applied physics, computer simulation, science policy, and higher education. He has served on or chaired numerous boards and study commissions including the National Science Board, the Executive Board of the National Academies and its Policy and Global Affairs Division, the Nuclear Energy Advisory Committee of the US Department of Energy, the National Commission on the Future of Higher Education, and the Advisory Committee on Cyberinfrastructure of the National Science Foundation. He has received numerous awards and honorary degrees including the National Medal of Technology for exemplary service to the nation. At the University of Michigan he currently chairs the program in Science, Technology, and Public Policy in the Gerald R. Ford School of Public Policy and directs the Millennium Project, a research center exploring the impact of over-the-horizon technologies on society.

Tuesday, April 3

2:15-3:30 p.m.

Chesapeake Ballroom

CLOSING PLENARY SESSION

Key Trends in Teaching & Learning: Aligning What We Know About Learning to Today's Learners

Phillip Long

Professor

Director, Centre for Educational Innovation & Technology

University of Queensland

We are beginning to understand more and more about how we learn. Data from neuroscience, cognition and memory research is telling us about how we engage with each other, filter and encode the deluge of data continually washing past us, and recreate the relationships among data to act on our world. Technologies are our ways of extending our senses and augmenting our bodies. They enable us individually as well as situate us socially. Our opportunity to engage in learning individually, in groups, or in massive collectives is transforming the landscape of higher education. For many institutions, today's trends are simply things worth noting, the impact of which are still distant or unrecognizable. For others the trends emerging have the potential to shake the foundations on which they rest. What are the most significant trends in teaching, learning, and technologies that surround us today? It is hubris to suggest a road map. Perhaps instead we should look at the topology of our current landscape, with a slightly international perspective. Which among the experiments currently underway are important? Why? What do they represent? While some describe our current learning environment as 'academically adrift,' tens of thousands are choosing to loosely aggregate in massive collective learning ventures. Technology has always been touted as a potentially disruptive innovation, but its impact on higher education has been measured. Hang onto your hats. Things are just getting interesting.

About the speaker:

Phillip Long is Professor of Innovation and Educational Technology in the School of Information Technology & Electrical Engineering and the School of Psychology, founding director of the Centre for Educational Innovation and Technology (CEIT) at the University of Queensland (UQ), dedicated to research on learning environments that have the potential to innovate teaching, learning and creativity. The Centre fosters a community of scholarship among technology innovators, and researchers within UQ, across Australasia, and around the world. Professor Long's current research interests focus on emerging technologies, the cognitive interactions with them, and the spaces, physical and virtual wherein they occur. His professional activities include serving on the boards of the New Media Consortium (NMC), NMC Horizon Project, Pearson Australia Advisory Board, Association for Authentic, Experiential & Evidence-based Learning (AAEEBL), and the Global Online Laboratory Consortium (GOLC). He has served as chair or co-chair of various conferences and meetings, including the first Learning Analytics & Knowledge conference. He also holds a Visiting Researcher appointment at the Massachusetts Institute of Technology.

**CNI SPRING 2012 MEMBERSHIP MEETING
BREAKOUT SESSIONS**

**MONDAY
APRIL 2
2:45-3:45 PM**

Monday, Apr. 02, 2012 2:45-3:45 PM Chesapeake Ballroom

National Status of Data Management: Current Research in Policy and Education

Martin Halbert
Dean of Libraries
University of North Texas

William E. Moen
Associate Dean for Research
College of Information
University of North Texas

Rachel Frick
Director, Digital Library Federation
Council on Library and
Information Resources

Spencer D. C. Keralis
Postdoctoral Research Associate, Libraries
University of North Texas

The demands of big data pose significant challenges for research institutions and academic libraries. This panel features project updates from three interrelated projects examining the data management ecosystem to determine requirements and emerging best practices in policy, graduate education, and professional enrichment. The University of North Texas (UNT) Libraries' DataRes project, funded by the Institute of Museum and Library Services (IMLS), is examining the policy landscape at top research institutions across the nation and is conducting a comprehensive investigation of the concerns of key stakeholders in the research lifecycle (including researchers, administrative officials, librarians, funding agency officials, research equipment vendors, and others) with regard to data management plan mandates from funding agencies, the long-term management of research data generated in universities, and the role of information professionals in such efforts.

The UNT College of Information's iCAMP project, also funded by the IMLS, is developing a four-course, online, graduate academic certificate program in data curation and data management for information professionals, disciplinary researchers, and scholars. The Council on Library and Information Resources' (CLIR) research on scholarly practitioners in data curation has conducted an environmental scan of the state of data curation education in and out of the academy, and an anthropological study of the development needs of research professionals thrust into data curation roles. This project builds upon an existing successful program that brings scholars into libraries to develop a rigorous training program in data curation for individuals with domain expertise, and to propose next steps in the implementation of a sound data management curriculum.

<http://txcdk-v22.unt.edu/icamp/content/icamp-project>
<http://datamanagement.unt.edu/>
<http://www.clir.org/initiatives-partnerships/data-curation>
<http://research.library.unt.edu/datares/>

Monday, Apr. 02, 2012 2:45-3:45 PM Potomac Room

ResourceSync: Towards a Web-Based Approach for Resource Synchronization

Herbert Van de Sompel
Scientist
Los Alamos National Laboratory

Web applications frequently leverage resources made available by remote Web servers. As resources are created, updated, deleted, or moved, these applications face challenges to remain in lockstep with the server's change dynamics. Several approaches exist to help meet this challenge for use cases where "good enough" synchronization is acceptable. But when strict resource coverage or low synchronization latency is required, commonly accepted web-based solutions remain elusive.

Motivated by the need to synchronize resources for applications in the realm of cultural heritage and research communication, the National Information Standards Organization (NISO) and the Open Archives Initiative (OAI) have recently launched the ResourceSync project that aims at designing an approach for resource synchronization that is aligned with the web architecture and that has a fair chance of adoption by different communities.

This project briefing will discuss some motivating use cases and will provide a perspective on the resource synchronization problem that results from initial discussions. It will also give an overview of an experiment in which a synchronization approach was explored to sync with rapidly evolving linked data, and that is based on a server pushing notifications out when its resources change, and recipients of those notifications pulling the changed resources.

