



# **Coalition for Networked Information Fall 2019 Membership Meeting**

December 9-10, 2019

Washington, DC

**#cni19f**

Online, mobile-friendly meeting schedule:  
[cnifall2019membershipmeeting.sched.com](http://cnifall2019membershipmeeting.sched.com)

## **CNI Code of Conduct**

CNI is committed to maintaining a welcoming and inclusive environment for inquiry, constructive disagreement, and intellectual freedom and honesty. We do not tolerate personal attacks, harassment of any kind, violence, or disruptive behavior. Please be respectful of our community's diversity and generous of others' views. If you have concerns, please talk to a member of the CNI staff. In case of emergency, dial 911.

**[cni.org](http://cni.org)**

# CNI Fall 2019 Membership Meeting

---

## DECEMBER 9 • MONDAY • Schedule at a Glance

---

8:30am – 11:00am	<b>Executive Roundtable (by prior registration only)</b> <i>Moderators: Clifford Lynch</i>	Governor's
11:00am – 12:15pm	<b>Meet and Greet and Orientation for First-Time Attendees</b> <i>Speakers: Joan K. Lippincott, Clifford Lynch</i>	Hampton
11:00am – 4:00pm	<b>Registration Opens</b>	Blue Room Prefunction
12:15pm – 1:00pm	<b>Break</b>	Blue Room Prefunction
1:00pm – 2:15pm	<b>Opening Plenary: Overview of the 2019-20 CNI Program Plan</b> <i>Speakers: Clifford Lynch</i>	Regency Ballroom (West Wing)
2:15pm – 2:30pm	<b>Break</b>	Blue Room Prefunction
2:30pm – 3:30pm	<b>1.1 Artificial Intelligence: Impacts and Roles for Libraries</b> <i>Speakers: Jason Griffey, Keith Webster</i>	Hampton
2:30pm – 3:30pm	<b>1.2 Watchful Eyes on Digital Preservation Action</b> <i>Speakers: Laura Alagna, Gaelle Bequet, Peter Burnhill, Dan Noonan, Jessica W. Meyerson, Jane Mandelbaum, Trevor Owens</i>	Capitol
2:30pm – 3:30pm	<b>1.3.1 Community Development Model for Digital Community Archives</b> <i>Speakers: Nathan Hall</i>	Calvert
	<b>1.3.2 Challenges and Rewards of Community-Engaged Collection-Building: The Lehigh Valley Engaged Humanities Consortium Digital Archive</b> <i>Speakers: Janna Avon, Nora Egloff, Charlotte Nunes</i>	
2:30pm – 3:30pm	<b>1.4.1 Data Sharing from the Ground Up: Building Data Communities</b> <i>Speakers: Danielle Cooper, Rebecca Springer</i>	Governor's
	<b>1.4.2 The Evolution of a Research Data Management and Curation Program: Candid Reflections and Considerations</b> <i>Speakers: Moira Downey, Sophia Lafferty-Hess</i>	
2:30pm – 3:30pm	<b>1.5 Perspectives on Digital Scholarship Programs</b> <i>Speakers: Joan K. Lippincott</i>	Executive (West Wing)
2:30pm – 3:30pm	<b>1.6 Maps, Gaps, Landscapes, and Ecosystems: Sorting Out the State of Online Scholarly Publishing</b> <i>Speakers: Katherine Skinner, David Lewis, Terry Ehling</i>	Congressional A (West Wing)
2:30pm – 3:30pm	<b>1.7.1 Experimenting on a Digital Strategy</b> <i>Speakers: Laurie Allen, Meghan Ferriter, Leah Weinryb Grohsgal, Jaime Mears</i>	Congressional B (West Wing)
	<b>1.7.2 Accessibility Task Force: Determining Compliance and Organizing Action</b> <i>Speakers: Claire DeMarco, Suzanne Wones</i>	
3:30pm – 4:00pm	<b>Break</b>	Blue Room Prefunction
4:00pm – 5:00pm	<b>2.1.1 Ready or Not: Here Comes Voice Search</b> <i>Speakers: Tim Smith, Twila Camp</i>	Hampton
	<b>2.1.2 Can We Talk? Adding a Smart Assistant Interface to Library Services</b> <i>Speakers: Lisa Smith, Greg Davis</i>	
4:00pm – 5:00pm	<b>2.2.1 Building and Sustaining Community Infrastructure: An Update from the Research Organization Registry (ROR)*</b> <i>Speakers: John Chodacki</i>	Capitol
	<b>2.2.1-5 Short Updates I</b>	

4:00pm – 5:00pm	<b>2.2.2 A Progress Report on Open Repositories in Canada*</b> <i>Speakers: Geoffrey Harder</i>	Capitol
	<b>2.2.3 Shared Repository Infrastructure: Two Years Later*</b> <i>Speakers: Jimmy Ghaphery</i>	
	<b>2.2.4 Shareyourpaper.org: Simplifying Self-Archiving and Cutting the Cost of Mediated Repository Deposit*</b> <i>Speakers: Leila Bella Sterman</i>	
	<b>2.2.5 Crowdsourcing Inputs and Outputs of a Digital Photo Archive*</b> <i>Speakers: Theresa Westbrook</i>	
4:00pm – 5:00pm	<b>2.3.1 Developing a Community-Based Strategic Agenda for the Transformation of Archival Discovery and Delivery</b> <i>Speakers: Mark Matienzo, Audra Eagle Yun, Tom Cramer, Hillel Arnold</i>	Calvert
	<b>2.3.2 Webrecorder, Web Archiving for All: Past, Present and Future</b> <i>Speakers: Ilya Kreymer</i>	
4:00pm – 5:00pm	<b>2.4.1 No One is Using That Anymore: Assessing the Impact of Digital Availability on Print Usage</b> <i>Speakers: Thomas H. Teper</i>	Governor's
	<b>2.4.2 Creating a Virtual Reading Room at UC San Diego Library</b> <i>Speakers: Roger Smith</i>	
4:00pm – 5:00pm	<b>2.5 Human-Centered Artificial Intelligence: The Landscape of Autonomy</b> <i>Speakers: Ben Shneiderman</i>	Executive (West Wing)
4:00pm – 5:00pm	<b>2.6.1 Center for Research Data and Digital Scholarship at the University of Colorado Boulder</b> <i>Speakers: Thea Lindquist, Thomas Hauser</i>	Congressional A (West Wing)
	<b>2.6.2 Connecting Digital Scholarship: Building Communities of Support at University of Michigan</b> <i>Speakers: Joe Bauer, Anne Cong-Huyen</i>	
4:00pm – 5:00pm	<b>2.7 Data Curation Network Update</b> <i>Speakers: Timothy M. McGeary, Cynthia Hudson Vitale, Lisa Johnston</i>	Congressional B (West Wing)
5:00pm – 5:15pm	<b>Break</b>	Blue Room Prefunction
5:15pm – 6:15pm	<b>3.1 Memory Institutions and Deep Digital Disruption: Beyond the Technical Challenges of Born-digital Preservation</b> <i>Speakers: Clifford Lynch, Carol Mandel</i>	Hampton
5:15pm – 6:15pm	<b>3.2.1-5 Short Updates II</b> <b>3.2.1 Connect, Collaborate, and Contribute in the Research Commons: Reflections from Three Years of Service Delivery*</b> <i>Speakers: Meris Longmeier</i>	Capitol
	<b>3.2.2 Teaching Data Curation for Reproducibility (Data CuRe)*</b> <i>Speakers: Limor Peer</i>	
	<b>3.2.3 Piloting Digital Scholarship Support at the Library of Congress*</b> <i>Speakers: Eileen Jakeway</i>	
	<b>3.2.4 Sprinting Toward a Lab: Network Building, Knowledge Sharing and Transforming Communities in Galleries, Archives, Libraries and Museums through a Book Sprint*</b> <i>Speakers: Caleb Derven</i>	
	<b>3.2.5 InstantILL: Simplifying Content Delivery With or Without Subscriptions*</b> <i>Speakers: Tina Baich</i>	
5:15pm – 6:15pm	<b>3.3.1 Bringing Computational Access to Book-length Documents Via an ETD Pilot</b> <i>Speakers: William Ingram</i>	Calvert
	<b>3.3.2 Designing a Migration Path: Final Report and Recommendations</b> <i>Speakers: Aaron Birkland, Jennifer Gilbert, Tim Shearer, David Wilcox</i>	
5:15pm – 6:15pm	<b>3.4.1 Scanning for Science: Astronomy Data Rescue as a Learning Opportunity</b> <i>Speakers: Elisabeth Long</i>	Governor's

5:15pm – 6:15pm	<b>3.4.2 The Texas GeoData Portal: A New System for Enhancing Access to Geospatial Data</b> <i>Speakers: Michael Shensky</i>	Governor's
5:15pm – 6:15pm	<b>3.5.1 Data Doubles: Student Perceptions of Privacy and Learning Analytics in Higher Education</b> <i>Speakers: Andrew Asher, Michael Perry</i>	Executive (West Wing)
	<b>3.5.2 Skills, Knowledge, and Values for Prioritizing Privacy in Library Learning Analytics</b> <i>Speakers: Lisa Hinchliffe, Kyle M. L. Jones</i>	
5:15pm – 6:15pm	<b>3.6.1 Into the Dataspace: Data Science Services on the Ground</b> <i>Speakers: Mike Nutt</i>	Congressional A (West Wing)
	<b>3.6.2 Encouraging the Ecosystem: Growing Unique Partnerships Through Data Science</b> <i>Speakers: Joan Peckham, Karim Boughida</i>	
5:15pm – 6:15pm	<b>3.7 Flipping Open Access Away from APCs</b> <i>Speakers: Nick Shockey, Amy Buckland, Ellen Dubinsky, Nicky Agate</i>	Congressional B (West Wing)
5:15pm – 6:15pm	<b>3.2.1-5 Short Updates II</b>	Capitol
6:15pm – 7:30pm	<b>Reception</b>	Blue

# CNI Fall 2019 Membership Meeting

---

## DECEMBER 10 • TUESDAY • Schedule at a Glance

---

7:30am – 9:00am	<b>Breakfast</b>	Blue
7:30am – 2:30pm	<b>Registration Opens</b>	Blue Room Prefunction
9:00am – 10:00am	<b>4.1 The New Normal: Why Libraries are Teaching AI, ML, DH, NLP, VR...</b> <i>Speakers: Harrison Dekker, Matt Burton, Indrani Mandal, Vicky Steeves, Tim Dennis</i>	Hampton
9:00am – 10:00am	<b>4.2.1 Guerrilla Governance and Program Review: Organizational Transformation in Collaborative Models</b> <i>Speakers: Joanne Kossuth, Bradley Daigle</i>	Capitol
	<b>4.2.2 Responsible Operations: Shaping a Community Research Agenda</b> <i>Speakers: Sarah Shreeves, Thomas Padilla</i>	
9:00am – 10:00am	<b>4.3.1 Evolving with Agile into a New Technical Landscape: Cultivating a Sustainable DAMS Ecosystem</b> <i>Speakers: Dave Ronn, Megan Will</i>	Calvert
	<b>4.3.2 Developing an Open Source Digital Scholarly Research System: Local and Global Possibilities</b> <i>Speakers: Ray Uzwyshyn</i>	
9:00am – 10:00am	<b>4.4.1 An Institutional Research Data Discovery Tool: Open Source Technology and Cross-University Collaboration</b> <i>Speakers: Patty Hinegardner, Na Lin</i>	Governor's
	<b>4.4.2 Mobilizing Computable Biomedical Knowledge, and Making it FAIR</b> <i>Speakers: Marisa Conte, Peter Boisvert</i>	
9:00am – 10:00am	<b>4.5.1 Electronic Lab Notebooks: Implementation, Evaluation, and Lessons Learned</b> <i>Speakers: Harish Maringanti, Xuemao Wang, Daureen Neddill</i>	Executive (West Wing)
	<b>4.5.2 Organizational Strategies to Support Emerging Needs in Areas Such As: Data Science, Technology Rich Spaces, and Experiential Learning</b> <i>Speakers: David Woodbury, Mira Waller</i>	
9:00am – 10:00am	<b>4.6.1 Experimenting with a Machine Generated Annotations Pipeline</b> <i>Speakers: Joshua Gomez</i>	Congressional A (West Wing)
	<b>4.6.2 A Demonstration of Annotation Interoperability</b> <i>Speakers: Mark Patton, Sayeed Choudhury</i>	
9:00am – 10:00am	<b>4.7.1 The End of a Statewide Digital Preservation System: Coping with the Fallout</b> <i>Speakers: Fletcher Durant, Todd Digby</i>	Congressional B (West Wing)
	<b>4.7.2 Project Canopus: Rethinking Preservation Infrastructure in Toronto, Ontario, and Beyond</b> <i>Speakers: Steve Marks</i>	
10:00am – 10:30am	<b>Break</b>	Blue Room Prefunction
10:30am – 11:30am	<b>5.1.1 Building a Digital Preservation Strategy Across a Broad University System</b> <i>Speakers: Edson Smith, Mary Elings, Todd Grappone</i>	Hampton
	<b>5.1.2 Toward Collaborative Models for Sustaining Digital Scholarship</b> <i>Speakers: Katrina Fenlon</i>	
10:30am – 11:30am	<b>5.2.1 Fedora 6 and the Oxford Common File Layout*</b> <i>Speakers: Andrew Woods</i>	Capitol
	<b>5.2.2 IIIF: What's New and What's Next with A/V and Discovery*</b> <i>Speakers: Josh Hadro</i>	
	<b>5.2.3 It Took a Village: The Evolution of Samvera*</b> <i>Speakers: Carolyn Caizzi</i>	

