http://www.ils.unc.edu/digccurr2009/

April 1-3, 2009 | University of North Carolina at Chapel Hill, NC USA

PRACTICE, PROMISE & PROSPECTS





National Archives and Records Administration (NARA)





UNC

SUPPORTERS

CONFERENCE PROGRAM

DigCCurr2009

- 2 WELCOME
- 4 ACKNOWLEDGMENTS
- 6 PROGRAM AT-A-GLANCE
- 8 KEYNOTE
- 9 SCHEDULE
- **20** OPENING RECEPTION
- **21** CONFERENCE VENUE
- 22 DINNER
- 23 LOCAL INFORMATION

Table of Contents

Welcome

Helen R. Tibbo Christopher (Cal) Lee

Conference Chairs

School of Information and Library Science University of North Carolina at Chapel Hill, NC, USA t is our pleasure to welcome you to beautiful Chapel Hill, North Carolina for DigCCurr2009: Digital Curation Practice, Promise and Prospects. DigCCurr2009 is a component of the "Preserving Access to Our Digital Future: Building an International Digital Curation Curriculum" (DigCCurr) project. DigCCurr is a three-year (2006-2009), Institute of Museum and Library Services (IMLS)-funded project to develop a graduate-level curricular framework, course modules, and experiential components to prepare students for digital curation in various environments (IMLS grant #: RE-05-06-0044).

The project's first symposium, DigCCurr2007: An International Symposium in Digital Curation, was held April 18-20, 2007, attracting nearly 300 attendees from ten countries. Participants explored the definition of digital curation and what skills are necessary for digital curation professionals working in libraries, archives, museums, data centers, and other data-intensive organizations.

DigCCurr2009 continues the same general themes as our symposium two years earlier, focusing on current practice and research surrounding digital curation with a look toward the future, and trends in preparing digital curation professionals. Submissions to DigCCurr2009 took the form of papers, posters and panels, which were peer-reviewed by a panel of international experts. The DigCCurr2009 program was also greatly enhanced by several invited papers and panels by a distinguished body of international speakers. Additionally, we introduced a special panel format for DigCCurr2009 called "Tools and Demos," featuring demonstrations of leading-edge tools to support digital curation work.

In total, DigCCurr2009 features more than 100 presenters from the United States, Canada, Australia, the United Kingdom, Austria, Germany, Greece, Italy, and the Netherlands. Our program includes 18 panels, 27 papers, 13 demos, and 9 posters. John Wilkin, Associate University Librarian for Library Information Technology at the University of Michigan Library and Executive Director of the HathiTrust, is delivering our keynote address, "Building the Universal Library: The Promise and Challenges of HathiTrust."

DigCCurr2009 would not be possible without the support of many people and organizations, including our supporters, listed on the back of your program, as well as our DigCCurr2009 Planning Committee, DigCCurr Advisory Board, and volunteer support from our exceptional students at the School of Information and Library Science at the University of North Carolina at Chapel Hill. A special thank you to all authors and presenters for their substantial contributions to the quality of the conference and proceedings, as well as our superb reviewers.

And lastly, we would like to thank you for joining us in Chapel Hill for DigCCurr2009. We hope you find our program engaging and informative, and that your participation contributes to your profesional interest and work in the area of digital curation.

• • •



igCCurr2009 is the result of the work of many dedicated individuals and groups. Our success would not be possible without their much valued and appreciated support. On behalf of the DigC-Curr2009 conference chairs, Helen Tibbo and Christopher (Cal) Lee, I would like to give a very sincere thanks to all who contributed to the success of DigCCurr2009.

DigCCurr2009 was made possible by the generous support of the following: the School of Information and Library Science, University of North Carolina at Chapel Hill (UNC-CH); the University Library, UNC-CH; the Institute of Museum and Library Services (IMLS); the National Archives and Records Administration (NARA), and the Coalition for Networked Information (CNI).

The rich program presented at DigCCurr2009 is the result of the excellent work of all DigCCurr2009 contributors, as well as dozens of international reviewers. Invaluable guidance and support in shaping the conference was provided by the members of the DigCCurr2009 Planning Committee: Helen R. Tibbo, Cal Lee, and Rachael Clemens, *UNC-CH*; Wendy Duff, *University of Toronto*; Maria Guercio, *University of Urbino*; Seamus Ross, *University of Toronto*; Kenneth Thibodeau, *NARA*; and Elizabeth Yakel, *University of Michigan*.

Hundreds of planning and service hours were needed for this three-day conference. We would like to acknowledge the dedicated students at the School of Information and Library Science (SILS) at *UNC-CH* who contributed their time and skills, including our Carolina Digital Curation Fellows (Master's) – John Blythe, Alexandra Chassanoff, T. Mike Childs, Lisa Gregory, Samantha Guss, Jennifer Mantooth, Katie-Rose Repp, Mark Swails, and Tara Wink; our Carolina Digital Curation Doctoral Fellows – Heather Bowden, Michael Brown, and Sarah Ramdeen; and Amber Cushing, Gardener Treneman, and Danica Cullinan.

ACKNOWLEDGEMENTS

Further, thank you to SILS administrative staff for their invaluable help in conference preparation and promotion, including Scott Adams, Wanda Monroe, Michael Penny, Stephanie Peterson, and Marcia Tauber.

With the help of a dedicated group of SILS students, we produced an outstanding print proceedings for the conference. We would particularly like to thank Sarah Carrier, Rachael Clemens, Lori Eakin, and Lisa Gregory for their assistance in the production of our DigCCurr2009 proceedings volume.

Additionally, a special thanks to Songphan Choemprayong, SILS Ph.D. student, for his exceptional graphic design work for this program as well as our proceedings.

DigCCurr2009 is one component of a three-year, IMLS-funded project. We would like to thank our DigCCurr Advisory Board members for their continuous and valued contributions to the DigCCurr project, including their participation in DigCCurr2009: Stephen Chapman, Harvard University Library; Robin Dale, University of California, Santa Cruz; Raymond van Diessen, IBM Business Consulting Services; Wendy Duff, University of Toronto; Philip Eppard, State University of New York, Albany; Anne Gilliland, University of California, Los Angeles; Maria Guercio, University of Urbino; Hans Hofman, Nationaal Archief Netherlands; Anne Kenny, Cornell University Library; Steven Knight, National Library of New Zealand; Clifford Lynch, CNI; Richard Marciano, UNC-CH; Seamus Ross, University of Toronto; Donald Sawyer, VIE Inc.; Kenneth Thibodeau, NARA; and Elizabeth Yakel, University of Michigan.