Project team: Todd Carpenter (NISO), Bernhard Haslhofer, (Cornell University), Martin Klein (Los Alamos National Laboratory), Nettie Lagace (NISO), Carl Lagoze (Cornell University), Peter Murray (NISO), Michael L. Nelson (Old Dominion University), Robert Sanderson (Los Alamos National Laboratory), Herbert Van de Sompel (Los Alamos National Laboratory), Simeon Warner (Cornell University)

<http://live.dbpedia.org>

Monday, Apr. 02, 2012 2:45-3:45 PM Loch Raven I

Interoperating Requirements for a Media-Specific Repository

James Shulman
President
ARTstor

Randy Stern
Manager of Systems Development
Office for Information Systems, Library
Harvard University

David Germano
Associate Professor, Religious Studies
Director of SHANTI
University of Virginia

William Ying
Chief Information Officer
ARTstor

Shared Shelf, a cloud-based cataloging and image management repository built by ARTstor (in partnership with eight universities and colleges and the Society of Architectural Historians), provides a real world test of how a media-specific repository needs to fit in with the campus ecosystem of library and instructional technology services. Images that are produced locally on a campus need to be cataloged, stored, searched, and used in a variety of ways and in a variety of services on a campus. Institutions have a range of approaches to digital preservation, so a "use repository" like Shared Shelf needs to complement archival preservation processes on the back-end. At the same time, because images are used in so many ways, the repository in which they are stored needs to fit in with the other software utilities that discover and draw upon the content in flexible ways.

This presentation will include a report of the progress on interoperation efforts between Shared Shelf and Harvard's Digital Repository Service, and a discussion of the University of Virginia's needs for APIs to search into and get metadata/data from Shared Shelf. It will also include a discussion of the opportunities and challenges associated with the full range of Shared Shelf's interoperating aims, including the International Image Interoperability Framework (IIIF).

<http://www.artstor.org>
<http://www.sharedshelf.org>
<http://www.uvatibetcenter.org>
<http://www.shanti.virginia.edu>

retroReveal.org: An Open-Source Process for the Discovery and Recovery of Lost or Obscured Content

Joyce L. Ogburn
Dean and University Librarian
University of Utah

Kenning Arlitsch
Associate Dean for IT, Library
University of Utah

Harold M. Erickson
Research Associate, Cardiothoracic
Surgery, Health Sciences Center
University of Utah

New technologies are making it possible to recover content that has been lost or obscured due to human or natural causes. A team at the University of Utah has developed a process that reveals hidden information in various media, including print, microfilm, and photographs. This presentation is a public pre-launch of retroReveal.org, a project supported and hosted by the J. Willard Marriott Library. The process provides automated, forensic-style enhancement of digital images of varying quality from cameras and scanners, uploaded by anyone. Using different algorithms, dozens of surrogate versions are rendered automatically. Users can then select and annotate the version that best reveals hidden aspects of the image. After processing, users may move images to the public upload gallery, copy to another location, or delete them. In addition, retroReveal.org is a community-oriented site that facilitates scholarly collaboration on interpretation of results.

The beauty of this process is the ability to upload and process any digital format. It is an accessible, inexpensive, and highly effective approach to the problem of revealing hidden information. Though originally targeted at archivists, curators, and conservators, during the alpha phase, the retroReveal algorithms have proven useful in archaeological and other scholarly applications. Examples include recovering content from greatly overexposed microfilms of objects that can no longer be accessed; reading through endpapers to a vellum letter used as a book's sewing support; recovering a composer's water-damaged instructions to a publisher concerning musical details of a score; enhancing a poor-quality, low-light aerial photograph that revealed a major archaeo-astronomical complex; visualizing weathered pictographs/petroglyphs; and reading exposure-faded Oregon Trail axle-grease messages on stone. The process also has shown the potential of improving optical character recognition (OCR) to create machine-readable text from image files.

<http://retroreveal.org/>

Monday, Apr. 02, 2012 2:45-3:45 PM Camden Room

The California Digital Library and the Public Knowledge Project Partnership: A New Model of Collaborative Institutional Repository Publishing Services Development

Lisa Schiff
Technical Lead, Access & Publishing
Services, California Digital Library
University of California

Brian Owen
Associate University Librarian
Simon Fraser University

Catherine Mitchell
Director, Access & Publishing Services
California Digital Library
University of California

The California Digital Library (CDL) and the Public Knowledge Project (PKP) have recently joined forces, with the CDL signing on as a major PKP development partner. This relationship has grown out of CDL's recent work to incorporate a customized version of PKP's Open Journal System (OJS) into the back-end submission and publishing system for eScholarship, the University of California's open access institutional repository and publishing platform. This development work marks an important step toward fully integrated, open-source institutional repository and journal publication services, and the CDL and PKP have ambitious plans for extending this work to the larger PKP community. This panel will describe:

- How OJS was customized to meet the needs of eScholarship journals (including user interface modifications, the extension of a single OJS instance to support almost 50 independent journals, PDF generation, and more)
- Which of these and other features may be available in a future release of OJS as a result of this new partnership
- How the PKP development partnership program is shaping the direction of OJS and other PKP scholarly communication services
- How the relationship with PKP is likely to affect future development and service directions for eScholarship
- How this work fits into the larger effort of both of these organizations to refine their services in support of new practices and opportunities within the scholarly publishing environment

<http://www.escholarship.org>
<http://pkp.sfu.ca/>

Opportunities and Challenges for the 21st Century Federal Depository Library Program (FDLP)

James Jacobs
Government Information Librarian
Stanford University

Suzanne Sears
Assistant Dean for Public Services
University of North Texas

David Walls
Preservation Librarian
United States Government Printing Office

The vast majority of all US Government documents published today are "born digital," published electronically and available through the Internet, and will never be printed by the federal government. The lack of a systematic process for capturing, preserving, and disseminating born-digital government information challenges the ability of the Federal Depository Library Program (FDLP) in being able to provide permanent and equal access to online-only government information to all citizens. However, the Government Printing Office (GPO) and the FDLP community have begun to make strides on this most critical issue.

This project briefing will describe several exciting initiatives currently underway to capture, preserve, and provide access to born-digital government information, including GPO's Federal Digital System (FDsys) and web harvesting initiatives, and the agency's partnerships with federal agencies; the CyberCemetery, the Congressional Research Service Reports archive, and robust digitization program and digital repository of the University of North Texas; and the LOCKSS-USDOCS collaborative program.