10:30am – 11:30am	<b>5.2.4 Update on Public Access Submission System (PASS)*</b> <i>Speakers: Hanh Vu</i>	Capitol
	<b>5.2.5 Acknowledging Core Facilities and Collections Use with ORCID*</b> <i>Speakers: Eric Olson</i>	
10:30am – 11:30am	<b>5.3.1 InvenioRDM: A Collaborative Next-Generation Research Data Management and Repository Solution</b> <i>Speakers: Thomas Morrell, Kristi Holmes</i>	Calvert
	<b>5.3.2 Opportunities for Academic Libraries to Shape National Approaches to Research Data Management: A Canadian Perspective</b> <i>Speakers: Susan Haigh, Lee Wilson, Jason Brodeur</i>	
10:30am – 11:30am	<b>5.4.1 Visualizing Use and Performance Data from a Global Cross-platform Set of Institutional Repositories</b> <i>Speakers: Jonathan Wheeler, Minh Pham, Nikolaus Nova Parulian, Kenning Arlitsch</i>	Governor's
	<b>5.4.2 North Broad Press: A Collaborative Library/Press Publishing Program</b> <i>Speakers: Mary Rose Muccie, Annie Johnson</i>	
10:30am – 11:30am	<b>5.5 Refreshing the Agenda for Collaboration: Library, IT, and New Partners</b> <i>Speakers: Clifford Lynch, Joan K. Lippincott</i>	Executive (West Wing)
10:30am – 11:30am	<b>5.6.1 Collaborative Empowerment, Empowering Collaboration: The Carpentries at the University of Toronto</b> <i>Speakers: May Chan</i>	Congressional A (West Wing)
	<b>5.6.2 Python Camp: Meeting the Demand for Computational Skills Through Open Technology and Reusable Curriculum</b> <i>Speakers: Lorena A. Barba, Hannah Sommers, Megan Potterbusch, Laura Wrubel</i>	
10:30am – 11:30am	<b>5.7 The Future of OA: The Impact of Open Access on Readership and Subscription Decisions</b> <i>Speakers: Jason Priem, Heather Piwowar</i>	Congressional B (West Wing)
11:30am – 1:00pm	<b>Lunch</b>	Blue
1:00pm – 2:00pm	<b>6.1 Machine Learning in Research Libraries: A Snapshot of Projects, Opportunities and Challenges</b> <i>Speakers: Zheng (John) Wang, Elizabeth Lorang, Harish Maringanti</i>	Hampton
1:00pm – 2:00pm	<b>6.2.1 Modern Endangered Archives Program at the UCLA Library*</b> <i>Speakers: Rachel Deblinger</i>	Capitol
	<b>6.2.2 The Foundations of Discovery: A Short Summary of the Assessment of the Impacts of CLIR's Cataloging Hidden Collections Program, 2008-2019*</b> <i>Speakers: Joy M. Banks</i>	
	<b>6.2.3 Reaching the Researchers: Using Geographic and Chronological Metadata to Facilitate Access to New Acquisitions*</b> <i>Speakers: Christian Casey</i>	
	<b>6.2.4 Preservation of Electronic Government Information (PEGI) Project Next Steps*</b> <i>Speakers: Roberta Sittel</i>	
	<b>6.2.5 Lever Press Update*</b> <i>Speakers: Beth Bouloukos</i>	
1:00pm – 2:00pm	<b>6.3.1 Pay to Play: Licensing Local Television News Content</b> <i>Speakers: Morgan Gieringer</i>	Calvert
	<b>6.3.2 Unlocking Opportunity: Using the JSTOR Platform to Get Library Special Collections into the Research Workflow Without a Paywall</b> <i>Speakers: Bruce Heterick</i>	
1:00pm – 2:00pm	<b>6.4.1 Release of Digital Preservation Risk Analysis Documents by the National Archives (NARA)</b> <i>Speakers: Leslie Johnston</i>	Governor's
	<b>6.4.2 Making Access Happen: New Projects at the National Archives</b> <i>Speakers: Andrew Wilson</i>	
1:00pm – 2:00pm	<b>6.5 Update on Funding Possibilities, Priorities, and Trends</b> <i>Speakers: Josh Sternfeld, Lucy Barber, Becca Quon, Ashley Sands</i>	Executive (West Wing)

1:00pm – 2:00pm	<b>6.6.1 When Research Data Requires Controls: Institutional Support for Regulated Research Environments</b> <i>Speakers: Jeremy Frumkin</i>	Congressional A (West Wing)
	<b>6.6.2 Health Sciences Data Archive</b> <i>Speakers: Mara Blake, Sayeed Choudhury</i>	
	<b>6.6.3 Almost Open: Benefits, Challenges, and Design of an Authorized-access Research Data Enclave</b> <i>Speakers: Jeffrey Spies, Rick O. Gilmore</i>	
1:00pm – 2:00pm	<b>6.7.1 The Academic Library as IT Partner: Supporting Sponsored Research at Auburn University</b> <i>Speakers: Denise Baker, Mallory Lucier-Greer, Aaron Trehub</i>	Congressional B (West Wing)
	<b>6.7.2 Redesigning the Researcher Library Experience: Case Studies, Key Questions</b> <i>Speakers: Tom Hickerson, John Brosz</i>	
1:00pm – 2:00pm	<b>6.2.1-5 Short Updates IV</b>	Capitol
2:00pm – 2:15pm	<b>Break</b>	Blue
2:15pm – 3:30pm	<b>Closing Plenary: Forgetting and Being Forgotten: Growing Up in a Digital Era</b> <i>Speakers: Kate Eichhorn</i>	Regency Ballroom (West Wing)

**PLENARY SESSIONS**



---

**DECEMBER 9 • MONDAY**

---

1:00pm – 2:15pm

**Opening Plenary: Overview of the 2019-20 CNI Program Plan**

Regency Ballroom (West Wing)

*Speaker: Clifford Lynch (CNI)*

This plenary presentation will look at key developments that the Coalition for Networked Information (CNI) has been tracking over the past year, highlight some specific emerging developments that we believe to be of particular importance, and summarize CNI's 2019-2020 *Program Plan*.  
<https://www.cni.org/program>.

**About the Speaker**

Clifford Lynch is Executive Director at the Coalition for Networked Information (CNI), a joint program of the Association of Research Libraries and EDUCAUSE, based in Washington, DC. More information is at <https://www.cni.org/about-cni/staff/clifford-a-lynch>.

---

**DECEMBER 10 • TUESDAY**

---

2:15pm – 3:30pm

**Closing Plenary: Forgetting and Being Forgotten: Growing Up in a Digital Era**

Regency Ballroom (West Wing)

*Speaker: Kate Eichhorn (The New School)*

A convergence of technological and economic changes has eroded our ability to forget and be forgotten by others. For children and youth, this reversal has also resulted in a series of high-stakes tradeoffs. While young people have finally gained access to the media technologies needed to represent their lives and broadcast these representations on a wide scale, in the process, their ability to take risks without consequence and to carefully curate which memories they carry forward into adulthood has been compromised. This talk explores the decline of forgetting and its impact on social identity development, paying specific attention to the impact this shift will have on young people who have historically had the most to gain from the freedom associated with forgetting and being forgotten.

**About the Speaker**

Kate Eichhorn has spent the past two decades researching the impact of new media technologies on youth cultures, subcultures, and social movements. Her most recent book, *The End of Forgetting: Growing Up with Social Media*, published by Harvard University Press in 2019, examines the impact of social media platforms on identity development, reputation management, and memory. Her previous books include *Adjusted Margin* (MIT, 2016) and *The Archival Turn in Feminism* (Temple University Press, 2013). Eichhorn is Associate Professor and Director of Culture and Media Studies at The New School University in New York City.

---

---

**DECEMBER 9 • MONDAY**

---

2:30pm – 3:30pm

**1.1 Artificial Intelligence: Impacts and Roles for Libraries**

Hampton

*Speakers: Jason Griffey (NISO), Keith Webster (Carnegie Mellon U.)*

Jason Griffey will present on the evolution of artificial intelligence (AI) and potential impacts on libraries, drawing upon his recent book on the topic. Keith Webster will present on the opportunities for libraries to support AI education and research, based on work at Carnegie Mellon, the most prolific AI research institution in the United States. He will also touch on opportunities afforded by AI to advance library priorities.

---

2:30pm – 3:30pm

**1.2 Watchful Eyes on Digital Preservation Action**

Capitol

*Speakers: Laura Alagna (Northwestern U.), Gaele Bequet (ISSN-IC), Peter Burnhill (Independent), Dan Noonan (Ohio State), Jessica W. Meyerson (Educopia), Jane Mandelbaum (Independent), Trevor Owens (LC)*

The session brings together updates on activities that are key to supporting digital preservation in areas relevant to the future of research libraries. It begins with presentations on the IMLS-funded 'Beyond the Repository' and on the transition to the ISSN International Centre of the Keepers Registry for archived digital content issued serially (e-journals and other continuing resources). These will be followed by lightning talks with updates on the Digital Preservation Storage Criteria, developments in the Software Preservation Network (related to virtualization and emulation as methods for enduring access to digital content), the new version of the NDSA Levels of Digital Preservation, and the recently published Library of Congress Digital Content Management Compendium.

Digital Preservation Storage Criteria <https://osf.io/sjc6u>

NDSA Levels of Preservation Working Group <https://ndsa.org/working-groups/levels-of-preservation>

Software Preservation Network <https://www.softwarepreservationnetwork.org>

Digital Collections Management Compendium <https://www.loc.gov/programs/digital-collections-management>

<https://www.imls.gov/grants/awarded/lg-70-18-0168-18>

[portal.issn.org](http://portal.issn.org) [thekeepers.org](http://thekeepers.org)

---

2:30pm – 3:30pm

**1.3.1 Community Development Model for Digital Community Archives**

Calvert

*Speakers: Nathan Hall (Virginia Tech)*

This session will provide an update on the Institute of Museum and Library Services National Leadership Grant project (LG-15-19-0137-19), "Community Development Model for Digital Community Archives," which addresses a gap between community collections and research library infrastructure through collaborative digital repatriation and non-custodial collection development. Small regional institutions and community organizations produce and collect rich cultural heritage collections but lack a robust digital infrastructure preservation strategy. Well-resourced institutions with robust digital infrastructure are able to digitize and provide access to such collections, though too often it occurs through physical acquisition, resulting in smaller institutions losing their unique materials. While physical collection acquisition will always play a part in collection development, it should not be the only means for providing digital access and preservation services. Attendees will learn about the project's methods and strategies for developing and fostering equitable partnerships to expand access to community archives.

