



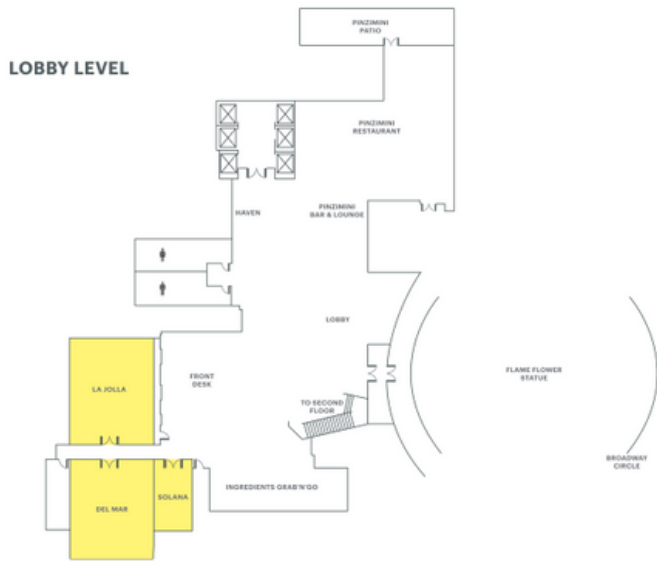
SPRING 2024 SCHEDULE

COALITION FOR NETWORKED INFORMATION

MEMBERSHIP MEETING
WESTIN GASLAMP QUARTER, SAN DIEGO, CA
MARCH 25-26, 2024

WESTIN GASLAMP QUARTER

LOBBY LEVEL



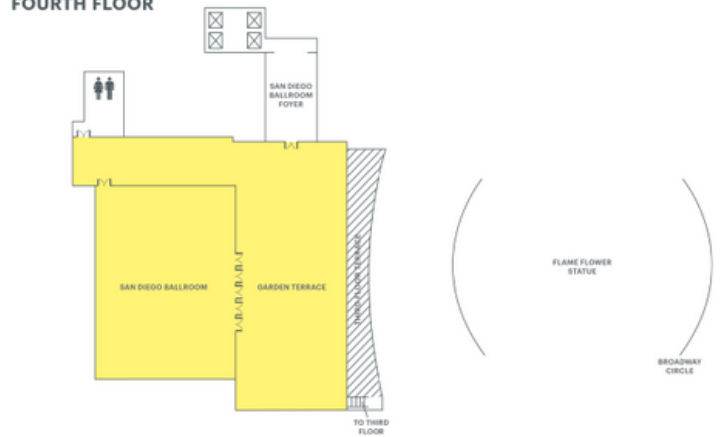
SECOND FLOOR



THIRD FLOOR



FOURTH FLOOR





Coalition for Networked Information

Coalition for Networked Information Spring 2024 Membership Meeting

March 25–26, 2024

San Diego, CA

#cni24s

Network: Westin_CONFERENCE

Wi-Fi Passcode: CNIs24conf

For the most up-to-date information visit the digital,
mobile-friendly schedule:

<https://cnispring24mtg.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. Please bring concerns to our attention by contacting a member of the CNI staff.

cni.org

CNI Spring 2024 Schedule-at-a-Glance

MONDAY, MARCH 25

11:00am Registration Opens (*California, Plaza, and Santa Fe Foyers*)

11:15am First-time Attendees (*Santa Fe*)

12:00pm Refreshment Break (*California, Plaza, and Santa Fe Foyers*)

1:00pm Opening Plenary: Computing Futures in a Changing World (*California A&B*)

2:15pm Refreshment Break (*California, Plaza, and Santa Fe Foyers*)

2:45pm PROJECT BRIEFINGS

1.1 The T in GPT: Transformers for Cultural Heritage Work	<i>California A&B</i>
1.2 Future-Proofing Research Data Repositories: Keeping Up With the ML/AI Revolution	<i>California C</i>
1.3 Scaling Instrument Science in the FAIR Age	<i>Santa Fe</i>