These projects offer examples of how the FDLP community, in partnership and under formal agreements with GPO, can work collaboratively to assure the long-term preservation of born-digital government information to "keep America informed."

**<http://fdlp.gov>
<http://lockss-usdocs.stanford.edu>
<http://www.gpo.gov/fdsys/>
<http://digital.library.unt.edu/>
<http://govinfo.library.unt.edu/>
<http://digital.library.unt.edu/explore/collections/CRSR/>**

**CNI SPRING 2012 MEMBERSHIP MEETING
BREAKOUT SESSIONS**

**MONDAY
APRIL 2
4:15-5:00 PM**

Monday, Apr. 02, 2012 4:15-5:00 PM Chesapeake Ballroom

Archiving Large Swaths of User-Contributed Digital Content: Lessons from Archiving the Occupy Movement

Howard Besser
Director, Moving Image Archiving &
Preservation MA Program
New York University

David Millman
Director, Digital Library Technology
Services
New York University

Sharon M. Leon
Director of Public Projects
Center for History & New Media
George Mason University

Kristine Hanna
Director, Archiving Services
Internet Archive

Archiving born-digital content from the "Occupy" movement can serve as a prototype for archiving all kinds of user-contributed content. In this presentation, several organizations will discuss the tools and methods they have developed for ingesting, preserving, and offering discovery services to large numbers of digital works where they cannot really rely on the contributors to follow standards and metadata assignment. Topics covered will range from automatic extraction of time-stamp and location metadata (and an empirical analysis of which upload services strip these out), to app development for uploading content along with permission forms, to maintaining lists of frequently-changing URL nodes for web-crawling, to issues in educating content creators in best practices. Speakers will also discuss issues in trying to document a social movement while it is happening.

**<http://activist-archivists.org/>
<http://www.archive.org/details/occupywallstreet>
<http://occupyarchive.org/>**

Monday, Apr. 02, 2012 4:15-5:00 PM Potomac Room

A Feasibility Study of a National Science Foundation (NSF) Open-Access Repository

Sayeed Choudhury
Associate Dean for Research Data
Management
Johns Hopkins University

Mark Cyzyk
Scholarly Communications Architect
Johns Hopkins University

The Johns Hopkins University (JHU), in partnership with the Council on Library and Information Resources (CLIR) and the University of Michigan (UM), has conducted a study of the feasibility of building an open access repository to collect, store, manage, preserve, and make broadly available the reports of research funded by the National Science Foundation (NSF), should NSF choose to implement such a repository. The study examined the technical, policy, and business dimensions of such an undertaking through, first, a series of workshops that convened a diverse range of interested parties and potential stakeholders, including libraries, publishers and professional societies. The analysis team, secondly, considered four categories of approaches that might be taken in the construction of an open access repository of NSF materials:

- Category 1 - Locally installed system(s)
- Category 2 - Large-scale hosted systems
- Category 3 - A federation of systems
- Category 4 - A custom solution

The protocol for evaluation was based upon a framework developed previously through a JHU-led analysis of open-source electronic publishing systems. This presentation will focus on the results of the analysis and recommendations for NSF.

Monday, Apr. 02, 2012 4:15-5:00 PM Loch Raven I

New York University's Implementation of Sakai Open Academic Environment (OAE)

Jennifer Stringer
Director of Academic Technology Services
New York University

The Sakai Open Academic Environment (OAE) project continues to look more broadly than standard learning management system functionality to engage and connect learners around their academic activities, enhancing the contexts for content authoring, sharing, and discovery, and emphasizing authentic assessment. With broad support for the creation and sharing of templates, offering a Widget Standard Development Kit (SDK) and Widget Marketplace, and supporting IMS Learning Tools Interoperability standards, OAE continues to push the boundaries for ways for anyone to straightforwardly contribute new capabilities.

This presentation will discuss New York University's current pilot implementation of the Sakai OAE. It will also detail the roadmap for development of the Sakai OAE as well as give background information on the bodies of community representatives developing it, the OAE Steering Group, the OAE Technical Reference Group (TRG), and the OAE User Reference Group (URG).

<https://oae-community.sakaiproject.org/>
<https://confluence.sakaiproject.org/display/3AK/OAE+Year+2+Roadmap>
<http://chartingatlas.blogspot.com/>

Monday, Apr. 02, 2012 4:15-5:00 PM Loch Raven II

Changing for Excellence: Libraries and IT Experience with Campus Consultants

Deborah Ludwig
Assistant Dean, Libraries
University of Kansas

Paul Farran
Chief of Staff, Information Technology
University of Kansas

In 2011, the University of Kansas hired the consulting firm HURON to undertake an intensive review of how the University conducts business, and to help the institution identify ways to increase its effectiveness. This effort, "Changing for Excellence" (CFE), encompassed campuses in Lawrence, Kansas City, and Wichita and was divided into phases in which Information Technology (IT) and the Libraries have been involved: Assessment and Analysis, Business Case Development, and Implementation.

This presentation will include descriptions of the business cases developed in both the IT and in the Libraries, work with the consultants and campus stakeholders, movement forward into implementation, and the connections between CFE, the University's strategic planning effort (Bold Aspirations), and possible organizational transformation.

<http://chancellor.ku.edu/changingforexcellence/>

Monday, Apr. 02, 2012 4:15-5:00 PM Camden Room

JSTOR: Exploring New Economic Models and Markets for Scholarly Content

Heidi McGregor
Vice President
Marketing & Communications
ITHAKA

People around the world are finding scholarly content online every day, but often do not have ready access. JSTOR, the online home for more than 1,400 academic journals and other content, turns away more than 100 million attempted accesses each year. We are undertaking new access experiments to better understand and quantify the interest in scholarly content globally, and to find new, sustainable economic models for providing access to more people. This presentation will include discussion of the strategy behind new initiatives, including the JSTOR Register & Read BETA, and early lessons and results from these activities.

**<http://about.jstor.org/individuals>
<http://about.jstor.org/rr>**

Competing Priorities: Sustainability, Growth, and Innovation in Digital Collections

Jennifer Riley
Head, Carolina Digital Library and Archives
University of North Carolina at Chapel Hill

Academic libraries with digital collections programs are faced with a difficult task in simultaneously growing capacity, promoting sustainability, allowing room for innovation, ensuring success within an environment that has limited and, in many cases, shrinking resources, and giving staff the tools they need to be effective contributors to technology initiatives. The University Library at the University of North Carolina at Chapel Hill is embarking upon a number of efforts designed to support the transition of digital library work from project to program, a transition that will allow the Library to better address these competing needs.