<https://di.lib.vt.edu/community-catalysts>

---

### 1.3.2 Challenges and Rewards of Community-Engaged Collection-Building: The Lehigh Valley Engaged Humanities Consortium Digital Archive

*Speakers: Janna Avon, Nora Egloff, Charlotte Nunes (Lafayette)*

Lafayette College Libraries Digital Scholarship Services (DSS) partners on the Lehigh Valley Engaged Humanities Consortium or LVEHC (<https://sites.lafayette.edu/lvehc/>), a four-year grant initiative supported by The Andrew W. Mellon Foundation through June 2021. The goal of the grant is to foster collaboration across campuses, libraries, cultural institutions, and community partners in the Lehigh Valley region of eastern PA, in exploration of themes such as the diversity of communities, the changing nature of work, and the environment and sense of place in the Lehigh Valley during the past half-century. As a partner on the LVEHC, DSS recently launched the LVEHC Digital Archive (<https://lvehc-archive.lafayette.edu>), a regional digital archive representing diverse collection-building efforts across the Lehigh Valley. Contributors include faculty, students, academic librarians, public librarians, and community partners such as members of the local NAACP. The range of primary source materials collected in the archive includes oral histories, family photographs, historic maps, ephemera, and other materials that document ethnic and immigrant communities, emergent economies in the de-industrial era, relationships between culture and the environment, and other topics. Featuring perspectives from three DSS staff members who have played key roles to actualize the archive, this project briefing will address the challenges and rewards of building a regional, community-driven digital archive hosted by an academic library.

<https://lvehc-archive.lafayette.edu>

### 1.4.1 Data Sharing from the Ground Up: Building Data Communities

*Speakers: Danielle Cooper, Rebecca Springer (Ithaka S+R)*

There is a growing consensus that research can progress more quickly, more innovatively, and more rigorously when scholars share data with each other. Policies and supports for data sharing are being put in place by stakeholders such as research funders, publishers, and universities, with overlapping effects. By contrast, many scholars are not engaging in data sharing and remain skeptical of its relevance to their work. As organizations and initiatives designed to promote data sharing multiply -- within, across, and outside academic institutions -- there is a pressing need to understand scholars' perspectives and needs and use that evidence to decide strategically on the best ways to move forward. In a cutting-edge study, Ithaka S+R has drawn on fifteen years of research into scholars' practices to propose a new mechanism for conceptualizing and supporting data sharing. In this session, we will show how successful data sharing happens within "data communities" -- formal or informal groups of scholars who share a certain type of data with each other, regardless of disciplinary boundaries. This framework will enable a participatory discussion of how stakeholders who wish to promote data sharing -- including librarians, information technologists, scholarly communications professionals, and research funders -- can identify and support emergent data communities. We will also consider the implications of this broader approach within the current landscape of repositories, funder and journal mandates, and the public/private data divide. This session will demonstrate how focusing on data communities can lead to strategic support interventions that maximize scholar buy-in and promote greater rigor and innovation through data sharing.

<https://sr.ithaka.org/publications/data-communities/>

### 1.4.2 The Evolution of a Research Data Management and Curation Program: Candid Reflections and Considerations

*Speakers: Moira Downey, Sophia Lafferty-Hess (Duke)*

Research data management and curation programs situated within academic libraries have been growing and evolving over the past decade. Shaped by the increasing focus on open science, reproducibility, data publishing, and research integrity, these programs provide a lens through which to view how academic libraries are engaged in this changing space. While the mechanics of building such programs have been much discussed, continuing to foster open dialogue regarding the institutional contexts, communication strategies, challenges, and lessons learned facilitates cross-institutional learning and community building. How can we implement standards-driven, scalable, and approachable data curation workflows? What strategies have been effective in reaching and communicating with researchers? Who are key partners and how can we meaningfully engage with these partners? How does internal library culture impact our service models? How do we measure success? This presentation will provide candid reflections from Duke University Libraries information professionals regarding the evolution of our research data management and curation program since its formation in early 2017.

*Speakers: Joan K. Lippincott (CNI)*

As part of the planning process for implementing digital scholarship programs, centers, and labs, there are some fundamental issues that should be discussed, analyzed, and addressed. There is no one preferred model for developing a digital scholarship program because institutional factors are so important in designing a program to meet the needs of the local constituency. In this session, we will examine such aspects as program goals, scope, governance, administration, staffing models, funding sources, program elements, physical spaces, and assessment based on examples from CNI's work in this area.

<https://www.cni.org/events/cni-workshops/digital-scholarship-centers-cni-workshop/>

<https://er.educause.edu/articles/2018/3/moving-ahead-with-support-for-digital-humanities>

### **1.6 Maps, Gaps, Landscapes, and Ecosystems: Sorting Out the State of Online Scholarly Publishing**

Congressional A (West Wing)

*Speakers: Katherine Skinner (Educopia), David Lewis (IUPUI), Terry Ehling (MIT)*

This session will focus on the findings from two major, recently issued Andrew W. Mellon Foundation-funded reports on scholarly publishing and knowledge infrastructure: Mind the Gap and the Mapping the Scholarly Commons. We will discuss the feedback these two efforts have received on their reports' audit of over 100 individual active projects and the accompanying sustainability analysis that will inform future adoption, development, and resourcing. During this session, we plan to address questions such as: Where are the primary gaps in the landscape? What communities are not being well served by existing tools? Where is there an overlap between technologies and/or possibilities for interoperability? How do we better understand and address the risk that exists individually and collectively within the 'system'? Where can we set helpful industry standards, and what areas are still best left to open experimentation? What place is there for maintaining information about our shared infrastructure, and how can this information help inform our individual and collective investment decisions?

<https://mindthegap.pubpub.org/>

<https://scholarlycommons.net/>

#### **1.7.1 Experimenting on a Digital Strategy**

Congressional B (West Wing)

*Speakers: Laurie Allen, Meghan Ferriter, Leah Weinryb Grohsgal, Jaime Mears (LC)*

About one year ago, the Library of Congress released its first Digital Strategy. The Library of Congress Labs is part of the Digital Strategy Directorate and it is a space for experimentation and trying new approaches to sharing and using digital collections and data. As the team matures its practice of experimentation and evaluation, it would like to share the different types of pilots, prototypes, and presentations that help create pathways to achieving the goals of the Digital Strategy. Experiments in Labs aim to work with staff to improve services, start new conversations with users and other partner communities, and transform how the Library uses technology to connect to the public. This briefing will showcase staff from across the Library, including curatorial, Labs, and technical staff, and experiments that have led to Library-wide programs (crowdsourcing) and other experiments that we still have a lot to learn from (web archive data releases).

<https://labs.loc.gov>

#### **1.7.2 Accessibility Task Force: Determining Compliance and Organizing Action**

*Speakers: Claire DeMarco, Suzanne Wones (Harvard)*

In June of 2019, Harvard University announced that it would be adopting a campus-wide digital accessibility policy as of December 1, 2019. With a host of systems and digital offerings, the Library needed a comprehensive way to audit our user-facing products and create actionable deliverables for remediation, both by developers and content authors with varying levels of technical expertise. Having had much success working in the agile framework for development projects in the Digital Strategies and Innovation group, we proposed running the review work as a set of Digital Accessibility sprints from July to November. We quickly put together a 12-person, cross-functional team of librarians, UX experts, designers, and developers to review systems, identify issues, and set priorities. The end result will be actionable deliverables to facilitate remediation of library digital products and serve as a framework for ongoing accessibility compliance. This approach reflects our commitment to iterative, innovative work and we hope it will be a model for other institutions adopting similar policies and prioritizing universal accessibility.

[library.harvard.edu](http://library.harvard.edu)

**2.1.1 Ready or Not: Here Comes Voice Search**

*Speakers: Tim Smith, Twila Camp (U. Oklahoma)*

Today's children can be power users of the Internet, YouTube, and other technology before they even know how to read. These same children don't think twice about asking Alexa anything they want to know. Voice searching and virtual assistants are creating a dramatic paradigm shift in information-seeking behavior, and it is barreling towards libraries at lightning speed. This talk discusses the impact of voice search, how libraries can prepare and harness its potential, and the caveats for this artificial intelligence-driven technology. This presentation includes a live demo of an Alexa skill that allows library users to perform a voice search of LibGuides or Primo using vendor APIs.

---

**2.1.2 Can We Talk? Adding a Smart Assistant Interface to Library Services**

*Speakers: Lisa Smith, Greg Davis (Iowa State)*

The Iowa State University (ISU) Parks Library has been working with a third-party developer to create a smart assistant app, currently based on Amazon Alexa technology, to interface with various library services. The smart assistant, called Parks Libro, uses APIs provided by various ISU library systems to return a variety of library-related information, including results of searches of library collections, information related to the user's library account, and information related to library hours and events. This session will provide background information on the work to develop Parks Libro, a demonstration of the current Parks Libro capabilities, and show the roadmap for the next phase of Parks Libro development.

<https://www.lib.iastate.edu/spaces-computers/computers/parks-libro>

---

**2.2.1 Building and Sustaining Community Infrastructure:  
An Update from the Research Organization Registry (ROR)**

*Speakers: John Chodacki (UC CDL)*

**2.2.2 A Progress Report on Open Repositories in Canada**

*Speakers: Geoffrey Harder (CARL)*

**2.2.3 Shared Repository Infrastructure: Two Years Later**

*Speakers: Jimmy Ghaphery (VCU)*

**2.2.4 Shareyourpaper.org: Simplifying Self-Archiving and Cutting the Cost of Mediated  
Repository Deposit**

*Speakers: Leila Bella Sterman (Montana State)*

**2.2.5 Crowdsourcing Inputs and Outputs of a Digital Photo Archive**

*Speakers: Theresa Westbrook (U. Northern Iowa)*

---

### 2.3.1 Developing a Community-Based Strategic Agenda for the Transformation of Archival Discovery and Delivery

*Speakers: Mark Matienzo, Tom Cramer (Stanford), Audra Eagle Yun (UC Irvine), Hillel Arnold (Rockefeller Archive Center)*

Archives and special collections are at a critical point of reconsidering discovery, access, and delivery for their collections given development and adoption of key technical and community infrastructure that supports access, delivery, and use of collections. Systems and platforms supporting these functions, like discovery systems, fulfillment systems, media delivery, and specialized access environments, are becoming increasingly sophisticated and complex, with users and researchers demanding more sophisticated functionality over time. Inadequate systems integration for archival discovery and delivery leads to a frustrating user experience for researchers and archivists, introducing unnecessary friction to ongoing research and fulfillment processes. Systems integration also impacts (and may be driven by) distributed systems support models such as those between departments, in consortial environments, and when using software as a service solutions. Given ongoing concerns about the sustainability of digital infrastructure and equitable access to content managed in it within the cultural heritage sector, there is an opportunity to get an in-depth understanding of how systems integration impacts archival discovery and delivery, and to develop a forward-looking agenda to address these functions holistically.

Through our experience with technical and community development supporting archival discovery and delivery through initiatives like Arclight (an archival discovery environment developed by Stanford University, University of Michigan, Indiana University, Duke University, and Princeton University), Project Electron (developed by Rockefeller Archive Center), and the International Image Interoperability Framework's Archives Community Group, we see an opportunity to align both strategic and tactical work to develop an agenda describing a future for access and use of archives and special collections that cares for communities that use and are represented in them through carefully-considered integrations. This activity is a core focus of Lighting the Way: A National Forum on Archival Discovery and Delivery, a project facilitated by Stanford University Libraries and funded by the Institute for Museum and Library Services, which is convening a series of national meetings on strategic alignment focused on front-end systems integration for archives and special collections.

<https://lightingtheway.stanford.edu/>

<https://projectelectron.rockarch.org/>

<https://library.stanford.edu/projects/arclight>

<https://iif.io/community/groups/archives/>

---

### 2.3.2 Webrecorder, Web Archiving for All: Past, Present and Future

*Speakers: Ilya Kreymer (Rhizome.org)*

Over the past several years, the Webrecorder project demonstrated a user-focused, smaller-scale approach to web archiving, prioritizing user experience and quality over quantity. This presentation will provide a broad overview and briefing on the Webrecorder project's software and tools developed over the last several years and a glimpse of what's next. Topics will include an intro to high-fidelity user-driven web archiving, the intersection of software preservation and web archives with web browser preservation, approaches to automating capture of complex sites, desktop apps for do-it-yourself web archiving, and new research in web archive replay technology that could enable better integration of web archives into existing digital repository and scholarly publishing workflows. The presentation will include lessons learned and focus on approaches for making web archiving more accessible, reliable and less expensive for a wide variety of institutional and individual users, especially those interested in stewarding their own web archive data.