Lastly, a very appreciative thank you to Rachael Clemens for her devoted and indispensable work on behalf of DigCCurr2009. She has had her hand in every element of your experience here at DigCCurr2009, along with Helen, Cal and me, and her contributions are enormous.

Carolyn Hank

Program Chair

School of Information and Library Science University of North Carolina at Chapel Hill, NC, USA





National Archives and Records Administration (NARA





Supporters

WEDNESDAY, APRIL 1

Wilson Library University of North Carolina at Chapel Hill

6:00-7:30 PM

Opening Reception

THURSDAY, APRIL 2

William & Ida Friday Center, Chapel Hill, NC

8:00-8:30 Registration

8:30-8:50 Welcome and Introduction

8:50-10:00 Keynote: Building the

Universal Library: The Promise and Challenges of HathiTrust

10:00-10:30 Break

10:30-12:00 Concurrent Session #1

A. Common Workflows: Health and Social Science Data Curation Collaborations (Panel)

B. Comparing Curricula for Digital Library and Digital Curation Education (Panel)

C. Distributed Custodial Frameworks for Archival Preservation (Panel)

D. Funders' Perspectives (Panel)

12:00-1:00 Lunch

1:00-1:30 Poster Session (Atrium) **1:30-3:00** Concurrent Session #2

A. Education for Digital Curation (Paper Session)

B. Recordkeeping and Government Information Curation (Paper Session)

C. Digital Curation of Humanistic, Multimedia Materials: Lessons Learned and Future Directions (Panel)

D. Digital Curation Tools and Demos I

Program at-a-glance

THURSDAY, APRIL 2

3:00-3:30 Break

3:30-5:00 Concurrent Session #3

A. Metadata (Paper Session)

B. Curation of Scientific Datasets: Trends, Current Initiatives, and Solutions (Panel)

C. Technology Learning for Digital Curators (Panel)

D. Digital Curation Tools and Demos II

5:15-8:30 Dinner at the Weathervane at

A Southern Season

FRIDAY, APRIL 3

William & Ida Friday Center, Chapel Hill, NC 8:00-8:30 Coffee, Tea, and Pastries 8:30-10:00 Concurrent Session #4

A. Digital Curation and Preservation Training and Education: A Panel to Consider Options and Intersections (Panel)

B. Snapshot of Digital Preservation in Federal Libraries (Panel)

C. Digital Curation Vignettes: Personal, Academic, And Organizational Digital Information (Panel)

D. Cooperative Approaches to Digital Preservation (Panel)

10:00-10:30 Break

10:30-12:00 Concurrent Session #5

A. Extending the Data Curation Curriculum to Practicing LIS Professionals (Panel)

B. Archives in the Wild (Paper Session)

C. Moving Web Archiving into the Classroom (Panel)

D. Gaps and Persistent Challenges (Panel)

FRIDAY, APRIL 3

12:00-1:00 Lunch

1:00-2:30 Concurrent Session #6

A. Personal Digital Archiving (Panel)

B. Digital Curation Policy Issues (Panel)

C. Scientific and Humanities Data Curation (Paper Session)

D. Digital Curation Research (Paper Session)

2:30-3:00 Break

3:00-4:30 Concurrent Session #7

A. Digital Curation for Cultural Institutions (Paper Session)

B. Skills for Significant Properties: Debating Pragmatics and Philosophy in an Area of Digital Curation (Panel)

C. Change Management (Panel)

D. Digital Curation Tools & Strategies (Paper Session)

4:30-5:00 Closing Thoughts and Future

Directions

ROOM ASSIGNMENTS FOR CONCURRENT SESSIONS

A: SUNFLOWER ROOM

B: AZALEA ROOM

C: MOUNTAIN LAUREL

D: DOGWOOD



Building the Universal Library: The Promise and Challenges of HathiTrust

John P. Wilkin, University of Michigan

HathiTrust, a multi-institutional effort to create the universal library, is bringing together as comprehensive a body of works as possible and doing it in a way to ensure access, permanence, content preservation, and an advanced environment for research. In short, it is an effort born of libraries, working to bring their lasting contributions to bear on the growing body of digital materials available to students and researchers. Much has been said and written about the silo effect of digital libraries, the way that our early technological efforts Balkanized content and failed to capitalize on economies of scale. With the creation of HathiTrust, many of the world's great research libraries will



work together to create a single, comprehensive library without walls. Our partners will work to coordinate their investments both in curating content and in building services, to create a whole greater than the sum of its parts.

John P. Wilkin is the Associate University Librarian for Library Information Technology (LIT) and is the Executive Director of HathiTrust. The LIT Division supports the Library's online catalog and related technologies, provides the infrastructure to both digitize and access digital library collections, supports the Library's web presence, and provides frameworks and systems to coordinate Library technology activities (e.g., authentication and authorization). Reporting units include Core Services, Digital Library Production Service, Library Systems, and Web Services.

Wilkin previously served as the Head of the Digital Library Production Service (DLPS) at the University of Michigan (UM), a position he held from its inception in 1996. DLPS was originally conceived as a federated organization, drawing on resources from UM's major information and technology organizations, including the University of Michigan Library. Although this particular aspect of DLPS has changed, now with all of the department's activities situated within the University Library, and nearly all of its funding on the Library's base

budget, DLPS continues to have strong campus and national relations. Among the units in the DLPS is UM's Humanities Text Initiative, an organization responsible for SGML document creation and online systems that Wilkin founded in 1994.