This presentation will introduce two major aspects of the work in this area. First, in early 2012, a new process was implemented for the allocation of library technology resources and the selection of projects requiring technology support. Through careful design and ongoing assessment, this process should significantly improve the Library's efforts to promote a culture of solid planning and accountability, reduce uncertainty that has historically caused inaction and missed opportunities, clearly determine and communicate ongoing support models, and reinforce the Library's confidence in its ability to live up to the commitments it makes. A key part of this new proposal, review, and approval process is staff from the Library's Carolina Digital Library and Archives (CDLA) department acting as project facilitators, and offering planning and coordination services where necessary to help technology proposals have their best chance of success.

The second area of work to be presented in this session involves the early planning stages of a new technical infrastructure that will provide coordinated and shared support for digital collections, digital humanities, and institutional repository work. The plan is to prioritize digital library initiatives that help build out and implement this technical infrastructure, reduce 'silozation' and promote reuse of content in multiple environments, increase efficiency and sustainability of curated digital collections, and provide tiered access and preservation services. This shift has the potential to re-frame the development work on the Carolina Digital Repository to occupy a more centralized role in the University's digital library efforts.

**CNI SPRING 2012 MEMBERSHIP MEETING
BREAKOUT SESSIONS**

**MONDAY
APRIL 2
5:15-6:00 PM**

Networks and the Paradox of the Active Learner

Gardner Campbell
Director, Professional Development & Innovative
Initiatives; Associate Professor, English
Virginia Polytechnic Institute and State University

Experimentation for the last three years has been conducted with versions of a course called "From Memex to YouTube" with varying populations but roughly the same syllabus and objectives. The course aims to help participants thrive and innovate within the rapid pace of change in information and communications technology (in other words, a fundamental human-computer interface) by introducing them to the powerful visions and conceptual frameworks underlying the development of networked interactive computing.

The primary text used in the course is *The New Media Reader* (MIT, 2003), beginning with Vannevar Bush and ending with Tim Berners-Lee. The populations have been undergraduates at the University of Mary Washington, first-year students at Baylor University, faculty (interdisciplinary) and staff (IT and library) at Baylor and at Virginia Polytechnic Institute, honors program undergraduates at Virginia Polytechnic Institute, and this semester (in a cross-listed course) honors undergraduate and graduate students at Virginia Polytechnic Institute. In addition, for three semesters, the faculty-staff development seminar has been networked among colleges and universities including Rice University, Baylor University, Houston Community College, Benedictine University, the University of California at Berkeley, St. Lawrence University, Whitman College, Virginia Polytechnic Institute, Tulane, the University of South Carolina Upstate, and a group in Second Life including educators across the US as well as in Aruba and Belgium.

This session provides an overview of the seminar design, a description of outcomes and lessons learned over the three years, and a consideration of pedagogical vs. andragogical strategies, with particular attention to difficulties and opportunities outlined in Carroll & Rosen's classic human-computer interaction essay, "Paradox of the Active User."

<http://blogs.lt.vt.edu/vtclis12>
<http://blogs.lt.vt.edu/vtnmfss12/>
<http://blogs.lt.vt.edu/vtnmsf11/>
<http://www.gardnercampbell.net/hrc-grand>
http://courseblogs.atlhub.net/baylor_nms_f10/
http://courseblogs.atlhub.net/baylor_nmfs_f10/
<http://gardnercampbell.wetpaint.com/> (syllabi from the last several iterations)
<http://www.newmediareader.com/>
<http://dl.dropbox.com/u/16760174/Papers/Paradox.pdf>

Monday, Apr. 02, 2012 5:15-6:00 PM Potomac Room

Update on the Activities of the Board on Research Data and Information (BRDI)

Clifford Lynch
Executive Director
Coalition for Networked Information

The National Research Council (NRC) established the Board on Research Data and Information (BRDI) in 2008 with the mission "to improve the management, policy, and use of digital data and information for science and the broader society." As BRDI begins its second term of membership, with Francine Berman (Rensselaer Polytechnic Institute) and Clifford Lynch (CNI) serving as co-chairs, the major activities in the near term will include: a consensus study on the future career opportunities and educational requirements for digital curation, an international data attribution and citation initiative, a sustainability study on publicly funded research databases, an international symposium on intellectual property rights in scientific databases, and of course the continuing meetings of the Board itself. This session will provide an update and overview of the Board's recent and planned activities, including the February 2012 Forum on CODATA-World Data Systems Cooperation.

<http://sites.nationalacademies.org/PGA/brdi/index.htm>

Monday, Apr. 02, 2012 5:15-6:00 PM Loch Raven I

Why Google Scholar Has Trouble Indexing Institutional Repositories

Kenning Arlitsch
Associate Dean for IT Services
University of Utah

Patrick O'Brien
Search Engine Optimization Manager
University of Utah

Google Scholar (GS) has difficulty indexing the contents of institutional repositories (IRs) because most IRs use Dublin Core metadata, which cannot express bibliographic citation information adequately for academic papers. GS's Webmaster Inclusion Guidelines site cautions to "use Dublin Core only as a last resort," and recommends other metadata schemas instead. It also recommends specific guidelines to facilitate crawlers, including writing metadata from the repository database to HTML headers. Surveys of institutional and disciplinary repositories across the United States were conducted and the inquiries revealed indexing ratios to support the hypothesis that IRs that do not follow these metadata and crawl guidelines suffer from a low indexing ratio. Survey results also demonstrate that the low indexing ratio problem cuts across institutions and repository software. Three pilot projects were conducted that transformed the metadata of a subset of papers from USpace, the University of Utah's institutional repository, and examined the results of Google Scholar's harvest. The pilot projects were successful, achieving a 90% indexing ratio.

This presentation will cover the highlights of a paper that is being published in March in *Library Hi Tech*. The broader research initiative emphasizes search engine optimization for all digital repositories, including general digital library collections, and has recently been funded by a 3-year National Leadership Grant from the Institute of Museum and Library Services.