<https://webrecorder.io/>

---

**2.4.1 No One is Using That Anymore: Assessing the Impact of Digital Availability on Print Usage**

*Speakers: Thomas H. Teper (UIUC)*

Librarians speculate that the digitization and delivery of items through the HathiTrust may reduce or eliminate demand for the corresponding print content. This belief feeds into a perception that monographs housed within academic libraries and delivered via such services are ripe for deduplication or outright withdrawal from research libraries, often while other institutions remain dependent upon those institutions to provide access for their patrons. Embracing HathiTrust's emerging Shared Print Monograph Program, many member institutions committed to retain print monographs that correspond to those digitized from their collections. However, such commitments are not universal among the membership. Moreover, such commitments may be utilized by other institutions to withdraw against these holdings without fully understanding the potential local impact. Developing an evidence-based understanding of how the availability of digital access to these items might impact both local circulation and the rate of ILL/DD lending for such items is a critical step in determining how our institutions might approach the management of these collections in the future. The author ingested 10.7-million bibliographic records, 8.3-million circulation transactions, and 751,000 records corresponding to local holdings digitized and available in the HathiTrust into a single MSSQL database and examined the records to determine if the digitization status had any identifiable impact on circulation.

---

**2.4.2 Creating a Virtual Reading Room at UC San Diego Library**

*Speakers: Roger Smith (UCSD)*

The University of California, San Diego (UCSD) Library has established a Virtual Reading Room (VRR) as a function built into its Samvera Digital Asset Management (DAMS) infrastructure. The core UCSD DAMS historically provided three levels of access control: open access, campus IP restriction, and curator only. The Special Collections and Archives Program identified a need to provide access to digital collections that have use restrictions that did not fit into the DAMS access model. A specification was developed that called for the assignment of temporary access credentials to researchers for specific collections or objects. Download is restricted, a click-through agreement articulates use expectations, and files are watermarked. Requests are made via an API integration with the Aeon platform. Development of this function required a collaborative effort between the in-house development group, a regional consultant with expertise in Samvera repository code (Notch 8), and Atlas Systems' Aeon developers. We have already digitized and made available two collections that were highly requested in our physical reference space. We hope that we can share our experience with institutions interested in such an access model for restricted digital content.

**2.5 Human-Centered Artificial Intelligence: The Landscape of Autonomy** Executive (West Wing)

*Speakers: Ben Shneiderman (UMD)*

The next generation of user experiences will produce 1000-fold improvements in human capabilities. These new tools will amplify, augment, enhance, and empower people, just as the Web, email, search, navigation, digital photography, and many other applications have already done. These new human-centered tools will produce comprehensible, predictable, and controllable applications that promote self-efficacy and social participation at scale. The goal is to ensure human control while increasing the level of automation. In short, the next generation of tools will make more people, more creative, more often.

Improved designs will give billions of users comprehensible interfaces that hide the underlying complexity of advanced algorithms. Every use will build confidence that users can reliably accomplish their goals and increase the trust that the machine is under their control.

<http://www.cs.umd.edu/~ben>

---

4:00pm – 5:00pm

### **2.6.1 Center for Research Data and Digital Scholarship at the University of Colorado Boulder**

Congressional A (West Wing)

*Speakers: Thea Lindquist, Thomas Hauser (U. Colorado Boulder)*

The Center for Research Data and Digital Scholarship at the University of Colorado Boulder is a campus research center that supports data-intensive research undertaken by the CU Boulder campus community. The center is a partnership between the University Libraries and the Office of Information Technologies' Research Computing group. The center provides consulting and training in methods and practices supporting digital research (e.g., digital humanities, research data management, programming languages), offers cloud and other cyberinfrastructure to support data-intensive research, and provides interdisciplinary educational opportunities. The center's directors will discuss this partnership in practice as well as some of our infrastructure and educational programming, including data publishing for large data sets and the Digital Humanities Graduate Certificate.

<https://www.colorado.edu/crdds/>

---

### **2.6.2 Connecting Digital Scholarship: Building Communities of Support at University of Michigan**

*Speakers: Joe Bauer, Anne Cong-Huyen (U. Michigan)*

When institutional pressure to show progress and produce cutting edge research often means a desire to solve problems through budgetary or technological means, the Connecting Digital Scholarship group offers an alternate model based on building strong interpersonal connections. This presentation will include discussion of ongoing work in developing a cross-unit program founded upon the value of trust in social connections that brings academic IT into collaboration with library digital scholarship support to build a learning community of scholars and support partners.

4:00pm – 5:00pm

### **2.7 Data Curation Network Update**

Congressional B (West Wing)

*Speakers: Timothy M. McGeary (Duke), Cynthia Hudson Vitale (Penn State), Lisa Johnston (UMN)*

The Data Curation Network (DCN) enables organizations to support researchers in making data FAIR through a shared staffing model for curation expertise across institutions. Since launching in May 2018 through the generous support of the Alfred P. Sloan Foundation, the Data Curation Network has grown to include 10 organizations nationwide. The outcomes of this collaborative model include expertly-curated data regardless of repository destination, standardized data curation practices (CURATED steps), and a unique platform for advocacy and research on data sharing issues. Our work has implications in the broader data curation community through the development of a number of educational initiatives, with funding from the Institute of Museum and Library Services (IMLS), including hands-on workshops and our community-published "data curation primers." Over the past year, the DCN has enlisted the consulting services of LYRASIS and an external advisory panel to assist us in developing a sustainability plan. This presentation will provide an update on these DCN activities and discuss how other libraries and organizations can become involved.

<https://datacurationnetwork.org>

5:15pm – 6:15pm

### **3.1 Memory Institutions and Deep Digital Disruption: Beyond the Technical Challenges of Born-digital Preservation**

Hampton

*Speakers: Clifford Lynch (CNI), Carol Mandel (NYU)*

The presenters will share the work that each has been undertaking to frame and address the profound changes in collection and stewardship posed by born-digital content. Today's memory institutions have evolved over several centuries, shaped by a pre-digital world. Existing societal and institutional structures do not map neatly, or even adequately, into the challenges posed by new forms of digital content, including proprietary news platforms and social media, diffuse personal "archives," scattered cultural heritage initiatives, changes in institutional focus from collecting to service, and assumed reliance on retrospective collecting in a world of dissolving content. While the technical challenges of preserving born-digital content are many and need ever-continuing work, collection and stewardship are essential pre-requisites to having access to knowledge in the future. The presenters will describe their investigations, discuss issues, and facilitate a discussion with the audience about priorities and strategies.

**3.2.1 Connect, Collaborate, and Contribute in the Research Commons: Reflections from Three Years of Service Delivery**

*Speakers: Meris Longmeier (Ohio State)*

**3.2.2 Teaching Data Curation for Reproducibility (Data CuRe)**

*Speakers: Limor Peer (Yale)*

**3.2.3 Piloting Digital Scholarship Support at the Library of Congress**

*Speakers: Eileen Jakeway (LC)*

**3.2.4 Sprinting Toward a Lab: Network Building, Knowledge Sharing and Transforming Communities in Galleries, Archives, Libraries and Museums through a Book Sprint**

*Speakers: Caleb Derven (U. Limerick)*

**3.2.5 InstantILL: Simplifying Content Delivery With or Without Subscriptions**

*Speakers: Tina Baich (IUPUI)*

---

---

*Speakers: William Ingram (Virginia Tech)*

Virginia Polytechnic Institute and State University (Virginia Tech) Libraries, in collaboration with Virginia Tech Department of Computer Science and Old Dominion University Department of Computer Science, is the recipient of an IMLS National Leadership Grant for Libraries award to fund research into bringing computational access to book-length documents, through a research and piloting effort employing electronic theses and dissertations (ETDs). The three-year project is motivated by the following library and community needs:

- (1) Despite huge volumes of book-length documents in digital libraries, there is a lack of models offering effective and efficient computational access to these long documents.
- (2) Nationwide open-access services for ETDs generally function at the metadata level. Much important knowledge and scientific data lie hidden in ETDs, and we need better tools to mine the content and facilitate the identification, discovery, and reuse of these important components.
- (3) A wide range of audiences can potentially benefit from this research, including but not limited to librarians, students, authors, educators, researchers, and other interested readers.

Our research focuses on extracting and analyzing segments of long documents (chapters, reference lists, tables, figures), as well as methods for automated classification and summarization of individual chapters of longer texts to increase findability. The project brings cutting-edge machine/deep learning technologies to advance discovery, use, and potential for reuse of the knowledge hidden in the text of books and book-length documents. By focusing on libraries' ETD collections, the research will enhance ETD programs, devising effective and efficient methods for opening the knowledge currently hidden in the rich body of graduate research and scholarship.

### **3.3.2 Designing a Migration Path: Final Report and Recommendations**

*Speakers: Aaron Birkland (JHU), Jennifer Gilbert (NLM), Tim Shearer (UNC Chapel Hill), David Wilcox (LYRASIS)*

In 2018, the Institute of Museum and Library Services (IMLS) awarded DuraSpace a planning grant (LG- 72 -18-0204) to investigate the barriers that prevent hundreds of U.S. based libraries and archives from upgrading to a supported version of Fedora. Continued community reliance on the unsupported Fedora 3 puts the stability, security, accessibility, and functionality of these repositories at risk. The planning project revealed a pair of crucial themes: effort and value. Any software upgrade/migration requires effort (labor, time, and cost), so it must provide enough value to justify the cost. This insight has informed the design of Fedora 6. This version includes a commitment to developing tools, documentation, case studies, and defined paths that will provide resources to support a valuable upgrade process. This panel session will feature representatives from the project advisory board, Fedora 6 pilot program, and technical team who will present a summary of the report findings, discuss how community stakeholders are engaging with Fedora 6 development and lay out a roadmap for next steps based on the recommendations from the report.

<https://duraspace.org/fedora>

*Speakers: Elisabeth Long (U. Chicago)*

Converting analog data into digital form for use by scientists raises many questions. How good is good enough? Are library practices developed for cultural heritage projects sufficient for scientific inquiry? Can off-the-shelf equipment provide necessary precision? And most importantly, can questions like these provide a learning opportunity for students? The University of Chicago Library has partnered with faculty and students in the Department of Astronomy and Astrophysics to run a pilot study to determine how to scan a large collection of glass plate slides of astronomical images in a way that will facilitate meaningful scientific inquiry. This data rescue project has been an opportunity to question our own practices and to engage undergraduate students in a new kind of learning. In this project, we are exploring the feasibility of using off-the-shelf equipment instead of the specialized, custom-built approach being used by several other significant astronomical data rescue projects. The students are helping to do the necessary mathematical and scientific analysis of our test images as well as pursuing their own research agendas, and we, in turn, are teaching them about good data management. The collaboration has also been an opportunity to learn how scientific data rescue projects may require different assumptions from cultural heritage projects, as well as to explore methods for collaborating with researchers to make hard decisions about what to prioritize in the face of abbreviated timelines and modest budgets. This session will include a discussion of how we have approached this project as a teaching and collaboration opportunity, what we have learned so far in our research, the next steps, and how this can serve as a model for approaching other data rescue projects.