Wilkin earned graduate degrees in English from the University of Virginia ('80) and Library Science from the University of Tennessee at Knoxville ('86). In 1988 he assumed responsibility for the UM's English and American language and literature collection development, as well as library research support for English faculty and graduate students. Soon after, he implemented a campus-wide service for the analysis of electronic text and encoding text in SGML. In 1992, he began work at the University of Virginia as the Systems Librarian for Information Services, where he shaped the Library's plan for establishing a group of electronic centers, led and provided technical support for those centers, and consulted for the University's Institute for Advanced Technology in the Humanities (IATH) in textual

Schedule

WEDNESDAY, APRIL 1, 2009

Wilson Library, University of North Carolina at Chapel Hill

6:00-7:30 Opening Reception

- Opening: Helen R. Tibbo, Professor and DigCCurr PI, SILS, UNC-Chapel Hill
- Welcome: Sarah Michalak, University Librarian and Associate Provost for University Libraries, UNC-Chapel Hill

THURSDAY, APRIL 2, 2009

William & Ida Friday Center, Chapel Hill, NC

8:00-8:30 Registration (Atrium)

8:30-8:50 Welcome and Introduction (Grumman Auditorium)

- Helen R. Tibbo, SILS, UNC-Chapel Hill
- José-Marie Griffiths, Dean, SILS, UNC-Chapel Hill
- Bernadette Gray-Little, Executive Vice Chancellor and Provost, UNC-Chapel Hill

8:50-10:00 Keynote: Building the Universal Library: The Promise and Challenges of HathiTrust (Grumman Auditorium)

 John Wilkin, Associate University Librarian for Library Information Technology (LIT) and Executive Director of HathiTrust, University of Michigan

10:00-10:30 Break (Note: Posters on display in atrium)

10:30-12:00 1. Concurrent Session

A. Panel: Common Workflows: Health and Social Science Data Curation Collaborations (Sunflower)
Peter Granda and David Thomas, ICPSR, University of Michigan, USA; Chris Grasso, The Fenway
Institute, Fenway Community Health, USA; & Cindy Teixeira, The Roper Center, University of Connecticut, USA

This panel session will describe collaborations between different organizations to establish common technologies and procedures to process and preserve quantitative data in health and the social sciences for dissemination to the research community.

B. Panel: Comparing Curricula for Digital Library and Digital Curation Education (Azalea)

Jeffrey Pomerantz, Barbara M. Wildemuth & Carolyn Hank, SILS, UNC-Chapel Hill, USA

Two related curriculum development projects are currently underway, one concerning digital libraries and one concerning digital curation. This session explores the convergence and divergence of these two federally-funded projects' approaches to curriculum development.

Schedule (continued)

C. Panel: Distributed Custodial Frameworks for Archival Preservation (Mountain Laurel)

Helen. R. Tibbo (Moderator) and Richard Marciano, SILS, UNC-Chapel Hill, USA; Caryn Wojcik, Michigan Historical Center, USA; Mark Conrad, National Archives and Records Administration (NARA), USA; & Eliot Wilczek, Tufts University, USA

In this panel, presenters discuss case studies of preservation environments they are jointly building. These include TPAP (Transcontinental Persistent Archives Prototype), DCAPE (Distributed Custodial Archival Preservation Environments), Fedora / iRODS integration, and iRODS (integrated Rule-Oriented Data Systems) prototypes, and represent distributed custodial frameworks from the federal, state, university, and cyberinfrastructure perspectives.

D. Panel: Funders' Perspectives (Dogwood)

Joel Wurl (Moderator), National Endowment for the Humanities (NEH), USA; Joyce Ray, Institute of Museum & Library Services (IMLS), USA; Neil Grindley, Joint Information Systems Committee (JISC), UK; & Kathleen Williams, National Historical Publications & Records Commission (NHPRC), USA

This panel brings together leaders from four funding agencies in the US and the UK. The panelists will speak on the ways in which their organizations are supporting digital curation initiatives and future directions for funding.

12:00-1:00 Lunch (*Trillium*)

1:00-1:30 Poster Session (Atrium)

- Kelly M. Gordon, Cornell University, USA: A Residential Data Curation Internship: Opportunities and Challenges
- Lisa M. Schmidt, Michigan State University, USA: Preserving Electronic Mailing Lists: The H-Net Archive
- Deborah Keller, US Army Corps of Engineers, HECSA Library, USA & Ann Harrison, Library of Congress, FLICC/FEDLINK, USA: Federal Libraries Digital Preservation Census
- Allen H. Renear, Lauren C. Teffeau, Patricia Hswe, Molly Dolan, Carole L. Palmer, Melissa H.
 Cragin, & John M. Unsworth, Graduate School of Library & Information Science (GSLIS), University of Illinois at Urbana-Champaign, USA: Extending an LIS Data Curation Curriculum to Include
 Humanities Data
- Barrie. E. Hayes, James L. Harroun, & Brenda Temple, UNC-Chapel Hill, USA: Data Management and Curation of Research Data in Academic Scientific Research Environments
- Robert G. Capra, Rachael Clemens, Christopher A. Lee, & Laura Sheble, SILS, UNC-Chapel Hill, USA: Contextual Information from Blogs in Video Digital Curation
- James Waters & Robert B. Allen, College of Information Science and Technology, Drexel University, USA: What Should We Teach about METS in a Digital Preservation Course?
- Lori Eakin, SILS, UNC-Chapel Hill, USA: Sustaining Digital Preservation Organizations: What Discourse Analysis Can Tell Us about Market Demand and Long-Term Survival
- Lisa Gregory, SILS, UNC-Chapel Hill, USA: Getting the Tar Off Our Heels: Moving Forward with Archiving University of North Carolina at Chapel Hill Websites

1:30-3:00 2. Concurrent Session

A. Paper Session: Education for Digital Curation (Sunflower)

Maria Guercio (Chair), University of Urbino, Italy

Magia Krause, School of Information, University of Michigan, USA: Thinking Like a Digital Curator: Creating Internships in the Cognitive Apprenticeship Model

Effective formal learning about digital curation must take place both in the classroom and in the field. This paper discusses how the cognitive apprenticeship model is being applied in the new Preservation of Information specialization at the University of Michigan School of Information to foster learning inside and outside the classroom.

 Stefan Strathmann, Göttingen State and University Library, Germany: Qualification & Education in Digital Curation: the nestor Experience in Germany

Being a relatively new topic in research and education, digital curation is, for a number of reasons, currently not very well covered by university curricula. Within the project "nestor", a transnational partnership of academic institutions in Germany, Switzerland, and Austria, we report on the establishment of a comprehensive qualification program not only based on e-learning tutorials but also on schools, seminars and an (open access) encyclopedia in digital curation.