Monday, Apr. 02, 2012 5:15-6:00 PM Loch Raven II

Creating the Digital Preservation Network

James Hilton
Vice President and Chief Information Officer
University of Virginia

The Digital Preservation Network (DPN) will address risk to the very long-term preservation of the scholarly record by creating a federated approach to preservation of academic content. The DPN ecosystem ensures reliable, long-term digital preservation through a federated network of diverse, non-overlapping preservation strategies sustained by committed institutions of higher education, which mitigates the threat of a single point of failure - organizational, technical, physical, or political - jeopardizing centuries of scholarship. This presentation will discuss progress toward the launch of DPN and opportunities to engage in the effort.

<http://digitalpreservationnetwork.wordpress.com/>

Monday, Apr. 02, 2012 5:15-6:00 PM Camden Room

**Advances in Discovery:
An EBSCO Service**

Michael Gorrell
Executive Vice President of Technology
Chief Information Officer
EBSCO Publishing

Discovery Services have emerged to become a key element of libraries' efforts to allow their patrons to satisfy their research needs. Harvesting and indexing millions of scholarly journal articles, books, biographies, reviews, and a vast array of other content types from thousands of sources, allowing users to find the best matches for their needs and presenting this information in a clear and understandable way is a tall order. Challenges include determining relevance for search results, providing users with ways to understand the depth and breadth of the collection being searched, and overall site usability. EBSCO has taken a data driven approach to solving these problems by testing various aspects of its Discovery Service, and using other data mining techniques. This session will describe the various methodologies that have been used and describe ways in which the service has evolved based on these efforts.

Monday, Apr. 02, 2012 5:15-6:00 PM Severn I

Developing a Portal for Geospatial Resources

Alan Darnell
Director, Scholars Portal
University of Toronto

21 university libraries in the Ontario Council of University Libraries have collaborated to develop a portal for delivering numeric and geospatial resources to students and scholars across the province. Using the Data Documentation Initiative (DDI) for numeric data markup and the ISO 19115 metadata standard for geographic information system (GIS) resources, the portal allows users to discover, preview, select, and download data from key government and commercial publishers. The features of the current portal will be presented along with a description of the collaborative model employed for design and development.

<http://geo.scholarsportal.info>

**CNI SPRING 2012 MEMBERSHIP MEETING
BREAKOUT SESSIONS**

**TUESDAY
APRIL 3
9:00-10:00 AM**

Tuesday, Apr. 03, 2012 9:00-10:00 AM Chesapeake Ballroom

Building the Grateful Dead Archive Online: The Golden Road to Unlimited Devotion

Virginia Steel
University Librarian
University of California, Santa Cruz

Robin Chandler
Project Manager
University of California, Santa Cruz

The University of California Santa Cruz (UCSC) Libraries, recipient of a 2009 two-year Institute for Museum and Library Services (IMLS) grant, is building the socially constructed Grateful Dead Archive Online (GDAO) website using Omeka open source software. The Grateful Dead Archive (GDA) represents one of the most significant popular culture collections of the 20th Century and documents the band's activity and influence in contemporary music from 1965 to 1995.

Donated to the UCSC Library in 2008, the GDA contains over 600 linear feet of material including business records, photographs, posters, fan envelopes, tickets, video, audio (oral histories and interviews) and 3-dimensional objects such as stage props and band merchandise. With the release of GDAO in July 2012, the Archive will actively begin collecting artifacts from an enthusiastic community of Grateful Dead fans.

This presentation will discuss the donation of the collection to UCSC; the challenges of merging a traditional archive with a socially constructed one; rights clearances issues and the intellectual property strategy; crawling and harvesting strategies employed for collecting web resources; plugins and workflows supporting data exchange between CONTENTdm and Omeka; and integrating "the crowd" in the curation of user-submitted content preserved by the California Digital Library's Merritt repository. Future directions, such as the integration/development of better curation tools and what the Libraries hope to learn from opening the archive to contributions from a large community of fans, will also be discussed.

**<http://library.ucsc.edu/gratefuldeadarchive/gda-home>
<http://deadcentral.blogspot.com/>**

Tuesday, Apr. 03, 2012 9:00-10:00 AM Potomac Room

Taming the Data Shrew: The National Science Board's Priorities and Recommendations on Scientific Data Management

José-Marie Griffiths
Vice President for Academic Affairs
Bryant University

The increasing ease of gathering large amounts of varied data (including digital data, research specimens, artifacts, etc.), and funding of large-scale collaborative projects, have caused the broad policy issues surrounding the management of scientific and engineering research data to become critically important. The National Science Board (NSB) appointed a Data Policy Task Force to explore these various issues and develop a set of recommendations as to how collected data are shared and managed to ensure broad, timely, and long-term availability and accessibility to the entire research community.

This session will present the results of the Task Force's work and the NSB priorities and recommendations on scientific data management based on the Task Force's input. It will also include a discussion of the impacts, challenges, potentials and possible next steps for higher education, publishing, information technology, scholarly and professional organizations, foundations, and libraries and library organizations, especially as it relates to NSF-funded grants and projects.

Tuesday, Apr. 03, 2012 9:00-10:00 AM Loch Raven I

The Digital Public Library of America Platform

David Weinberger
Co-Director, Harvard Library Innovation
Lab
Harvard University

Paul Deschner
Developer, Harvard Library Innovation Lab
Harvard University

The platform for the Digital Public Library of America (DPLA) is being built to make maximally available what libraries and other curatorial institutions know about works, ideas, relationships, usage patterns, and more. This description is broad and general because the DPLA is a loose federation of major institutions (as well as smaller ones) that has not yet come to full agreement about what specifically it wants to present to end users. The group's charge is to build a metadata server with open APIs that will enable and encourage developers to create applications that cannot be anticipated. The metadata may eventually include records for millions of items and many thousands of collections, all of which do not adhere to established standards. The service becomes more valuable as it is able to present developers with richer and deeper responses that make sense of the jumble of content at which the server points. The problems are large and knotty in every direction.