<https://www.lib.uchicago.edu/about/news/mining-historical-glass-slides-for-astronomical-data/>

5:15pm – 6:15pm

### **3.4.2 The Texas GeoData Portal: A New System for Enhancing Access to Geospatial Data**

*Speakers: Michael Shensky (UT Austin)*

Governor's

In 2019 the University of Texas Libraries successfully completed development of its new Texas GeoData portal which has been designed to enhance discoverability of the geospatial data contained in the Libraries extensive collections and facilitate use of these datasets in geographic information system (GIS) software. The technical work for this project was carried out in an Agile development environment by a team of UT Libraries Information Technology developers who worked in coordination with a diverse stakeholder group with a mix of technical, administrative, and collection content expertise. Together this project team was able to overcome a variety of data processing, metadata aggregation, and interface customization challenges by developing a robust system architecture comprised of both open source (GeoBlacklight) and proprietary (ArcGIS Server) software products. These applications were carefully integrated to allow the Libraries to provide access to its datasets in the widest possible variety of formats while still adhering to open standards, utilizing open technologies developed by the university library community, and leaving open the ability to integrate the portal with additional systems in the future. This presentation will discuss how this new portal ties into the broader ecosystem of UT Libraries collections content and research data services, and will offer insight into the lessons learned while developing its underlying infrastructure.

<https://geodata.lib.utexas.edu/>

---

5:15pm – 6:15pm

### **3.5.1 Data Doubles: Student Perceptions of Privacy and Learning Analytics in Higher Education**

Executive (West Wing)

*Speakers: Andrew Asher (IU Bloomington), Michael Perry (Northwestern U.)*

The Data Doubles research project is a student-centered, IMLS-funded project that seeks to understand student perspectives on privacy issues associated with academic library participation in learning analytics (LA) initiatives. In this project briefing, we will be reporting on the results of the first phase of the project which entailed 112 qualitative interviews with undergraduate students across 8 institutions and discussing the creation and validation of a survey to be deployed during the project's second phase in spring 2020. The phase I interviews uncovered that students lacked awareness of educational data mining and analytic practices and the data necessary for these initiatives. Students see potential in learning analytics but also thought about when and where data should be shared and the value of informed consent and expressed nuanced opinions on the acceptability of personal data collection and use. Notions of institutional trust were also discovered, as well as what actions by institutional actors might violate that trust. The phase II survey builds on these findings by developing a quantitative instrument to evaluate these issues in a statistically-representative format and to examine demographic differences that may affect students' opinions. All data collection protocols and instruments from the Data Doubles projects will be published open access via the Open Science Framework.

<http://datadoubles.org/>

---

**3.5.2 Skills, Knowledge, and Values for Prioritizing Privacy in Library Learning Analytics**

Executive (West Wing)

*Speakers: Lisa Hinchliffe (UIUC), Kyle M. L. Jones (IUPUI)*

Learning analytics work extends long-standing library assessment and evaluation practices and they may help further demonstrate library impact on student learning, faculty productivity, and more. Nonetheless, regardless of the benefits, learning analytics work unquestionably presents challenges to student privacy, thus straining the professional ethics commitments that librarians make to uphold user confidentiality, respect privacy in information seeking and use, and support intellectual freedom. When facing these "privacy conundrums," librarians may refrain from engaging with campus learning analytics projects, meaning that librarian values around privacy and confidentiality are missing from those campus conversations and that libraries are marginalized in campus discussions of student success. As such, there is a pressing need to train librarians to handle the particular data ethics issues that arise in learning analytics work-especially the privacy issues-before they begin pursuing learning analytics projects.

Indeed, there is no lack of documentation of the pressing need for training on privacy in learning analytics. Reports from recent IMLS-funded projects (Library Values & Privacy in Our National Digital Strategies, A National Forum on Web Privacy and Web Analytics Prioritizing Privacy, and Library Integration in Institutional Learning Analytics) all express this need. This session shares the in-development curriculum for Prioritizing Privacy, an IMLS-funded professional development program that addresses these needs.

Prioritizing Privacy will teach academic library practitioners about privacy and other related ethical issues associated with learning analytics, provide them structured experiences to reflect on ethical issues intentionally and purposefully, and support the development of privacy protections for their learning analytics projects.

<https://www.imls.gov/grants/awarded/re-18-19-0014-19>

**3.6.1 Into the Dataspace: Data Science Services on the Ground**

Congressional A (West Wing)

*Speakers: Mike Nutt (NCSSU)*

Over the past five years, the North Carolina (NC) State University Libraries have engaged in a rapid experimentation and evaluation cycle aimed at developing and refining the Libraries' services supporting data science. The Dataspace at the James B. Hunt Jr. Library opened in August 2018, and Libraries staff started providing walk-up data science consultation services at the Datapoint at the D. H. Hill Jr. Library in January 2019. User feedback and data from these two efforts generated substantive insights that are now guiding our planning for a new service point in the Hill Library, the Data Experience Lab. As part of a major renovation of the Hill Library, NC State University Libraries will open the Data Experience Lab (DXL) in fall 2020. The DXL will be like a makerspace, but for data science, visualization, and digital scholarship. It will be a place that is as much about people as it is about technology—a hands-on place where community is created and collaborations happen, where powerful computers, specialized software, and peer-to-peer expertise are all in the same space. This 2500 square foot space will be the culmination of a five-year-long, iterative period of needs assessment, learning from others, prototyping spaces, evolving our service portfolio, and optimizing in areas where we have seen demonstrated success. We will share the successes, challenges, and lessons learned along the way, and describe a model for space-based data science support that is centered around a team of peer consultants and professional librarians.

<https://www.lib.ncsu.edu/services/data-visualization>

<https://www.lib.ncsu.edu/spaces/dataspace>

5:15pm – 6:15pm

### 3.6.2 Encouraging the Ecosystem: Growing Unique Partnerships Through Data Science

Congressional A (West Wing)

*Speakers: Joan Peckham, Karim Boughida (URI)*

The evolving national data science ecosystem is encouraging libraries across the country to support the development of data-focused analytic and problem-solving skills. By reaching out to K-12 schools, participating in professional development programs, supporting students, instructors and researchers, university libraries are uniquely positioned to answer the growing demand for data experts and a data-savvy workforce. Building on libraries' success as hosts of makerspaces, artificial intelligence labs, and design thinking spaces, we will examine the benefits of centralized spaces where scholars can partner in data-driven consultations and collaborations. With federal support for the development of new theories and tools, and a focus on a modern, data educated workforce, almost every institution of higher education in Rhode Island is participating in this exciting new multidimensional and interdisciplinary movement. Connecting these efforts to reports from the National Academies and other scholarly and data science educational groups, this briefing will explore the unprecedented partnerships forming nationally, in Rhode Island and at the University of Rhode Island (URI), to meet data-focused challenges in machine learning, artificial intelligence, and big data. We will examine the role of university libraries as data providers and share the joys and challenges of developing educational and research-focused data science programs.

---

5:15pm – 6:15pm

### 3.7 Flipping Open Access Away from APCs

Congressional B (West Wing)

*Speakers: Nick Shockey (SPARC), Amy Buckland (U. Guelph), Ellen Dubinsky (U. Arizona), Nicky Agate (Columbia)*

The APC (article processing charge) path to an open access scholarly publication ecosystem has dominated the scholarly communication discussion since the release of PlanS in Europe in the fall of 2018. Though this path may be feasible for some, it is impossible for others, and the "transformative" nature of an author-pays model seems in doubt when it (a) perpetuates injustices and biases ingrained in the traditional scholarly publishing ecosystem and (b) does little to stem the flow of increasing amounts of money into the system. True and lasting change requires collective investment in new infrastructure models and a commitment to new community-based models of scholarly publishing that open up research communication for equitable participation and equitable access.

Libraries from the University of Arizona, Columbia University, and the University of Guelph have recently redirected funding away from APCs and toward "investments" in open - projects and initiatives that have potential for impactful and substantial change for the many, rather than the few. These institutions are directing funding toward community-led, open infrastructure, advocacy, and other projects which might accelerate the transformation of the existing publication ecosystem to a more equitable and sustainable non-commercial open access model. Panelists will discuss their revised investment funds, describe other community-driven reinvestment opportunities, and detail the truly transformative possibilities of non-APC dependent open access publication models.

University of Guelph: <https://www.lib.uoguelph.ca/get-assistance/publishing-support/open-scholarship-support>

University of Arizona: <https://new.library.arizona.edu/about/awards/oa-fund>

SPARC: <https://sparcopen.org/our-work/landscape-analysis/>

Columbia University: In development

---

9:00am – 10:00am

**4.1 The New Normal: Why Libraries are Teaching AI, ML, DH, NLP, VR...**

Hampton

*Speakers: Harrison Dekker, Indrani Mandal (URI), Matt Burton (U. Pittsburgh), Vicky Steeves (NYU), Tim Dennis (UCLA)*

Search cutting-edge academic libraries' homegrown services and you will find artificial intelligence (AI), machine learning (ML), digital humanities (DH), natural language processing (NLP), and virtual reality (VR) experts, workshops, consultations, and interdisciplinary projects. Why is that? While on the surface this segment of information and research services may be considered beyond the accustomed library role, the growing popularity of these offerings, at a time when traditional library services are on the wane, suggests that they are fast becoming the new normal. By connecting and fostering collaboration across the disciplines, these libraries are enhancing the teaching, learning and research experience and helping reassert the library's key role as the center of scholarship. In this panel, data-centric librarians and scientists explore the origins of this trend and possible future directions.

---

9:00am – 10:00am

**4.2.1 Guerrilla Governance and Program Review: Organizational Transformation in Collaborative Models**

Capitol

*Speakers: Joanne Kossuth (DLF), Bradley Daigle (UVA, APTTrust)*

Across the library and archival profession, collaborative programs form a critical role in sharing information, developing new skills, and launchpads for new strategic efforts. This talk will focus on recent activity within the CLIR (Council on Library and Information Resources) program umbrella - namely the Digital Library Federation (DLF) and the preservation-focused National Digital Stewardship Alliance (NDSA). Both organizations have undergone a significant review of mission, goals, and achievable outcomes in the past year. The Digital Library Federation has provided a collaborative platform for digital library practitioners since 1995. As it approaches its 25th anniversary, CLIR has embarked upon a thorough review of the program to ensure it remains relevant, vibrant, and welcoming into the next quarter-century and beyond. Consultant Joanne Kossuth, Principal at 1MountainRoad and a longtime dean of CLIR's Leading Change Institute, is leading this review on behalf of CLIR. Incorporated into the review process are conversations with leaders of DLF's working groups and other community members; an assessment of the annual Forum; an analysis of where and how CLIR can best provide support to individual practitioners and related organizations to help their work advance; and how, through this strong and trusted program, CLIR can continue to work with other committed individuals to support the creation, preservation of, and access to our global cultural heritage and knowledge as a public good. At this session, Kossuth will provide an overview of her findings to date. Community input and feedback will be welcomed and encouraged. The NDSA transitioned from the Library of Congress to DLF in 2016 but despite an entirely different organizational structure, there was little to no change in how the program was run. With a growing volunteer membership of over 250 organizations, it became clear that the NDSA needed to take a closer look at both the mechanics and function of its network. Beginning in January 2019, the NDSA leadership worked to revise, clarify, and extend its governance structure. A work in progress, this talk will share our experience as we look to global collaboration.

<http://www.diglib.org> <https://ndsa.org/>

<https://ndsa.org/about/>

Governance documents: <https://docs.google.com/document/d/1v5IToyV6YECeNsBX6b7ucT9bmlx6-HyGTZwH0EgP4E/edit>

---

#### 4.2.2 Responsible Operations: Shaping a Community Research Agenda

*Speakers: Sarah Shreeves (U. Arizona), Thomas Padilla (OCLC Research)*

In light of widespread interest, OCLC Research commissioned the development of a community research agenda that charts library engagement with data science, machine learning, and artificial intelligence. "Responsible Operations: Shaping a Community Research Agenda" is the result. This work was guided by an advisory board and constituted by interviews, an in-person working event, and asynchronous contributions from March 2019 to September 2019. Suggested areas of investigation are interdependent in nature, spanning responsible operations, description and discovery, shared methods and data, machine-actionable collections, workforce development, data science services, and interprofessional and interdisciplinary collaboration. Given the scale of positive and negative impacts presented by the work that lies ahead, committing to interdependence across challenges is more important than ever. This session aims to report out on Responsible Operations and start a conversation focused on collective action.