• Jean Dryden, College of Information Studies, University of Maryland, USA: Educating Archivists about Copyright: How Can We Do It Better?

This paper reports on a recent study of Canadian archivists' copyright knowledge in relation to repository practices in digitizing archival holdings for Internet access. If a good knowledge of copyright is an essential component of digital curation, this study suggests there are some problems in copyright education, and proposes possible ways of addressing them.

 Margaret Henty, Australian National Data Service, Australia: Building Australia's eResearch Capability: The Challenge of Data Management

The Australian National Data Service (ANDS) provides an opportunity to devise a national approach to the provision of skills for improving both institutional and individual eResearch capability. One of the four proposed ANDS programs is Building Capabilities. This paper discusses curriculum development, one of three broad areas of activity under this program.

B. Paper Session: Recordkeeping and Government Information Curation (Azalea)

Anne J. Gilliland (Chair), UCLA, Department of Information Studies, USA

 Kenneth Thibodeau, NARA, USA: The Survival of Records (and Records Management) in the Twenty-First Century

This paper discuss the nature and role of records and records management in relation to the contemporary cyber-landscape, and describes how the principles of records management inform the retention of electronic records and what changes in perspective and method are needed given the digital domain.

Kelly Eubank, North Carolina State Archives; and Jennifer Ricker and Amy Rudersdorf, State
Library of North Carolina, USA: Pathways to Preservation: Digital Curation Strategies in North
Carolina State Government

This paper discusses the impact of digital publishing, e-mail, and electronic records management on the North Carolina State Archives and the State Library of North Carolina, the entities responsible for gathering, providing access to, and permanently storing state agency electronic publications and records in North Carolina.

 Marc Fresko, Inforesight Limited, UK: MoReq2: a European Contribution to the Preservation of Electronic Records

This paper discusses MoReq and its successor MoReq2, European model specifications of requirements for Electronic Records Management systems. This paper concentrates specifically on features that address digital preservation, notably: migration; "components;" automated rendition; import and export; preservation metadata; and XML schema.



C. Panel: Digital Curation of Humanistic, Multimedia Materials: Lessons Learned and Future Directions (Mountain Laurel)

Megan Winget, University of Texas at Austin, School of Information, USA; Caroline Frick, University of Texas at Austin, School of Information & Department of Radio, Television & Film, and Texas Archive of the Moving Image, USA; and Jerome McDonough & Allen Rennear, GSLIS, University of Illinois at Urbana-Champaign, USA

This panel features interactive case studies of practical and theoretical experiences with multimedia and new media collections, providing alternative approaches to traditional models of appraisal, collection development, access, and preservation. Panelists will discuss what it means to curate a collection of multimedia and interactive media, how such collections might be used in the humanities, and what role curators play in creating, preserving, and promoting their use.

D. Digital Curation Tools and Demos I (Dogwood)

Jeffrey Pomerantz (Chair), SILS, UNC-Chapel Hill, USA

This invited demonstration session brings together an international collection of tools for performing and facilitating digital curation in the practice setting, as well as for use in the education setting. Tools to be demonstrated include:

- PLATO and Hoppla: Andreas Rauber, Department of Software Technology and Interactive Systems, Vienna University of Technology, Austria
- Digital Curation Exchange (DCE): Heather Bowden, SILS, UNC-Chapel Hill, USA
- Dioscuri: Jeffrey van der Hoeven, Koninklijke Bibliotheek, the Netherlands
- Preservation Manager and the Universal Virtual Computer: Raymond J. van Diessen, IBM, The Netherlands
- Persistent Digital Archives and Library System (PeDALS): Richard Pearce-Moses, Arizona State Library, Archives, and Public Records, USA

3:00-3:30 Break (Note: Posters on display in atrium.)

3:30-5:00 3. Concurrent Session

A. Paper Session: Metadata (Sunflower)

Wendy Duff (Chair), Faculty of Information, University of Toronto, Canada

Junte Zhang & Jaap Kamps, University of Amsterdam, The Netherlands: Effective Access to Digital Assets: An XML-based EAD Search System

This paper focuses on the question of effective access methods, by developing novel search tools that will be crucial on the massive scale of digital asset repositories. We illustrate concretely why XML matters in digital curation by describing an implementation of a baseline digital asset search system that is fully XML-driven. The system aims to provide better access to archival material through digital finding aids in the Encoded Archival Description (EAD) standard.

 Jewel Ward & Jonathan Crabtree, UNC-Chapel Hill, USA: Integrating Metadata into the NARA Transcontinental Persistent Archive Prototype via the OAI-PMH

This paper describes a collaborative project at UNC-Chapel Hill – an extension of NARA's transcontinental persistent archive prototype (TPAP) data grid with the new integrated Rule Oriented Data System (iRODS). Presented are the results of one part of that project, which is the development of a prototype service by which metadata can be transferred into the

NARA TPAP metadata catalogue (iCAT) via the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), using the Odum Institute Data Archive's (OIDA) Data Document Initiative (DDI) metadata as the test data.

 Jim Bradley, Ball State University Libraries, USA: The Russian Doll Effect: A Case Study in Digital Artifact Recontextualization

This paper explores a specific project to digitize and make available artworks of the Ball State University Museum of Art. By establishing partnerships and maintaining flexible metadata, the portability and recontextualization of digital artifacts (termed by the author as the 'Russian Doll Effect') has been maximized. The case study details the primary context of the digital assets, current recontextualizations, and future directions.

B. Panel: Curation of Scientific Datasets: Trends, Current Initiatives, and Solutions (Azalea)

Heike Neuroth, Max Planck Digital Library, Germany; Jane Greenberg, SILS, UNC-Chapel Hill, USA; John Kunze, California Digital Library, University of California, Oakland, CA, USA; & Alexander Ball, UKOLN, University of Bath, UK

This international panel will report on current initiatives addressing the management of scientific data, focusing on advances and solutions in the curation of datasets. The reporting will take place in the context of recommendations from funding agencies and international councils, and models for data curation such as the DCC Curation Lifecycle Model. The panel will provide recommendations for the scope and form of the effort required to address the challenge of scientific data curation and the implications for digital curation education.