<http://dp.la/dev>
<http://dp.la/dev/wiki>
<http://dp.la/dev/blog>

Tuesday, Apr. 03, 2012 9:00-10:00 AM Loch Raven II

Transforming Research Support Services

Jennifer Rutner
Senior Analyst
Ithaka S+R

Roger Schonfeld
Director of Research
Ithaka S+R

Rapidly evolving research methods and practices across disciplines are changing the nature of scholars' interactions with service providers such as libraries, computing support centers, humanities centers, scholarly societies, and publishers. As a result, many scholars have become less dependent on traditional library information services, and research support service providers would like to better understand the evolving research practices of their users, in order to transform their services in parallel. Ithaka S+R has launched the Research Support Services for Scholars program to engage scholars and research support professionals in building a deeper understanding of the needs of researchers, the support landscape, current and evolving practices, and the challenges both communities face in conducting and facilitating innovative research. The first disciplines to be covered in this new Ithaka S+R program are history and chemistry.

This session will provide an overview of the Ithaka S+R Research Support Services for Scholars History Project. Presenters will share preliminary research findings, including data gathered through interviews with twenty research support professionals and forty academic historians. In this session, presenters will facilitate discussion about the user research needs of the information services community, organizational challenges that research support organizations face in meeting evolving needs, and effective ways to consider research findings in the ongoing transformation of research support service environments.

<http://www.researchsupportservices.net/>

Tuesday, Apr. 03, 2012 9:00-10:00 AM Camden Room

National Science Foundation (NSF) EarthCube Program Update

Jennifer Schopf
Program Officer
National Science Foundation

The goal of the National Science Foundation (NSF) EarthCube Program is to transform the conduct of Geoscience research by supporting the development of community-guided cyberinfrastructure to integrate data and information for knowledge management. The project is fostering collaboration to address the multifaceted challenges of modern, data-intensive science, and education through an online forum and a series of virtual and face-to-face events.

This session will report on process and outcomes of community efforts to develop approaches to EarthCube, as well as the current activities and plans for the upcoming year.

<http://earthcube.ning.com/>
<http://www.nsf.gov/geo/earthcube/>

Tuesday, Apr. 03, 2012 9:00-10:00 AM Severn I

**Under the Blacklight:
Open Source Content Management for Collaborative Digitization
Projects Involving Mass Digitization of Archival Materials**

Eric C. Weig
Director, Digital Library Services
University of Kentucky

Institutions continue to digitize their unique holdings, with the current impetus to get more and more of that archival content online quickly. Traditionally to enable access, this has been a slow process of relating images to finding aids or digitizing a few select items from a collection. Content management systems have often done a poor job in connecting the finding aid to its digitized objects in meaningful ways.

The University of Kentucky has built upon its extensive experience with mass digitization of newspapers and archival description to envision an approach to mass digitization for archival collections. In the past year, the University has developed an open source content management system utilizing the Blacklight discovery software. It has also developed a method for quickly digitizing and loading complete archival collections described with Archivist Toolkit at the item or folder level while seamlessly integrating the finding aid and digital objects together with the assistance of generated METS files.

This session will include a presentation of Kentucky's new open source content management system as well as an outline of the approach developed to automate all steps in the digitization process beyond image capture and Encoded Archival Description (EAD) creation for archival collections. The automatically generated METS objects will be displayed and discussed as well as a demonstration of additional content encompassed by the new content management system including historic newspapers and oral histories.

Project contributors include Eric Weig, Dr. Michael Slone, Deirdre Scaggs, Mary Molinaro, and Dr. Doug Boyd.

<http://eris.uky.edu>
<http://eris.uky.edu/exploreuk>

**CNI SPRING 2012 MEMBERSHIP MEETING
BREAKOUT SESSIONS**

**TUESDAY
APRIL 3
10:30-11:30 AM**

Tuesday, Apr. 03, 2012

10:30-11:30 AM

Chesapeake Ballroom

Linked Data for Libraries: Why Should We Care? Where Should We Start?

Jennifer Bowen
Assistant Dean, Information Management
Services, River Campus Libraries
University of Rochester

Philip E. Schreur
Head, Metadata Department
Stanford University

Linked data has the potential to transform every aspect of how we create, acquire, and discover information. By creating simple assertions in Resource Description Framework (RDF) and linking them together, a semantic web of data is created. Current library metadata encoded in Machine Readable Cataloging (MARC) is an ideal place to begin this transformation. Its consistency and quality will immediately enrich the Semantic Web and position our data where people are now searching for it.

Linked data is poised to replace MARC as the basis for the new library bibliographic framework. For libraries to fully benefit from linked data, they must learn about it, experiment with it, demonstrate its usefulness, and take a leadership role in its deployment. The eXtensible Catalog Organization (XCO) facilitates these activities by providing open-source software for libraries that is "linked-data-ready". XC software prepares MARC and Dublin Core metadata for exposure to the Semantic Web using a platform based upon the OAI-PMH protocol, incorporating registered vocabularies for Resource Description and Access (RDA) elements and roles, and enabling the creation of linked data that represents Functional Requirements for Bibliographic Records (FRBR) Group 1 entities. XC can play a key role in transitioning libraries from their current record-based system infrastructure to linked data by facilitating the repurposing of metadata, without disrupting existing business processes.

This presentation will include a software demonstration, proposed software architecture for creation and management of linked data, a vision for how libraries can migrate from their current metadata environment to linked data, and an update on XCO progress toward linked data goals.

<http://www.eXtensibleCatalog.org>

Tuesday, Apr. 03, 2012

10:30-11:30 AM

Potomac Room

A Study in SCARLET

Matt Ramirez
Senior Augmented Reality Developer
University of Manchester

The SCARLET project is pioneering augmented reality (AR) using mobile devices to enhance students' use of special collections (SC) in libraries; bringing SCs into the age of the app. AR enables students to simultaneously experience the magic of primary materials while enhancing the learning experience by "surrounding" the object with digitized content. Learning and teaching is embedded at the heart of this innovative project, ensuring the focus remains on the student experience and not the technology. The session will describe current work to evaluate AR's effectiveness in different student groups and suggest other subject areas where this methodology may benefit learning.

<http://teamsscarlet.wordpress.com>

Tuesday, Apr. 03, 2012

10:30-11:30 AM

Loch Raven I

Taking Ownership of Electronic Journals and Books

Alan Darnell
Director, Scholars Portal
University of Toronto

While many libraries are moving to subscription or leasing models for electronic journals and books, the 21 libraries of the Ontario Council of University Libraries have collaborated to build shared facilities and services for housing and managing large collections of licensed content. The Books and Journals service of Scholars Portal contain over 25M digital articles and close to 400,000 electronic books, including digitized print books and born digital commercial ebooks. This session will describe the rationale behind this commitment to physical ownership of digital content and the maturing of the service from providing simple aggregation to supporting long-term digital preservation.