<https://hangingtogether.org/?p=7024>

<https://hangingtogether.org/?p=7320>

#### 4.3.1 Evolving with Agile into a New Technical Landscape: Cultivating a Sustainable DAMS Ecosystem

*Speakers: Dave Ronn, Megan Will (UT Austin)*

In the past few years, the University of Texas at Austin Libraries have pushed to modernize their digital tools and technology to better serve students, faculty, and researchers. This presentation describes the implementation of the UT Austin Libraries Digital Asset Management System (DAMS) Ecosystem. The goals were to create a new platform to ingest, manage and preserve digital assets, provide tools to allow researchers and the general public to view special digital collections online, and simultaneously modernize the underlying technology stack. We will discuss how we transitioned to enterprise workflow automation for asset ingest, containerized application hosting with automated tests, asynchronous messaging across systems, continuous integration for deployment, and challenges we faced in the process.

<https://collections.lib.utexas.edu/>

#### 4.3.2 Developing an Open Source Digital Scholarly Research System: Local and Global Possibilities

*Speakers: Ray Uzwyshyn (Texas State)*

During the last five years, Texas State University Libraries has developed a successful open-source digital scholarly ecosystem enabling university faculty and graduate students with their scholarly research online. This presentation overviews the elements that are necessary to develop a digital scholarship research ecosystem that is suitable for any university, college, research institution or academic research library interested in setting up such an infrastructure. Software systems, hardware, human resources, and timelines will be outlined with a pragmatic focus on open source software, best-in-class applications, and global best practices. Major digital scholarly system components in a larger digital ecosystem will be discussed: online institutional collection repositories (D-Space), online research data repositories (Dataverse), identity management systems (ORCID), electronic thesis and dissertation management systems (VIREO), academic journal systems (OJS3), digitization labs, user interface software (OMEKA) and synergistic possibilities. Because of the open-source affordances, the system is also suitable for any university, college, research institution or academic research library on more international levels interested in setting up such an infrastructure and enabling university faculty and graduate students with their scholarly research online with larger potential and prospects. These synergistic and global possibilities will also be discussed.

Texas State Digital Collections Repository: <https://digital.library.txstate.edu/>

Texas State Data Research Repository: <https://dataverse.tdl.org/dataverse/txstate>

Texas State Online Research Identity Management System:

<https://guides.library.txstate.edu/researcherprofile/orcid>

Texas State Electronic Thesis and Dissertation Management (VIREO): <https://www.tdl.org/etds/>

Texas State Digital & Web Services: <https://www.library.txstate.edu/about/departments/dws.html>

*Speakers: Patty Hinegardner, Na Lin (U. Maryland, Baltimore)*

Research data sharing, required by many funding entities including US government agencies and foundations, is shaping the way data is stored and discovered. Datasets are now maintained in repositories scattered around the web or residing on local area networks, hard drives, and even flash drives. To facilitate sharing of University of Maryland, Baltimore (UMB) research data, the UMB Health Sciences and Human Services Library created an institutional data catalog using open source technology developed at New York University. The UMB Data Catalog's mission is to identify UMB datasets, wherever they reside and provide in-depth curation for optimal findability and access. This presentation will provide an overview of the project including required resources, strategies for outreach, challenges of dataset description, development of metadata standards, and lessons learned throughout the project. It will also discuss the benefits of being a member of the Data Catalog Collaboration Project (DCCP). The DCCP was created to facilitate the discovery of biomedical research data. It consists of academic health science libraries that have or are considering implementing a local instance of NYU's open-source data catalog. This collaboration brings a cross-institutional perspective to addressing usability, data sharing workflows, metadata, and outreach for improving data discovery efforts in biomedical research. In 2018, the DCCP won the Clinical & Translational Science Award (CTSA) Great Team Science Award Contest for the Top Importance Category in a field of 170 entries.

UMB Data Catalog - <https://datacatalog.hshsl.umaryland.edu/>

DCCP - <https://www.datacatalogcollaborationproject.org/>

---

#### **4.4.2 Mobilizing Computable Biomedical Knowledge, and Making it FAIR**

*Speakers: Marisa Conte, Peter Boisvert (U. Michigan)*

This project briefing introduces emerging efforts to accelerate research and improve health by sharing computer algorithms on a large scale. These efforts include the work of the Mobilizing Computable Biomedical Knowledge (MCBK) community, and a knowledge infrastructure project to develop both standards for handling computable biomedical knowledge objects, and the methods needed to organize and disseminate them. MCBK is a 2-year-old community of researchers, developers, clinicians, librarians, informaticists and others, all dedicated to enabling the curation, dissemination, and application of computable (machine-executable) biomedical knowledge (CBK) in clinical, research, and public health contexts. Adopted in 2018, the MCBK Manifesto presents the group's vision: to mobilize CBK using machine-executable formats that can be integrated into health systems and applications, to ensure the trustworthiness of CBK as a reflection of the best and most current evidence and science, and to align CBK development and dissemination with the FAIR principles, such that it is Findable, Accessible, Interoperable and Reusable. The University of Michigan Medical School's Department of Learning Health Sciences is currently developing an open-source Knowledge Grid (KGrid) platform to facilitate the dissemination and deployment of computable knowledge objects (KOs). The KGrid team has recently expanded to include librarians with expertise in health sciences and digital preservation. This unique partnership leverages librarians' expertise to standardize and streamline the methods by which knowledge moves into practice by applying the FAIR principles to computable knowledge. Our presentation will conclude with a brief overview of the KGrid platform to date, including a demonstration. We seek CNI community input and support both on the facilitation of CBK validation, and on the deployment of validated KOs into existing platforms to enhance findability and access. Finally, we hope to encourage CNI members to engage more broadly and lend their expertise to MCBK initiatives.

Authors (listed alphabetically): Peter Boisvert, Marisa Conte, Allen Flynn, Lance Stuchell

Mobilizing Computable Biomedical Knowledge (MCBK) community: <https://mobilizecbk.med.umich.edu/>  
MCBK Manifesto:

<https://medicine.umich.edu/sites/default/files/content/downloads/MCBK%20Manifesto%20Ver%2010.7.18.pdf>

Knowledge Grid: <https://kgrid.org/>

---

**4.5.1 Electronic Lab Notebooks: Implementation, Evaluation, and Lessons Learned**

*Speakers: Harish Maringanti, Daureen Nesdill (U. Utah), Xuemao Wang (U. Cincinnati)*

Executive (West Wing)

Two institutions will share their experiences with electronic lab notebooks (ELN) as a follow-on to the fall 2019 CNI Executive Roundtable on that topic.

In 2017, the University of Utah initiated and implemented a campus-wide ELN solution. Partnering with the Office of the Vice President for Research and the University's central IT, the libraries took the leadership role in investigating ELNs as a component of its data management services. After three years of introducing the product to campus research groups, some unexpected use cases emerged and collaborations initiated. This project briefing will include an outline of the project rollout in the context of data management services offered by the Library, and a discussion of opportunities and challenges that arose with the implementation of ELNs.

From November 2016 to June 2018, the University of Cincinnati (UC) Libraries partnering with UC's Academic Medical Center, Office of Provost, and Office of Vice President for Research introduced a vendor based ELN pilot solution LabArchives. The LabArchives is a cloud-based ELN application that is compliant with several federal mandates and can be integrated with numerous software tools. This project briefing will describe the pilot program and discuss some of its findings, and the institution's learning experience.

**4.5.2 Organizational Strategies to Support Emerging Needs in Areas Such As: Data Science, Technology Rich Spaces, and Experiential Learning**

*Speakers: David Woodbury, Mira Waller (NCSU)*

The changing needs and expectations of faculty, researchers, and students around research, teaching, and learning provide great opportunities for libraries to add value to their academic institutions by providing broader and deeper services around emerging needs and technologies. At North Carolina State University Libraries we have been incubating, launching, and sustaining services to address emerging areas of emphasis such as data science, visualization spaces, and experiential learning. While working in these evolving and emerging areas we have faced challenges and learned strategies to use creative ways to grow and maximize staff skills, work together across library departments, create new staff roles, and better utilize student positions. In this session we will share the strategies we developed, the lessons learned, and how they might be adapted for future campus needs.

**4.6.1 Experimenting with a Machine Generated Annotations Pipeline**

*Speakers: Joshua Gomez (UCLA)*

Congressional A (West Wing)

The University of California, Los Angeles Library recently created a small subteam dedicated to conducting strategic software development experiments. This presentation is a review of our first experiment, MGAP, in which we built a prototype event-driven digital library whose images were harvested via IIIF protocols, annotated via the Web Annotations protocol using tags from computer vision services like Amazon Web Services Rekognition, Google Vision, and Clarafai. These machine-generated annotations were automatically merged into the search index of the prototype interface. We then conducted moderated and unmoderated user testing to evaluate whether the tags from these services actually improved the user experience of the digital library.

**4.6.2 A Demonstration of Annotation Interoperability**

*Speakers: Mark Patton, Sayeed Choudhury (JHU)*

Through a grant from the Andrew W. Mellon Foundation, Johns Hopkins and Tufts University are working with a group of partners to develop a framework for annotation interoperability. The overarching goal of this grant is a demonstration of making use of annotations created in one application for one purpose in another application for another purpose using web-based standards, protocols, and associated APIs. Through an initial planning workshop, we developed relevant ideas and approaches in consultation with application teams from Hypothes.is, T-Pen, Digital Mappa, Pelagios Commons, and with Rob Sanderson, Semantic Architect at the Getty Institute and member of the W3C Web Annotation working group. The Johns Hopkins team has modeled sample content from the Digital Library of Medieval Manuscripts and the Tufts team has modeled sample content from the Perseus Digital Library in a manner that both teams can display content within separate reading environments. This presentation will describe the proposed approach and demonstrate our initial results.

**4.7.1 The End of a Statewide Digital Preservation System: Coping with the Fallout***Speakers: Fletcher Durant, Todd Digby (U. Florida)*

Congressional B (West Wing)

This presentation will discuss how the University of Florida responded when the statewide digital preservation system was unexpectedly decommissioned, leaving a limited amount of time to plan and transition to a new workflow and system. We will examine the challenges and possible benefits that resulted in managing the response, which included building new partnerships, managing the complexities of moving 280 TBs of data, and examining new technology workflows.

<https://ufdc.ufl.edu>

**4.7.2 Project Canopus: Rethinking Preservation Infrastructure in Toronto, Ontario, and Beyond***Speakers: Steve Marks (U. Toronto)*

The University of Toronto Libraries and Scholars Portal embarked on a joint project two years ago to re-architect the way we think about and do digital preservation. In light of the ever-increasing amount of data our libraries are called to preserve, we realized that the approaches we had used previously were no longer keeping up. As a result, we have moved away from a monolithic repository model and have moved toward a more flexible, microservice-based approach. This approach has now been realized in a new DAMS we call Canopus. This briefing will include information about this new system and where this journey has taken us, as well as some lessons learned in the process, by way of contributing to the ongoing dialog about what it means to collect and preserve digital materials in 2019 (and hopefully beyond).

**5.1.1 Building a Digital Preservation Strategy Across a Broad University System**

Hampton

*Speakers: Edson Smith, Mary Elings, Todd Grappone (UCLA)*

In the fall of 2018, the University of California (UC) libraries, the largest university research library in the world, formed a multi-campus working group charged both with developing a practical, shared vision of digital preservation for library content, and with outlining a roadmap to guide the UC libraries in advancing that shared vision. The initial incarnation of the group was charged with conducting a survey of notable exemplars, detailing current digital preservation best practices in the field, and investigating the state of digital preservation at each individual UC campus. This talk will detail the group's activities, methodology, and findings, and describe how they will inform an overall digital preservation strategy within the UC system.

**5.1.2 Toward Collaborative Models for Sustaining Digital Scholarship***Speakers: Katrina Fenlon (UMD)*

Digital collections and projects created by humanities scholars constitute extensive, valuable, scattered bodies of historical and cultural evidence. Born outside of memory institutions, these collections and projects confront major barriers to sustainability. Though many at-risk digital collections and projects are within the scope of preservation missions of institutional libraries and archives, most institutions lack capacity to take in and sustain any more than a narrow swath of digital scholarship. Beyond that, research communities have deep investments in their collections; many seek new models of institutional partnership that keep varying levels of power over and responsibility for collections in the community's hands. Libraries and archives cannot and should not comprehensively collect and sustain the growing mass of digital community collections. The "Sustaining Digital Community Collections" project aims to develop context-driven sustainability models, which share responsibility for the long-term care of digital projects among libraries and research communities. This project builds on a prior study of digital humanities collections, which found that sustaining digital collections depends on understanding and maintaining idiosyncratic, distributed, collaborative workflows of collection development and maintenance. In order to build sustainable infrastructures for the cultural record in a growing diversity of communities and institutions, we need an understanding of the human and technical workflows on which collections depend. This project is currently developing cases of three digital humanities projects-including a large-scale linked data hub, a unique corpus of Islamicate texts, and a local community archive--to document collaborative workflows of collection development and maintenance and identify roles that libraries and archives may play to help realize community-determined, community-led strategies for sustaining digital collections.