C. Panel: Technology Learning for Digital Curators (Mountain Laurel)

Peter Botticelli and Bruce Fulton, School of Information Resources and Library Science, University of Arizona, USA; ; Richard Pearce-Moses, Arizona State Library, Archives, and Public Records, USA; Tyler Walters, Georgia Institute of Technology, USA; Alvin Hutchinson, Smithsonian Institution Libraries, USA; George Diez, National Library of Education, USA; & Laura Drake Davis, The Library of Virginia, USA

In this panel, we will explore current needs for technology learning in digital curation, and we will examine the role of hands-on learning methods in training digital curators.

D. Digital Curation Tools and Demos II (Dogwood)

Hans Hofman (Chair), Nationaal Archief Netherlands, The Haag, The Netherlands

This invited demonstration session brings together an international collection of tools for performing and facilitating digital curation in the practice setting, as well as for use in the education setting. Tools to be demonstrated include:

- Prometheus and MediaPedia: David Pearson, National Library of Australia, Australia
- ContextMiner: Chirag Shah, SILS, UNC-Chapel Hill, USA
- iRODS: Reagan Moore, SILS, UNC-Chapel Hill, USA
- CONTENTIAM: Geri Bunker Ingram, OCLC Digital Collection Services, USA
- DRAMBORA: Andrew McHugh, Humanities Advanced Technology and Information Institute (HATII), University of Glasgow, UK

5:00-8:30 Dinner at the Weathervane at A Southern Season (Dinner ticket is required)



FRIDAY, APRIL 3, 2009

William & Ida Friday Center, Chapel Hill, NC

8:00-8:30 Coffee, Tea, and Pastries (Atrium)

8:30-10:00 4. Concurrent Session

A. Panel: Digital Curation and Preservation Training and Education: A Panel to Consider Options and Intersections (Sunflower)

Nancy Y. McGovern, ICPSR, University of Michigan, USA; Helen R. Tibbo, SILS, UNC-Chapel Hill, USA; Melissa H. Cragin, GSLIS, University of Illinois at Urbana-Champaign, USA; Joy Davidson, Digital Curation Centre, HATII, University of Glasgow, UK; & Hans Hofman, Nationaal Archief Netherlands, The Haag, The Netherlands

This panel will present a variety of continuing education initiatives in digital curation, digital preservation, and data curation. Panelists will discuss the potential intersections of these initiatives and the opportunities for enhancing synergies and extending resources. The speakers represent curriculum development from a range of digital disciplines and organizational settings and will discuss the implications for sustainability of educational programs within these settings.

B. Panel: Snapshot of Digital Preservation in Federal Libraries (Azalea)

Deborah Keller, US Army Corps of Engineers, HECSA Library, USA; Patricia Murphy, U.S. Department of Agriculture, National Agricultural Library, USA; Christina Dunn, U.S. Department of Education, National Library of Education, USA; Richard Huffine, U.S. Geological Survey, National Library Program, USA; & Claretta Crawford, U.S. Army, Fort Leonard Wood, Bruce C. Clarke Library, USA

In this panel, we represent the variety within the federal library sector as seen through the digital projects being pursued. Panelists include librarians and digital project coordinators from both large and small federal libraries containing diverse subject collections, and address the challenges presented by digitizing different physical media.

C. Panel: Digital Curation Vignettes: Personal, Academic, And Organizational Digital Information (Mountain Laurel)

Joan E. Beaudoin, iSchool, Drexel University, USA; Maria Esteva, Texas Advanced Computing Center, University of Texas at Austin, USA; & Andrea Japzon, iSchool, Drexel University, USA

This panel presents variations on the theme of digital curation by examining the digital information management and preservation practices of three different populations. Personal digital information management, personal collections transferred to institutional repositories, and a digital archiving case in a private organization, offer a wide view of the types of contexts in which digital material is being produced "in the wild."

D. Panel: Cooperative Approaches to Digital Preservation (Dogwood)

Martin Halbert, Emory University, USA; Tyler Walters, Georgia Institute of Technology, USA; Aaron Trehub, Auburn University, USA; Richard Pearce-Moses, Arizona State Library, Archives, and Public Records, USA; & Jonathan Crabtree, UNC-Chapel Hill, USA

In this panel, representatives from four archives – the MetaArchive Cooperative, Alabama Digital Preservation Network, Data-PASS, and the Persistent Digital Archives and Library System – discuss the versatility, low cost, and compelling benefits of using cooperative distributed digital preservation networks to safeguard categories of digital content that

define our culture, identity, and history and that might otherwise be lost as a result of natural disaster, human error, or neglect.

10:00-10:30 Break (Atrium)

10:30-12:00 5. Concurrent Session

A. Panel: Extending the Data Curation Curriculum to Practicing LIS Professionals (Sunflower)

Jan Cheetham, Division of Information Technology, University of Wisconsin-Madison, USA; Melissa
H. Cragin, GSLIS, University of Illinois at Urbana-Champaign, USA; Leslie M. Delserone, University of
Minnesota Libraries, USA; & Marianne Stowell Bracke, Purdue University, USA

This panel presents an overview of and outcomes from the inaugural Summer Institute on Data Curation held at the Graduate School of Library and Information Science at the University of Illinois at Urbana-Champaign. The Institute addresses a growing need for continuing professional development in data curation. Panelists will present their experiences attending the Institute, and how it relates to the current and ongoing data curation activities at their own universities.

B. Paper Session: Archives in the Wild (Azalea)

Richard Marciano (Chair), SILS, UNC-Chapel Hill, USA

• Frank McCown, Computer Science Department, Harding University, USA: Everyone is a Curator: Human-Assisted Preservation for ORE Aggregations

In this paper, we investigate how the aggregations of web resources can be preserved outside of the typical repository environment and instead rely on the thousands of interactive users in the web community and the Web Infrastructure (the collection of web archives, search engines, and personal archiving services) to facilitate preservation. Inspired by Web 2.0 services such as digg, deli.cio.us, and Yahoo! Buzz, we have developed a lightweight system called ReMember that attempts to harness the collective abilities of the web community for preservation purposes instead of solely placing the burden of curatorial responsibilities on a small number of experts.