3:15pm Passing Break

3:25pm PROJECT BRIEFINGS

2.1 In Conversation with Daniel Reed	<i>California A&B</i>
2.2 Models of Support for Data Science: The Perspective of Two Libraries	<i>California C</i>

3:55pm Refreshment Break (*California, Plaza, and Santa Fe Foyers*)

4:25pm PROJECT BRIEFINGS

3.1 ARL/CNI Task Force on Scenario Planning for AI/ML Futures	<i>California A&B</i>
3.2 Reimagine Descriptive Infrastructure: Dreaming and Enacting Change	<i>California C</i>
3.3 Information Infrastructure to Address Societal Grand Challenges	<i>Santa Fe</i>

5:10pm Passing Break

MONDAY, MARCH 25 Continued

5:20pm

Lightning Round (*California A&B*)

6:05–7:30pm

Reception (*San Diego Ballroom/Garden Terrace*)

TUESDAY, MARCH 26

7:45am

Breakfast (including discussion tables) (*San Diego Ballroom/Garden Terrace*)

- Digital Preservation Coalition (DPC) in the USA
- Info Infrastructure for Grand Challenges
- Infrastructure for Open Scholarship
- Open Source Communities
- Scenario Planning for AI/ML Futures (4 tables)

Unmoderated

- Data Privacy and Security
- Developments in Open Access

Other topics and/or facilitators may be added; check online schedule for final list.

9:00am

PROJECT BRIEFINGS

4.1 Navigating the New Era: The Impact of GenAI on Information Discovery and Literacy	<i>California A&B</i>
4.2 Opening Collections of Marginalized Voices through Crowdfunding and Crowdsourcing	<i>California C</i>
4.3 Linked Data in Production: Moving Beyond Ontologies	<i>Santa Fe</i>

9:45am

Passing Break

10:00am

PROJECT BRIEFINGS

5.1 Cloud Labs and Self-Driving Laboratories Update and Futures	<i>California A&B</i>
5.2 So You Made an Institutional Strategy, Now What? A Canadian Approach to RDM Strategy Implementation	<i>California C</i>
5.3 Combining Micropublications into a Sustainable Back End and an Integrated Reading Environment	<i>Santa Fe</i>

10:30am

Refreshment Break (*California, Plaza, and Santa Fe Foyers*)

TUESDAY, MARCH 26 Continued

11:00am

PROJECT BRIEFINGS

6.1 Academic Applications of AI: Building Collaboration Among Libraries, IT, Faculty, and Students at SDSU	<i>California A&B</i>
6.2 Handling Academic Copyright and AI Research Questions as the Law Develops	<i>California C</i>
6.3 Navigating Generative AI: Early Findings and Implications for Research, Teaching, and Learning	<i>Santa Fe</i>

12:00pm

Lunch (*San Diego Ballroom/Garden Terrace*)

1:15pm

PROJECT BRIEFINGS

7.1 Making Research Data Publicly Accessible: Estimates of Institutional & Researcher Expenses	<i>California A&B</i>
7.2 NRP: Open Cyberinfrastructure for Research	<i>California C</i>
7.3 Recommitment and Recalibration: HathiTrust's Strategic Vision	<i>Santa Fe</i>

2:00pm

Passing Break

2:15pm

PROJECT BRIEFINGS

8.1 A New Approach to Data-Intensive Research Support: Computational Methods and Data at Yale University Library	<i>California A&B</i>
8.2 3D Digital Herbarium & 3D Exhibits4Learning	<i>California C</i>
8.3 Unveiling Whale Wisdom: Digitizing the Patagonian Right Whale Dataset	<i>Santa Fe</i>

2:45pm

Passing Break

3:00pm

Closing Session: Community Conversation (*California A&B*)