**5.2.1 Fedora 6 and the Oxford Common File Layout**

*Speakers: Andrew Woods (DuraSpace)*

**5.2.2 IIF: What's New and What's Next with A/V and Discovery**

*Speakers: Josh Hadro (IIF Consortium)*

**5.2.3 It Took a Village: The Evolution of Samvera**

*Speakers: Carolyn Caizzi (Northwestern U.)*

**5.2.4 Update on Public Access Submission System (PASS)**

*Speakers: Hanh Vu (JHU)*

**5.2.5 Acknowledging Core Facilities and Collections Use with ORCID**

*Speakers: Eric Olson (ORCID)*

---

---

10:30am - 11:30am

### **5.3.1 InvenioRDM: A Collaborative Next-Generation Research Data Management and Repository Solution**

Calvert

*Speakers: Thomas Morrell (Caltech), Kristi Holmes (Northwestern U.)*

Next-generation research data management and repository solutions are critical in today's research ecosystem to empower discovery and access by the community, support FAIR (Findable, Accessible, Interoperable, & Reusable) science, and collect, record, preserve, and disseminate a wide range of research products -- critical to enhance visibility and promote people and their expertise. Good RDM infrastructure requires a trusted framework for digital objects; good data practice workflows including support for automation and large files; incorporation of standards and persistent identifiers; incorporation of privacy considerations; and development of strategies to support implementation and incentivize individuals participating in such an ecosystem. InvenioRDM is a turn-key, integrated, born-interoperable RDM platform and repository to empower research data management, licensing, preservation, credit, discovery, and reuse of digital artifacts and data. The platform is being developed as part of a large, multi-organization collaboration with the European Organization for Nuclear Research (CERN), birthplace of the World Wide Web and developers of the Zenodo open data repository service. Invenio RDM comes with pre-configured profiles for institutional repositories (IRs) and research data management (RDM) systems, as well as domain-specific repositories for health and medical sciences. InvenioRDM development is driven by user needs and informed by best practices and standards, including those that help define Next Generation Repositories as a foundation for a distributed, globally-networked infrastructure for scholarly communication, discovery, and innovation. To help support sustainability of the platform, various levels of customization are available in InvenioRDM with platform support providers available to provide hosted repository services. We are cultivating an inclusive open source community with thoughtful governance to ensure a collaborative strategic direction. This highly interactive session will provide an update to the community about this work, describe our development roadmap, present key interdisciplinary use cases for this work, share local-level support resources, and gather real-time audience input.

This work is supported by the CERN Knowledge Transfer Fund and over 20 participating partners including Caltech Library, Northwestern University Feinberg School of Medicine, and the US Center for Data to Health (CD2H), supported by the NIH National Center for Advancing Translational Sciences grant number U24TR002306, to help guide the cultural and technological changes necessary for data and informatics to improve research and health care.

<https://inveniosoftware.org/>

<https://inveniosoftware-contrib.github.io/inveniosoftware.org/> (soon to be available at the

<https://inveniosoftware.org/> site)

---

### 5.3.2 Opportunities for Academic Libraries to Shape National Approaches to Research Data Management: A Canadian Perspective

Calvert

*Speakers: Susan Haigh (CARL), Lee Wilson (CARL Portage), Jason Brodeur (McMaster U.)*

The Canadian Research Data Management (RDM) landscape is evolving rapidly as a result of two national initiatives: a) the development and imminent release of an RDM policy applicable to all research supported by national funding agencies, and b) efforts by the federal government to reorganize the coordination and governance of Canada's broader Digital Research Infrastructure. With regard to these initiatives, Canadian academic libraries have played an important role in developing expectations and requirements for a national RDM approach through the work of the Canadian Association of Research Libraries' (CARL) Portage Network. Since its launch in 2015, Portage's operations have focused on expanding networks of expertise for RDM in Canada, while also developing, connecting, and sustaining the infrastructure and services required to support RDM needs at a national scale. In this project briefing, we discuss how the work of the Portage Network has shaped (and is shaping) the emerging Canadian RDM landscape at numerous levels of engagement. Following an overview of the current Canadian RDM context and Portage's role in its ongoing development, we will provide an update on recent expansions to national-scale systems and services--this includes the transition of national data repositories to full production and the employment of a suite of national-level coordinator positions to strengthen the Network's national data stewardship activities. Finally, we will summarize the outcomes of the Canadian Data Curation Forum--a Portage-led event that brought together relevant stakeholders and practitioners to articulate a vision for a national approach to data curation in Canada.

<https://portagenetwork.ca/>

<https://data-curation.github.io>

### 5.4.1 Visualizing Use and Performance Data from a Global Cross-platform Set of Institutional Repositories

Governor's

*Speakers: Jonathan Wheeler (UNM), Minh Pham (U. Missouri Columbia), Nikolaus Nova Parulian (UIUC), Kenning Arlitsch (Montana State)*

The Repository Analytics & Metrics Portal (RAMP) is a web service that has generated a large and unique dataset measuring aggregate use and performance of institutional repositories (IR). RAMP collects data about IR items from registered repositories that have appeared in the search results of all Google properties, including the position of each item in the results, the frequency of its appearance, and whether users clicked through to the IR that hosts the item. This briefing will show data aggregated from 35 institutional repositories around the world during the first five months of 2019. Utilizing the data analysis and visualization skills of fellows from Drexel University's LIS Education and Data Science (LEADS) program, we will demonstrate that higher use ratios are not always the domain of larger repositories. We will also show patterns of use based on repository content and platform, demonstrate the effect of position in search engine results, and show how device use varies by region and affects information-seeking behavior. Finally, we will discuss future research potential and provide attendees with a link to the publicly available dataset, documentation, and sample code for analyzing the data. Attendees will be invited to register their IR with RAMP to contribute to this growing dataset.

<https://osf.io/68xpr/>

### 5.4.2 North Broad Press: A Collaborative Library/Press Publishing Program

*Speakers: Mary Rose Muccie, Annie Johnson (Temple U.)*

The increase in the number of university presses reporting to their university libraries has brought opportunities for greater collaboration, communication, and alignment between two important pieces of the scholarly communications chain. This collaboration can be awkward. The perspectives of presses and libraries on key aspects of publishing can vary, and true partnerships are hard to identify. In early 2019, Temple University Libraries and Temple University Press launched a new joint imprint, North Broad Press. Building on the Libraries' longstanding commitment to and interest in textbook affordability, as well as the Press's expertise in project development, peer review, and production, North Broad Press publishes peer-reviewed open textbooks and other open educational resources written by Temple University faculty. Although funding primarily comes from the Libraries, staff from both the Press and the Libraries are involved in the editorial and production process. Speakers will present North Broad Press as a model for other libraries and presses looking for meaningful ways to collaborate. They will discuss how it was founded, how projects are acquired and produced, and how its staffing model works.

<https://temple.manifoldapp.org/projects/project-collection/north-broad-press>

<http://tupress.temple.edu/open-access/north-broad-press>

<https://library.temple.edu/services/7>

10:30am - 11:30am

## 5.5 Refreshing the Agenda for Collaboration: Library, IT, and New Partners

*Speakers: Clifford Lynch, Joan K. Lippincott (CNI)*

Executive (West Wing)

When the Coalition for Networked Information (CNI) was founded in 1990 by the Association of Research Libraries and the two organizations that became EDUCAUSE, it was intended to serve as a vehicle to allow librarians and information technologists to collaborate and to respond to the opportunities offered by emerging technologies, at the time most notably the Internet. The agenda in 1990 was fairly clear and the opportunities overwhelming.

30 years later, the world has changed: the organizational landscape, priorities, and missions of our universities have evolved, a generational shift in leadership has occurred, and the ubiquitous deployment of advanced technologies at scale creates new and different challenges and opportunities. Over the past year, CNI has convened a series of roundtables involving library and IT leaders -- including many new leaders -- to explore the changing character and purposes of library-IT collaborations and identify high priority and high payoff opportunities for the 2020s. The roundtables focused on: the broad landscape; teaching, learning, and student success; supporting the research enterprise; and institutional policy challenges in data governance and privacy. This session will report on the findings from these roundtables and we will encourage discussion of future CNI initiatives.

<https://www.cni.org/go/it-and-library-leaders-agenda-priorities-for-collaboration-cni-report>

---

10:30am - 11:30am

## 5.6.1 Collaborative Empowerment, Empowering Collaboration: The Carpentries at the University of Toronto

Congressional A (West Wing)

*Speakers: May Chan (U. Toronto)*

This session describes a pilot project that seeks to facilitate inclusive and equitable opportunities to learn core computational literacy and data science skills at the University of Toronto. The project team's approach to promoting such skill development actively reaches a diverse range of students, faculty, and staff in the university community regardless of the field of study or work. The heart of this pilot is driven by becoming a member organization of The Carpentries, a global community whose mission is to teach foundational computational and data science skills to researchers in academia, industry, and government through a transformative pedagogy and cooperatively developed curriculum. The session outlines how the project team has capitalized on membership benefits to cultivate a learning culture that relies on cross-functional collaboration at a large research university. Such collaboration empowers novice learners to acquire computational and data science skills in a mindfully scaffolded way, continue in self-directed learning, and give confidence to other novice learners in their networks. This presentation will be relevant to those interested in innovations in teaching and learning and infrastructure to support research.

<https://oneseach.library.utoronto.ca/carpentries-university-toronto>

---

## 5.6.2 Python Camp: Meeting the Demand for Computational Skills Through Open Technology and Reusable Curriculum

*Speakers: Lorena A. Barba, Hannah Sommers, Megan Potterbusch, Laura Wrubel (GWU)*

Learning Python, especially for data analysis, is a priority for students and faculty across disciplines, but good support to do so outside of a computer science course isn't typically available. George Washington University (GW) Libraries and Academic Innovation has been experimenting with a new model of Python programming instruction which leverages a GW engineering faculty member's openly licensed, Open edX curriculum, funded by the National Science Foundation Office of Advanced Cyberinfrastructure CyberTraining program. Designed for teaching computational thinking foundations to first-year engineering students, we adapted the curriculum into a three-day "Python Camp" to address the needs of a broad range of campus learners. The librarian instructors, also trained in the Carpentries pedagogical approach, emphasize live-coding and formative assessment. Key technical components of the mini-course include Open edX's platform's integration with Jupyter notebooks, auto-graded homework assignments, and the library-hosted JupyterHub platform. Learners who complete course requirements receive a certificate of completion. Python Camp sells out like a rock concert! While aligning with both the library mission and campus strategic priorities around increasing programming skills for research across disciplines, this project demonstrates the power of reusable computable content and open educational resources, as well as collaboration between library and faculty partners. The presenters will also discuss the challenges Python Camp raises for sustainability and scalability, as a hugely popular yet also resource-intensive offering.

---

10:30am – 11:30am

## **5.7 The Future of OA: The Impact of Open Access on Readership and Subscription Decisions**

*Speakers: Jason Priem, Heather Piwowar (Our Research)*

Congressional B (West Wing)

A recent\* large-scale study by the non-profit company Our Research (makers of Unpaywall) estimates that by 2025 44% of all journal articles will be available as open access (OA), and 70% of all article views will be to OA articles. We will update the meeting on the methodology and findings of this study, as well as the open data behind it. We'll close with a description of our new project based on this work, intended to make OA projections actionable for librarians making subscription decisions.