<http://journals.scholarsportal.info>
<http://books.scholarsportal.info>

Curation Practices for Born-Digital and Digitized Newspaper Collections

Martin Halbert
Dean of Libraries
University of North Texas

Katherine Skinner
Executive Director
Educopia Institute

Tyler Walters
Dean of University Libraries
Virginia Polytechnic Institute and State
University

This briefing will highlight and discuss the early findings of a National Endowment for the Humanities-funded project hosted by the Educopia Institute that is documenting and modeling the use of data preparation techniques and distributed digital preservation frameworks to collaboratively preserve digitized and born-digital newspaper collections. US libraries and archives have been digitizing newspapers since the mid-1990s using a highly diverse and ever-evolving set of encoding practices, metadata schemas, formats, and file structures. Increasingly, they are also acquiring born-digital newspapers in an array of non-standardized formats, including websites, production masters, and e-prints.

This project is exploring how existing standards (including the National Digital Newspaper Program's digitization standards) may be elaborated upon and applied to foster the preservation readiness of collections from the last two decades that were digitized according to evolving standards, as well as the born-digital content that institutions are steadily acquiring. This project is also documenting how curators can effectively exchange their preservation-ready content across repository systems, focusing on the use of distributed digital preservation (DDP), a collaborative approach in which content is exchanged and replicated across multiple sites, and actively monitored using various network-driven technologies (e.g., LOCKSS, iRODS, CODA).

This briefing will share initial project results, including the following:

1. A "state of the field" report (based on surveys conducted by the researchers) regarding the challenging collections with which academic libraries are contending, including legacy content from more than two decades of digitization and a wide range of born digital content; and
2. Preliminary recommendations regarding what type and level of preservation preparation for these diverse newspaper collections might be considered essential, and what type and level might be considered optimal.

Climate Data for Our Future: Acquired, Analyzed, Archived

Andreas V. Hense
Professor
Bonn-Rhine-Sieg University
of Applied Sciences

Florian Quadt
Research Scientist
Max Planck Institute for Meteorology

Creating data in the field of meteorology is laborious; some data is generated in the course of a climate simulation, while other data is measured during a weather observation experiment. As a result, meteorologists are developing a culture of publishing and reusing data. This presentation will include a research project aimed at simplifying the meteorological data publication process through the use of a web-based workflow software which is hosted at the site of the long-term archive World Data Center for Climate. This software solution helps to prepare publications by guiding the meteorologist through the necessary metadata review, including details about scientific quality assurance, and by supporting the publisher by carrying out technical quality assurance. The publication procedure ends with the assignment of a Digital Object Identifier (DOI), which is automatically registered using DataCite Web-services.

This project is funded by the German Research Foundation (Deutsche Forschungsgemeinschaft; project partners are the Bonn-Rhein-Sieg University oAS, the University of Bonn, and the German Climate Computing Center (Deutsches Klimarechenzentrum).

<http://umwelt.wikidora.com>

Tuesday, Apr. 03, 2012

10:30-11:30 AM

Severn I

Fembot: Reinventing Scholarly Production and Communication

Karen Estlund
Digital Library Services, Head
University of Oregon

Carol Stabile
Professor, English Department
School of Journalism and Communication
Department of Women's
and Gender Studies
University of Oregon

The Fembot Collective is a collaborative of faculty, graduate students and librarians engaged in research on gender, new media and technology and a platform for communication about related issues. The collective includes faculty and graduate students from North America, the UK, Australia, and Asia and encourages interdisciplinary and international participation. Fembot aims to seize the means of scholarly production. One of its strategies for doing so is the creation of an open access journal, *Ada: A Journal of Gender, New Media, and Technology*, with a re-envisioned model of peer review and tools for multi-modal publication, community and promotion. Not content with creating traditional scholarship that will appear online, Fembot is redefining what scholarly communication means in a digital environment by transforming the concept of the "article" and embracing multi-modal technologies for production and distribution. Fembot is in its initial phase of development with its first issue of *Ada* in progress. The University of Oregon (UO) Libraries is partnering with the Fembot collective to provide support and consultation on digital preservation, presentation, and bibliometrics. Rather than act as a publisher for the content, the Libraries is functioning as a true partner with the expert scholars to transform scholarly publishing, with a librarian serving as a permanent member of the collective's advisory board.

<http://fembotcollective.org>

Tuesday, Apr. 03, 2012

10:30-11:30 AM

Severn II & III

Sustaining Open Source Projects: An Update from DuraSpace

Michele Kimpton
Chief Executive Officer
DuraSpace

Jonathan Markow
Chief Security Officer
DuraSpace

DuraSpace, a not for profit organization, was formed nearly three years ago with the goal of providing a home for open source projects DSpace and Fedora, and with a mission of creating a sustainable model to allow for continued growth and development of the projects. Within this model, DuraSpace also needed to figure out how to transition the organization from a completely grant-based funded operation to an organization that would have a diversified and sustainable business strategy.

Over the years, the organization has made significant progress in developing robust communities for both the DSpace and Fedora repository solutions. Today, there are over 1500 identified installations of DSpace and Fedora platforms globally. The organization has put in place a business strategy that brings in a revenue stream in the form of sponsorship from existing users of DSpace and Fedora, service revenue from its DuraCloud technology, and grant-funded projects for exploratory work and development. DuraSpace also placed several new initiatives on the roadmap for 2012 and 2013 in order to expand services and programs offered to the community. These revenues support the operation, and provide a path to sustainability that does not rely on grant funding.

This presentation will describe the path followed for building a robust open source community, and the strategy put in place for sustaining the not for profit organization; the latter will also cover new services and initiatives on the roadmap.