3:30pm

Meeting Adjourns

CNI Spring 2024 Membership Meeting

Opening, Lightning Round, Breakfast Discussions, Closing

MONDAY, MARCH 25

1:00–2:15 pm
California A&B

Opening Plenary: Computing Futures in a Changing World

Daniel Reed, Presidential Professor in Computational Science, University of Utah and Chair, U.S. National Science Board

The computing ecosystem writ large is in extraordinary flux, with profound implications for research and discovery, economic competitiveness, and national security. From the geopolitics of semiconductor foundries, their exponentially rising costs, and the slowing of Moore's Law through the rise of cloud hyperscalers and shifting loci of economic leverage to the profound and uncertain effects of deep learning and artificial intelligence, the future promises to be qualitatively different from the past. Overlaying these technical and economic issues are equally important questions of policy, law, and ethics, strategic research investment, and STEM workforce development. This talk will discuss these challenges, how we reached this point, some opportunities, and where we might go.

5:25–6:05 pm
California A&B

Lightning Round

Transforming Libraries for the Future: Elevating Service Innovation with Generative Artificial Intelligence and Prompt Engineering

Yinlin Chen, Virginia Tech University

Catalyzing African Community Archives for Social Good

Chris Prom, University of Illinois at Urbana-Champaign

National Information Standards Organization's Interoperable System of Controlled Digital Lending (IS-CDL) Recommended Practice

Mike Furlough, HathiTrust

Let A Thousand Flowers Bloom: An Organic Funding Model to Incubate Library Transformation

Zhiwu Xie, University of California, Riverside

Infrastructure for Open Scholarship

Ken Klingenstein, Internet2

The Digital Preservation Coalition (DPC) Americas: DPC's New US-based Program

Jon Dunn, Indiana University

A brief roundup of breakfast discussion tables will finish out the Lightning Round session.

Full presentation descriptions are available on the CNI website:

<https://www.cni.org/mm/spring-2024/project-briefings-breakout-sessions-s24>

7:45–9:00 am
San Diego
Ballroom/Garden
Terrace

Breakfast Discussion Tables

During breakfast, attendees may join optional table discussions on pre-determined topics. Designated tables will be marked by table signs and located throughout the breakfast area. There is no signup; participation is first come, first served.

Digital Preservation Coalition (DPC) in the USA

Jon Dunn, Indiana University and Chris Prom, University of Illinois at Urbana-Champaign

Info Infrastructure for Grand Challenges

Donald Waters, Coalition for Networked Information

Infrastructure for Open Scholarship

Ken Klingenstein, Internet2

Open Source Communities

Heather Greer Klein, Samvera

Scenario Planning for AI/ML Futures (Four Tables)

Karen Estlund (Colorado State University), Leo Lo (University of New Mexico), Christy Long (University of Oregon), & Catherine Steeves (Western University)

Unmoderated

- **Data Privacy and Security**
- **Developments in Open Access**

Other topics and/or facilitators may be added.

3:00–3:30 pm
California A&B

Closing Session: Community Conversation

There will be no formal closing plenary. Instead, the meeting will close with something new: a lightly structured half-hour interactive session, which will allow participants to share what most captivated or surprised them at the meeting and any big takeaways. This will leave attendees with a greater sense of synthesis and engagement of shared insights and collaborative purpose. Short reflections will be invited from all who wish to share them (or at least as many as there is time to accommodate). *This session will not be recorded.*

CNI Spring 2024 Membership Meeting

PROJECT BRIEFINGS

MONDAY, MARCH 25

2:45–3:15 pm
California A&B

1.1 The T in GPT: Transformers for Cultural Heritage Work

Peter Leonard (Stanford University)

Although commercial large language models such as ChatGPT have captured the public imagination, the math underlying these conversational systems has important implications for libraries in key areas such as speech-to-text, image description, and handwriting recognition. Collections previously isolated from text-mining practices now stand at the cusp of becoming computationally tractable. Also, in an era of increased attention to accessibility requirements, providing textual proxies of audio-visual and image-based content can lessen legal exposure. The session will consider applications of transformers-based artificial intelligence models from an international perspective, with a focus on what they enable and what changes they imply for digital libraries.