 Richard J. Cox, School of Information Sciences, University of Pittsburgh, USA: Digital Curation and the Citizen Archivist

This paper provides a conceptual argument for how projects such as DigCCurr ought to consider developing spinoffs for archivists training private citizens how to preserve, manage, and use digital personal and family archives. Rethinking how we approach the public, which will increasingly face difficult challenges in caring for their digital archives, also brings with it substantial promise in informing them about the nature and importance of the archival mission. Can DigCCurr provide tools that can be used for working with the public?

William Bradley Glisson, HATII, University of Glasgow, UK: Use of Computer Forensics in the Digital Curation of Removable Media

The purpose of this paper is to encourage the discussion of the potential place and value of digital forensics techniques when dealing with acquisitions on removable media in the field of digital curation. It examines a basic computer forensics process, discusses a typical file system for removable media, and raises questions about necessary processes and incentives for addressing data capture in the field of digital curation.

C. Panel: Moving Web Archiving into the Classroom (Mountain Laurel)

Molly Bragg, Internet Archive, USA; Edward A. Fox, Department of Computer Science, Virginia Tech, USA; Margaret Hedstrom, School of Information, University of Michigan, USA; & Christopher A. Lee, SILS, UNC-Chapel Hill, USA

Graduate-level professional education has only recently begun to address the role of web archiving within digital collection and digital library development. The members of this panel have led efforts to introduce web archiving into university and K-12 classrooms. Based on their experiences from the educational activities discussed above, members of the panel will

Schedule (continued)

discuss the following questions: What are the main principles, concepts and skills required to appropriately archive web resources? What are the most effective strategies for and biggest challenges associated with incorporating web archiving activities into the classroom?

D. Panel: Gaps and Persistent Challenges (Dogwood)

Raymond J. van Diessen (Moderator), IBM, The Netherlands; Clifford Lynch, Coalition for Networked Information, USA; Donald Sawyer, VIE, Inc., USA; & Kevin Ashley, University of London Computer Centre, UK

Where have we been? Where do we need to go to advance the principles and goals of digital curation? This panel brings together leading researchers and thinkers in the areas of digital curation and digital preservation to respond to this critical question. Panelists will offer their unique and informed perspectives on the scope, extent, relevance, and quality of current digital curation research, practices and literature.

12:00-1:00 Lunch (*Trillium*)

1:00-2:30 6. Concurrent Session

A. Panel Personal Digital Archiving (Sunflower)

Jeremy Leighton John, British Library, UK; Cathy Marshall, Microsoft Research, USA; David Pearson, National Library of Australia, Australia; & Andreas Rauber, Department of Software Technology and Interactive Systems, Vienna University of Technology, Austria

This panel addresses personal digital archiving, which is an essential, but often neglected, arena of digital curation. Panelists will discuss their efforts to understand, support and document personal digital archiving activities, as well as exploring general challenges and opportunities for digital curation professionals.

B. Panel: Digital Curation Policy Issues (Azalea)

Catherine Arnott Smith, School of Library and Information Studies, University of Wisconsin-Madison, USA; Grace Agnew, Rutgers, the State University of New Jersey, USA; Melissa H. Cragin, GSLIS, University of Illinois at Urbana-Champaign, USA; Jean Dryden, College of Information Studies, University of Maryland, USA; & Kristin R. Eschenfelder, School of Library and Information Studies, University of Wisconsin-Madison, USA

This panel's speakers will discuss policy issues associated with the curation of digital materials including privacy, copyright/intellectual property, cultural sensitivity and trustworthiness. Several collection types will be included in the discussion including health information, multimedia collections, scientific data collections and digital library, museum and archival cultural collections.

C. Paper Sesssion: Scientific and Humanities Data Curation (Mountain Laurel)

Donald Sawyer (Chair), VIE, Inc., USA

 Wendy Duff, Faculty of Information, University of Toronto, Canada: "Able To Develop Much Larger and More Ambitious Projects": An Exploration of Digital Projects Teams

This paper reports on a research project which is exploring the prevalence and nature of research teams undertaking digital projects. Drawing upon interview and survey data, it aims to identify the methods and patterns of interaction used by collaborating digital project teams and provide recommendations to support effective and efficient teams.

Jinfang Niu, School of Information, University of Michigan, USA: Documentation Evaluation
 Model for Social Science Data: An Empirical Test

This paper builds on our prior research in which a Documentation Evaluation Model (DEM) for social science data was constructed and hypotheses about impacting factors of perceived documentation quality were proposed. In this paper, results from interviews and a survey were used to validate the model and test those hypotheses. We found the DEM model is valid, and that perceived documentation quality varies with several characteristics of data and is weakly affected by users' absorptive capacity.

David Giaretta, Science and Technology Facilities Council (STFC), Rutherford Appleton Lab, UK:
 Data Access and Long Term Data & Knowledge Preservation for Earth Science: An Overview on Some ESA Initiatives

Earth Observation Space Missions provide continuous surveillance of Earth, producing huge amounts of data every year that need to be processed, elaborated, appraised and archived by dedicated systems. This paper provides an overview on some ESA initiatives carried out in collaboration with European entities and institutions, with the objective of guaranteeing long term data preservation. In particular, the paper focuses on the ESA participation and contribution in the CASPAR project, the PARSE Insight project and the Alliance for Permanent Access to the Records of Science coalition.

D. Paper: Digital Curation Research (Dogwood)

Philip Eppard (Chair), Department of Information Studies, College of Computing and Information, University at Albany, State University of New York, USA

William E. Underwood, Georgia Tech Research Institute, USA: Speech Acts and Electronic Records

All written records of human activity involve actions expressed in records. Archivists identify these acts when they describe records and when they review records for possible restrictions on disclosure. This paper reports the results of an analysis of Presidential records to determine the speech acts conveyed by the records and the role of these speech acts in describing the records. It also proposes a method for automatically recognizing these acts for use in support of archival description and review.