<http://Duraspace.org>

**CNI SPRING 2012 MEMBERSHIP MEETING
BREAKOUT SESSIONS**

**TUESDAY
APRIL 3
1:00-2:00 PM**

Tuesday, Apr. 03, 2012 1:00-2:00 PM Potomac Room

DMPTool: Phase 2

Andrew Sallans
Head of Strategic Data Initiatives, Libraries
University of Virginia

Todd Grappone
Associate University Librarian for
Digital Initiatives and IT
University of California, Los Angeles

In January 2011, the National Science Foundation (NSF) began to require data management plans for all grant submissions. This major requirement by the NSF was a leading step, now being followed by other funding bodies. In response to these emerging requirements, a group of US libraries and research institutions have formed a partnership to develop a web-based solution called the DMPTool. The partnership's two primary goals are: 1) assisting researchers in development of high-quality and realistic data management plans for proposals, and 2) matching researchers with support resources and personnel at their respective institutions. In the first three months of public use, the tool has had over 1,200 unique users across the US, over 250 institutions represented, and nearly 40 institutions contributing localized resource guidance.

This session will discuss where the project team is now working to advance the tool with further functionality, addition of more rich content and guidance, expansion of the user community, and development of a sustainable governance and business model. The DMPTool project includes the University of California, Los Angeles Library, the UC3 from the California Digital Library, the Smithsonian, University of Virginia Libraries, the University of California, San Diego Library, the University of Illinois at Urbana-Champaign (library and chief information officer), DataONE, and the Data Curation Centre.

<https://dmp.cdlib.org/>

Tuesday, Apr. 03, 2012 1:00-2:00 PM Loch Raven I

Federated Identity, with a Side of Scholar

Renee Shuey
IT Manager, Principal Lead of Identity and
Access Management
Information Technology Services
The Pennsylvania State University

Kenneth Klingenstein
Senior Director
Internet2

This session will provide updates on the continued rapid growth of federated identity, including assurance profiles, integration with social identity, federal government activities, research and scholarship attribute bundles, and interfederation. It will also include coverage of some emerging discussions on integrating scholarly identity into the mix, including author identities, scholarly publishing records, trusted citations, and access controls on research databases, as well as some anticipated developments in the next year that might result in tighter integration of these efforts.

<http://www.incommon.org>

Pay Once, Preservation Forever: A "Paid Up" Cost Model for Long-Term Preservation

Stephen Abrams
Associate Director, University of California
Curation Center
California Digital Library

Digital preservation and curation are rapidly maturing disciplines, able to draw on an increasingly rich set of community best practices, tools, and service providers to ensure the long-term viability and value of the digital assets that thoroughly pervade all aspects of contemporary culture, commerce, science, education, and entertainment. In many ways, however, the most significant risk to that long-term viability is financial, rather than technical. Unlike the conservation of analog materials, the effective preservation of digital resources necessitates ongoing and proactive intervention, and any interruption in these activities could result in irretrievable data loss. In an era of severe budgetary constraints, however, many institutions are finding it difficult to identify and dedicate ongoing funding streams in support of long-term preservation efforts.

To address this concern, the University of California Curation Center (UC3) at the California Digital Library (CDL) has developed a comprehensive cost model for long-term preservation, known as the "Total Cost of Preservation" or TCP model, that can be applied on either a "pay as you go" (PAYG) or a "paid-up" or "pay once, preserve-forever" (POPF) basis. The latter is particularly useful for data produced as research outputs of grant-funded projects. In the absence of a paid-up option, the status of project outputs often becomes problematic when project funding ceases. The TCP model is capable of representing the full economic costs of long-term preservation, but it can be easily customized to consider only specific subsets of those costs as determined by local policy. The POPF option was derived using a standard economic forecasting technique, discounted cash flow (DCF) analysis.

This briefing will review prior work on preservation cost modeling (including the CMDP, DataSpace, KRDS, and LIFE projects), define the conceptual model underlying the UC3 analysis, summarize the derivation of the relevant cost equations, address some specific shortcomings introduced by the reliance on the DCF technique, and illustrate the application of the model in its PAYG and POPF forms.

Tuesday, Apr. 03, 2012 1:00-2:00 PM Camden Room

Too Big to Know

David Weinberger
Co-Director, Harvard Library Innovation Lab
Harvard University

The Internet is probably making us both smarter and stupider, but it's also changing the nature of knowledge itself. In this informal session, David Weinberger will lay out some of the themes of his new book, *Too Big to Know*, to initiate an open discussion about how the changing nature of knowledge affects us, and what we as librarians, technologists, parents, and teachers can do to seize the opportunity to enter a new Renaissance, instead of a new Dark Ages.

Tuesday, Apr. 03, 2012 1:00-2:00 PM Severn I

Absolute Relevance? Ranking in the Scholarly Domain

Tamar Sadeh
Director of Marketing
Ex Libris Group

The greatest challenge for discovery systems is how to provide users with the most relevant search results, given the immense landscape of available content. In a manner that is similar to human interaction between two parties, in which each person adjusts to the other in tone, language, and subject matter, discovery systems would ideally be sophisticated and flexible enough to adjust their algorithms to individual users and each user's information needs. When evaluating the relevance of an item to a specific user in a specific context, relevance-ranking algorithms need to take into account, in addition to the degree to which the item matches the query, information that is not embodied in the item itself.

Such information, which includes the item's scholarly value, the type of search that the user is conducting (e.g., an exploratory search or a known-item search), and other factors, enables a discovery system to fulfill user expectations that have been shaped by experience with Web search engines. This session will focus on the challenges of developing and evaluating relevance-ranking algorithms for the scholarly domain. Examples will be drawn mainly from the relevance-ranking technology deployed by the Ex Libris Primo discovery solution.

Tuesday, Apr. 03, 2012 1:00-2:00 PM Severn II & III

The United States End-of-Term Web Archive

Abbie Grotke
Web Archiving Team Lead
Library of Congress

Kathleen Murray
Post-Doctoral Research Fellow, Libraries
University of North Texas

In the spring of 2008 an ad-hoc collaboration was formed to build a comprehensive archive of the United States Federal Government web domain before, during, and immediately after the transition to a new presidency. The Library of Congress, the Internet Archive, the California Digital Library, the University of North Texas and the Government Printing Office collaborated to assemble a comprehensive list of sites, provide a nomination tool to engage federal documents experts in site selection, and distribute the work of harvesting content. This presentation will include discussion of various aspects of the ongoing collaboration, including recent work to provide researchers access to the archive, which consists of over 3000 sites, and plans which are underway for collecting in 2012 and 2013. The archive will be demonstrated at this session. The speakers will also discuss a two-year grant from the Institute of Museum and Library Services (IMLS) funding research into comparing machine clustering of Web pages to classification by subject matter experts.

<http://eotarchive.cdlib.org/index.html>