\*Submitted to bioRxiv on October 6, 2019

[https://docs.google.com/document/d/1NtGVKTQU7i6WGDIEtOGdrYgAxyhc7I51dBMUnn\\_CDpg/edit#](https://docs.google.com/document/d/1NtGVKTQU7i6WGDIEtOGdrYgAxyhc7I51dBMUnn_CDpg/edit#)

---

1:00pm – 2:00pm

## **6.1 Machine Learning in Research Libraries: A Snapshot of Projects, Opportunities and Challenges**

Hampton

*Speakers: Zheng (John) Wang (Notre Dame), Elizabeth Lorang (U. Nebraska-Lincoln), Harish Maringanti (U. Utah)*

Adoption of machine learning (ML) in cultural heritage organizations and research libraries is broadening, and will only continue to grow, be it for gaining internal efficiencies via automating collections processing or enhancing access and discovery experiences for users. Over the last several years, each of our teams has been exploring a range of practical and theoretical questions related to the intersection of machine learning and libraries. In this presentation, project leaders will discuss their efforts and outline opportunities and complexities in using machine learning in research libraries.

---

1:00pm – 2:00pm  
Short Updates IV

### **6.2.1 Modern Endangered Archives Program at the UCLA Library**

Capitol

*Speakers: Rachel Deblinger (UCLA)*

### **6.2.2 The Foundations of Discovery: A Short Summary of the Assessment of the Impacts of CLIR's Cataloging Hidden Collections Program, 2008-2019**

*Speakers: Joy M. Banks (CLIR)*

### **6.2.3 Reaching the Researchers: Using Geographic and Chronological Metadata to Facilitate Access to New Acquisitions**

*Speakers: Christian Casey (NYU)*

### **6.2.4 Preservation of Electronic Government Information (PEGI) Project Next Steps**

*Speakers: Roberta Sittel (U. North Texas)*

### **6.2.5 Lever Press Update**

*Speakers: Beth Bouloukos (Amherst College)*

---

### 6.3.1 Pay to Play: Licensing Local Television News Content

*Speakers: Morgan Gieringer (U. North Texas)*

Local television news played a critical role in recording the social and cultural history of America for much of the 20th century, however, archival news film and video has not been widely preserved due to a myriad of factors including the cost to digitize analog content and the difficulty in managing large and poorly organized news archive collections. The University of North Texas (UNT) collaborated with the oldest television news station in Texas to preserve over 60 years of film, video, broadcast scripts and other related materials. Through this unique custodial partnership model, all digitized materials are freely available to view or stream on-line. Additionally, UNT also serves as the licensing agent for the news archive. Working with producers for a variety of outlets including PBS, ESPN, HBO and Netflix, the news archive generated over \$100,000 in revenue in the past year, with revenues expected to increase as more content is digitized. This briefing will provide an overview of the partnership between the university library and the news station, and will describe the growing marketplace for archival media content fueled by demand for original programming from streaming services like Netflix and Hulu.

<https://library.unt.edu/special-collections/collections/kxasbc-5/>

<https://texashistory.unt.edu/explore/collections/KXAS/>

<https://exhibits.library.unt.edu/nbc5>

---

### 6.3.2 Unlocking Opportunity: Using the JSTOR Platform to Get Library Special Collections into the Research Workflow Without a Paywall

*Speakers: Bruce Heterick (ITHAKA)*

Libraries, museums, and archives hold rich and unique special collections that are critical to research, teaching, and learning. Yet these collections are often time-consuming and expensive to support and can be difficult for scholars to discover and use. JSTOR has a history of unlocking the value of underleveraged materials, starting first with digitizing the back runs of journals. Today, the JSTOR platform is visited by millions of researchers every day, in thousands of institutions from 175 countries around the globe. We are working to take the infrastructure we have developed and make it openly available to libraries. Libraries will be able to host their special collections on JSTOR, integrated alongside books and journals, where they can reach scholars from around the world. We will discuss how we are working with libraries to help make JSTOR their platform and enable special collections to have their broadest possible reach.

[www.jstor.org](http://www.jstor.org)

---

### 6.4.1 Release of Digital Preservation Risk Analysis Documents by the National Archives (NARA)

*Speakers: Leslie Johnston (NARA)*

In early September 2019, NARA released its complete digital preservation framework for public comment on GitHub: a File Format Risk and Prioritization Matrix used to calculate risk levels for over 350 file format variants so far, as well as 15 draft Preservation Action Plans, each addressing a different record type (digital design, email, still images, word processing files, web archives, etc.). The plans propose processing and preservation actions for the formats and the potential tools to be used; in some cases, the plans identify where further research is needed into the best preservation approach to take, and where there are no easy decisions about which tools to use. NARA released these not just to provide transparency into its digital preservation risk assessment process, but to elicit comments, discussions, and suggestions for approaches and tools from our colleagues and researchers. This presentation will provide an overview of the digital preservation framework and the preservation action plans, discuss the reception and the feedback received during the 2-month long public comment period, and how the formal release of the plans in 2020 will incorporate the suggestions and inform practice in the digital preservation community.

<https://www.archives.gov/preservation/electronic-records>

<https://github.com/usnationalarchives/digital-preservation>

---

*Speakers: Andrew Wilson (NARA)*

The National Archives and Records Administration (NARA) has a strategic goal to Make Access Happen. In pursuit of that goal, NARA has digitized and provided online over 90 million pages of records and is on track to provide access to 500 million digital copies by the end of FY2024. As the haystack grows, the needles become more difficult to find. NARA is exploring numerous ways to improve discovery of the records, including the development of "next-generation finding aids." These finding aids provide digital paths into the records by using NARA's API and leveraging the cataloging work performed by NARA staff over the last 20 years. This discussion will review our recently launched project, the Record Group Explorer, discuss lessons learned, and provide an overview of additional access projects underway for FY 20. There will be time included for audience discussion.

<https://www.archives.gov/findingaid/explorer>

*Speakers: Josh Sternfeld (NEH), Lucy Barber (NHPRC), Becca Quon (CLIR), Ashley Sands (IMLS)*

Representatives of major federal funding agencies--National Endowment for the Humanities, Institute of Museum and Library Services, National Historical Publications & Records Commission, and Council on Library and Information Resources--will discuss the current status of programs, goals and processes most relevant to the CNI community, and will offer observations on trends and priorities in the fields they monitor. Grant programs to be discussed may be, but are not limited to, the following: Research and Development, Digital Projects for the Public, NEH-Mellon Fellowships for Digital Publication, and Digital Humanities Advancement Grants (NEH); National Leadership Grants for Libraries and Laura Bush 21st Century Librarian Program (IMLS); Recordings at Risk and Digitizing Hidden Special Collections and Archives (CLIR); and Access to Historical Records (NHPRC). Ample opportunity will be allowed for audience dialogue.

neh.gov imls.gov  
clir.org archives.gov/nhprc

*Speakers: Jeremy Frumkin (U. Arizona)*

The rapid growth and adoption of data science approaches to research bring new challenges and opportunities for managing research data. Interestingly, while there is a great deal of attention being paid towards open access of research outputs, and federal funding agencies (as well as other funders) are increasingly adopting open access requirements for scholarly and data outputs, at the same time many federal funding agencies are also beginning to require new controls for research data, and universities are taking more seriously the need to manage their institutional risk by providing institutional support for long-standing data control standards (such as HIPAA), as well as developing support for newer control standards (such as CUI - controlled, unclassified information). This case study will describe how the University of Arizona has approached supporting research data that requires controls, including successes and lessons learned, and will conclude with thoughts towards future opportunities and challenges both for research computing units and academic libraries.

<https://it.arizona.edu/cui>

**6.6.2 Health Sciences Data Archive**

*Speakers: Mara Blake, Sayeed Choudhury (JHU)*

The provost at Johns Hopkins University (JHU) asked the Libraries to update the existing institutional JHU Data Archive to host, share, and preserve health sciences data. Consequently, the Libraries have worked with faculty and the Institutional Review Board (IRB) from the School of Medicine, the JHU Data Trust, which sets policies and processes for use of clinical data, the School of Medicine and University Legal Counsel, and the Chief Information Security Officer. This process has illuminated numerous issues that are relevant to HIPAA, data use and sharing agreements, data policies, data transfer and curation workflows, and user interface development. There are important observations to share with the community in terms of de-identification or anonymization of data, scanning tools for personally identifiable information or personal health information, and workflows for both static and dynamic data from a variety of technology platforms (including high-performance computing facilities). The JHU Libraries have also considered how to integrate our updated Data Archive with the Public Access Submission System (PASS) such that researchers could submit cited data at the same time as they deposit articles into PubMedCentral and institutional repositories.

### 6.6.3 Almost Open: Benefits, Challenges, and Design of an Authorized-access Research Data Enclave

Congressional A (West Wing)

*Speakers: Jeffrey Spies (221B), Rick O. Gilmore (Penn State)*

Openness with minimal restrictions is ideal but sometimes not feasible. For example, openness can be impractical or impossible when identifiable data about human subjects must be protected. Furthermore, the research community has a long way to go until openness is both normative and adequately resourced. Regardless of the practical or cultural barriers to unmitigated openness, it is still critical to promote reuse, reproducibility, and replicability to increase the quality and efficiency of the research process and the breadth and depth of research impact. Research data enclaves, where access is restricted to a trusted network of individuals, can foster these values while providing an iterative step toward wider openness.

In this talk, we will discuss our experience in designing, running, and now updating the technology behind the free, open-source behavioral video repository and restricted-access, authorized research environment Databrary, which has been in operation at New York University since 2014 and now includes 540 authorized institutions. In particular, we will focus on the pathways we are developing to promote wider openness wherever possible and our current curation-first design principle. More broadly, we will explore the benefits and challenges of such environments and discuss "almost open" as a strategy for the open community.

<https://nyu.databrary.org/>

### 6.7.1 The Academic Library as IT Partner: Supporting Sponsored Research at Auburn University

Congressional B (West Wing)

*Speakers: Denise Baker, Mallory Lucier-Greer, Aaron Trehub (Auburn U.)*

Academic libraries are seeking new ways to better align with current university research practice and to engage as vital partners in campus research. For the past two years, the Auburn University Libraries have supported the University's research enterprise in a new way: by adopting a new collaborative model and serving as high-level information technology (IT) and data-management consultants to faculty researchers who are pursuing external funding. The Military REACH Project at Auburn University is a successful example of this model in action. Funded by the Departments of Agriculture and Defense (USDA/NIFA Award No. 2017-48710-27339), the project has resulted in an ongoing functional partnership between the library, a research team in another college, and federal funding agencies. Working closely with the Military REACH team in the College of Human Sciences, the library's IT department contributed to the original funding proposal and has provided guidance on network architecture, web development, IT tools and solutions, sustainability, data management, accessibility, usage statistics, and automated methods for identifying recently published research. The Military REACH Project is now in its third year at Auburn University and seems likely to continue; indeed, it has highlighted the library's value as an IT partner and led to research partnerships and collaborative funding proposals with other units on campus. Associate Professor and project Principal Investigator Mallory Lucier-Greer will describe the functional partnership between her college and the library from the researcher's perspective, with input from library IT managers Denise Baker and Aaron Trehub. We believe that this project briefing will serve as a useful segue into the one-day Mellon-funded symposium on "Critical Roles for Libraries in Today's Research Enterprise" on Wednesday, December 11.

<https://militaryreach.auburn.edu/#/>

### 6.7.2 Redesigning the Researcher Library Experience: Case Studies, Key Questions

*Speakers: Tom Hickerson, John Brosz (U. Calgary)*

Libraries have developed ways of designing their online user experience and student experience. What is our design method for research? Researchers no longer depend on libraries for "search" and the resources purchased and licensed are of diminishing centrality. Providing new functional services, infrastructure, and expertise is essential. Services such as data curation, visualization, and geospatial analytics are now core capabilities. Capacities such as metadata, digitization, and copyright are being deployed in new ways. Library expertise and collaborative spaces are of critical importance. Yet, these often remain hidden by a traditional image of libraries, or by organizational models obscuring vital points of intersection for both researchers and library staff. In this session, we will identify essential elements derived from case studies and offer recommendations and checklists for assessing, redesigning, and repositioning the library's presence in campus research.

<https://library.ucalgary.ca/libraryresearchplatform>