Bernadette G. Callery, School of Information Sciences, University of Pittsburgh, USA: Reconstructing the Digital Past: A Case Study of the Reconstruction of the Lost Pittsburgh Project

The Web-based Pittsburgh Project was administered by the University of Pittsburgh's School of Information Sciences between 1992 and 1996. The site disappeared in 2000 when the School switched servers. Graduate students reconstructed the entire site, and added documentation on the reconstruction process. This case study discusses educational strategies used in the reconstruction process, particularly the introduction of issues of the completeness and authenticity of the restored site, and considers the effectiveness of collaborative tools in the management of a group project.

Leslie Johnston, Library of Congress Office of Strategic Initiatives, USA: Identifying and Implementing Modular Repository Services

In recent work at the Library of Congress, we have been identifying requirements for digital repositories for locally created collections and collections received from partner institutions. Our most basic needs are not surprising: How do we know what we have, where it is, and who it belongs to? How do we get files – new and legacy – from where they are to where they need to be? And how do we record and track events in the life cycle of our files? This paper describes current work at the Library in implementing tools to meet these needs as a set of modular services – Transfer, Transport, and Inventory – that will fit into a larger scheme of repository services to be developed.

2:30-3:00 Break (Atrium)



3:00-4:30 7. Concurrent Session

A. Paper Session: Digital Curation for Cultural Institutions (Sunflower)

Robin Dale (Chair), UC Santa Cruz, USA

 Kyle Elizabeth Bryner & M. Robert Vidrine, Wake Forest University, USA: Web Access for the Museum of Anthropology's Collections

This paper is a case study of a phased project to digitally curate archaeological and ethnographic collections and associated archival materials of the Museum of Anthropology at Wake Forest University and make them freely accessible through the World Wide Web.

 Marina Niero, Ateneo Veneto Campo, Italy: Communicating Archives of Cultural Institutions: Venice as a Case Study

In this paper, we describe the "pilot project" of archival cooperation between cultural institutions in the digitalization process.

 Kelvin L. White & June Abbas, School of Library and Information Studies, University of Oklahoma, Norman, USA: Capturing a Plurality of Perspectives: A Framework for Developing Culturally Sensitive Curriculum and Digital Repositories

This paper presents a six-element framework – conceptual expansion; embeddedness; collaboration; leadership, activism and ethics; sustainability; and reflexivity – developed from a 2008 study of archival education in Mexico. The framework is useful for capturing a plurality of perspectives when developing a culturally sensitive graduate-level digital curation curricular framework and course modules. A case study of the applicability of the framework within digital curation education is presented to illustrate further the necessity of this conceptual approach.

B. Panel: Skills for Significant Properties: Debating Pragmatics and Philosophy in an Area of Digital Curation (Azalea)

Stephen Grace and Sheila Anderson, Centre for e-Research, King's College London, UK; & Christopher A. Lee, SILS, UNC-Chapel Hill, USA

A debate at DigCCurr 2009 will allow educators and practitioners the chance to reflect on the balance of practical and theoretical skills required of digital curators. The area of significant properties is used as a case study for drawing out some of these skills in research and teaching environments. A panel will debate issues, with the audience invited to contribute to the discussion.

C. Panel: Change Management (Mountain Laurel)

Fynnette Eaton, Electronic Archives Records Program (ERA), NARA, USA; Stephen Chapman, Harvard University Library, USA; & Jonathan Crabtree, UNC-Chapel Hill, USA

In this panel, Fynnette Eaton, Change Management Officer for the Electronic Records Archives Program at NARA (2002-2007), will provide an overview of types of issues anyone introducing major changes in an organization will face. Stephen Chapman will discuss the experiences in the Open Collections Program at his institution as they deal with changes in digitization mandates that will change how staff performs its work. Jonathan Crabtree will share experiences that Odum has had during the DataPASS project and migration to the current Dataverse archive distribution and management software.

D. Paper Session: Digital Curation Tools & Strategies (Dogwood)

Adrian Cunningham (Chair), National Archives of Australia, Australia

 Mike Smorul, University of Maryland, College Park, USA: An Implementation of the Audit Control Environment (ACE) to Support the Long Term Integrity of Digital Archives

In this paper, we describe the implementation of the Audit Control Environment (ACE) system that provides a scalable, auditable platform for ensuring the integrity of digital archival holdings. ACE allows for the monitoring of collections on a variety of disk and grid based storage systems. Each collection in ACE is subject to monitoring based on a customizable policy. The released ACE Version 1.0 has been tested extensively on a wide variety of collections in both centralized and distributed environments.

Dimitris Gavrilis, Digital Curation Unit, "Athena" Research Centre, Greece: A Digital Library Service for the Small

This paper presents MOPSEUS, a lightweight digital library service based on the Fedora system and created to address the needs of small libraries without support from technical staff. Hence, MOPSEUS attempts to balance flexibility against ease of installation, configuration and use. The service is available as a standard Java Web servlet, uses no external databases or other systems and can easily be deployed on top of any Fedora installation. Preliminary tests concerning the ease of installation and use are encouraging. We contend that facilitating the introduction of digital library infrastructures in the small may contribute to spreading digital curation practices.

 Robin Rice, EDINA and Data Library, University of Edinburgh, UK: Lessons Learned from the DISC-UK DataShare and Data Audit Framework Implementation Projects

This paper discusses some of the outcomes from two JISC-funded data curation projects in the UK from 2007-2009: DISC-UK DataShare and the Edinburgh Data Audit Framework Implementation project. DISC-UK DataShare involved investigating deposit of research data in institutional repositories including metadata solutions and policy development; the second was about understanding and improving data management practice through partnering with academic departments in the use of the Data Audit Framework.

 David Giaretta, STFC, Rutherford Appleton Lab, UK: Preservation Workflows, Strategies and Infrastructure

The OAIS Reference model provides the de-facto model for digital archives and forms the basis for the effort to produce an international standard for audit and certification of digital repositories. OAIS is however an abstract model; digital repositories must have concrete tools, strategies and appropriate support to implement the requirements which derive form OAIS. This paper introduces and briefly describes a number of fundamental strategies, workflows, and a support infrastructure to enable repositories to follow OAIS and to help them be better positioned for international certification.

4:30-5:00 Closing Thoughts and Future Directions (Dogwood)

- Helen R. Tibbo, DigCCurr Pl, SILS, UNC-Chapel Hill, USA
- Christopher A. Lee, DigCCurr Co-PI, SILS, UNC-Chapel Hill, USA

• • •

Opening Reception

Wednesday, April 1, 2009 6:00 pm-7:30 pm

The Louis Round Wilson Library, University of North Carolina at Chapel Hill

Overview:

The DigCCurr2009 Opening Reception is sponsored by the University Library at the University of North Carolina at Chapel Hill (UNC-CH). Opening comments and welcome by Helen R. Tibbo, DigCCurr Pl and Professor, School of Information and Library Science, UNC-CH, and Sarah Michalak, University Librarian and Associate Provost for University Libraries, UNC-CH. Hors d'oeuvres and drinks to be served.

Transportation:

Buses will be running between the conference hotels and Wilson Library from 5:15 pm until 8:30pm.

Conference hotels served by the buses are: Courtyard by Marriot; Hampton Inn & Suites; Holiday Inn; & Holiday Inn Express. For guests staying at the Carolina Inn, Wilson Library is a brief, five minute walk from your hotel.

About the Venue:

The Louis Round Wilson Library, located on Polk Place, is the home of the University Library Special Collections at the University of North Carolina at Chapel Hill. In the spirit of the University's mission, with its emphasis on the University of North Carolina at Chapel Hill being built by the people of North Carolina and serving all people of the state, all of the special collections in the Louis Round Wilson Library are open to all users.

The Manuscripts Department (including the Southern Historical Collection, the Southern Folklife Collection, and University Archives) and the North Carolina Collection (including the Research Library, Gallery, and Photographic Archives) provide rich primary resources, southern material, and North Caroliniana, in a variety of formats. The Rare Book Collection offers early manuscripts, printed books, and visual materials covering a wide range of human endeavor. The University Library's Digital Library, Documenting the American South, has its offices and digital publishing center here. The Special Collections Conservation Lab provides treatment for the Wilson Library collections. Each collection displays items from its holdings through a regular program of exhibitions.

Wilson Library served as the main University Library from 1929 until the opening of the Walter Royal Davis Library in 1984 and continues to provide temporary quarters for other library units. Wilson Library currently provides quarters for the Music Library, the Maps Collection, and the Zoology Library. Wilson Library also houses the Pleasants Family Assembly Room and a master classroom.

For more information on Wilson Library, please see: http://www.lib.unc.edu/wilson/



Conference Venue

Dinner

Thursday, April 2, 2009
5:30 pm-8:30 pm
The Weathervane at A Southern Season

Overview

The DigCCurr2009 Dinner will be held at the Weathervane at A Southern Season. Please remember to bring your dinner ticket provided in your conference packet. For those who did not sign up for dinner at the time of online registration, we are unable to take any on-site dinner reservations.

Schedule:

5:00-6:00: Buses depart from Friday Center to A Southern Season 5:30-7:00: Cash bar and hors d'oeuvres at the Weathervane; You are also welcome to shop in A Southern Season.

7:00: Seating for dinner

8:30-9:30: Buses depart for all conference hotels

About A Southern Season:

A Southern Season was founded in 1975 in Chapel Hill, North Carolina. At nearly 60,000 square feet, it is one of the largest specialty gourmet markets in the United States. A Southern Season has been featured in *The New York Times*, *Travel & Leisure*, *Southern Living*, and *Gourmet Retailer*, and profiled in a "Food Finds" segment on *FoodTV* Network, and featured in a "Blue Ribbon" segment on *Turner South*. In addition to a large selection of gourmet food and cookware, A Southern Season also offers state-of-the-art cooking classes and demonstrations featuring visiting culinary experts and celebrity chefs.

For more information on A Southern Season, please see: http://www.southernseason.com/default.asp

About the Weathervane:

The Weathervane is A Southern Season's acclaimed, fine-dining restaurant. The Weathervane, under the direction of Executive Chef and general Manager, Patrick Cowden, features an ever-changing menu of creative seasonal fare using many of the store's specialty products.

For more information on the Weathervane, please see: http://www.southernseason.com/cafe.asp

0 0 0

Local Information

Welcome to the University of North Carolina at Chapel Hill. This page contains information to facilitate your stay in Chapel Hill and DigCCurr2009-provided transportation to/from the conference venue. Please check with your hotel for more information on local events, dining, and other transportation options, including hotel shuttle services and Chapel Hill Transit information.

Getting Around:

DigCCurr2009 will provide limited transportation to/from conference locations through the Carolina Livery.

Wednesday, April 1, 2009, Wilson Library, UNC-CH:

5:00 pm-8:30 pm Buses will be running between conference hotels (except Carolina Inn) and Wilson

Library approximately every 15 to 30 minutes.

Thursday, April 2, 2009, Friday Center, UNC-CH:

7:30 am-9:00 am Buses will be running between conference hotels (except Carolina Inn) and Friday

Center approximately every 15 to 30 minutes.

5:00 pm- 6:00 pm Buses will be running between Friday Center and EITHER A Southern Season (for

conference dinner attendees) or conference hotels (except the Carolina Inn).

8:30 pm-9:30 pm Buses will be running between A Southern Season and conference hotels.

Friday, April 3, 2009, Friday Center, UNC-CH:

7:30 am-9:00 am Buses will be running between conference hotels (except Carolina Inn) and Friday

Center approximately every 15 to 30 minutes.

Chapel Hill Transit

Consider Chapel Hill Transit for exploring Chapel Hill, and travel between your hotel and conference venue when Carolina Livery is unavailable. Chapel Hill Transit is a safe, reliable and fare-free bus service serving the town of Chapel Hill with over 20 routes. Please check with your hotel for information on bus routes and stops available near you.

For more information on Chapel Hill transit, see http://www.townofchapelhill.org/index.asp?NID=72

About Chapel Hill and the University of North Carolina

The University of Carolina at Chapel Hill, or, simply, Carolina, is located in the Piedmont region of North Carolina. Carolina was the nation's first state university to open its doors and the only public university to award degrees in the 18th century. Today, Carolina enrolls over 27,000 graduate and undergraduate students. Nicknamed the southern part of heaven, Chapel Hill is home to about 51,000 residents. It is part of the "Research Triangle," a region anchored by Chapel Hill; Durham, the home of Duke University; and Raleigh, North Carolina's capital city and home of North Carolina State University.

Notes