2:45–3:15 pm
California C

1.2 Future-Proofing Research Data Repositories: Keeping Up With the Machine Learning/Artificial Intelligence Revolution

Stephanie Labou (University of California, San Diego)

Research data repositories must prepare for the inevitable influx of machine learning (ML) and artificial intelligence (AI) input and output data, training models, and documentation necessary to reproduce and reuse ML and AI components. A team from the University of California San Diego Library recently assessed ML objects in eight generalist and specialist repositories to identify how institutional repositories can adapt current structures and processes to better meet ML practitioner preferences and enhance the findability and reusability of repository content. As best practices for curation of ML and AI research evolve, this session will include suggestions for a set of relatively small changes with potentially significant impacts on positioning academic library repositories for the next generation of research data, as well as making existing repository content itself ML- and AI-ready.

2:45–3:15 pm
Santa Fe

1.3 Scaling Instrument Science in the FAIR Age

Vas Vasiliadis (University of Chicago)

A growing number of research activities require large-scale data management. In particular, high resolution imaging instruments such as cryogenic electron microscopes and synchrotron beamlines require automation of data flows to increase throughput and researcher productivity, as well as to ensure the instrument remains highly utilized. Combined with the increasingly collaborative nature of research, this necessitates infrastructure that makes the resulting data products

more FAIR (findable, accessible, interoperable, and reusable). The Globus team at the University of Chicago and Argonne National Laboratory have worked with many institutions to build solutions for automating instrument data management, from the point of capture through publication and re-use. The session will present scenarios from research universities and national facilities that illustrate the implementation of common use cases and describe how they were enabled by the Globus platform.

3:25–3:55 pm
California A&B

2.1 In Conversation with Daniel Reed

Daniel Reed (University of Utah and National Science Board)
Clifford Lynch (Coalition for Networked Information)

The session will be a continuation of Daniel Reed's opening plenary, Computing Futures in a Changing World. Drawing on the plenary and Reed's deeply informed view of developments in computational science and their interplay with the broader national political and policy spheres, Reed and CNI Director Clifford Lynch will explore a few topics in greater depth and examine additional issues of particular interest to the CNI community.

3:25–3:55 pm
California C

2.2 Models of Support for Data Science: The Perspective of Two Libraries

David Minor (University of California, San Diego)
Jan Brase (Göttingen State and University Library)
Bela Gipp (University of Göttingen)

The University of Göttingen and the University of California, San Diego (UCSD) both have new but robust data science departments. The libraries of both institutions have committed significant effort to support the emergence of new data-intensive disciplines, majors, and students' needs. But their focus has been from differing perspectives: the State and University Library Göttingen has created a department within the library that shares a faculty member, centering data science within the organization. The UCSD Library maintains a more traditional support role, focused on pedagogical support in student projects. Nevertheless, both institutions have started to explore how cooperation between these departments can be realized and how they can offer students international experience through joint exchanges and joint projects. The session will share successes and challenges in these experiences and explore the new collaboration between the universities.

4:25–5:10 pm
California A&B

3.1 ARL/CNI Task Force on Scenario Planning for Artificial Intelligence/Machine Learning Futures

Christy Long (University of Oregon)

Elisabeth Long (Johns Hopkins University)

Catherine Steeves (Western University)

Keith Webster (Carnegie Mellon University)

Clifford Lynch (Coalition for Networked Information)

Moderator: Judy Ruttenberg (Association of Research Libraries)

The joint ARL/CNI Task Force was formed in late 2023 and has met several times, including an in-person workshop in Washington, DC, in February 2024, to develop draft scenarios exploring possible futures for the research enterprise and research libraries as the deployment of artificial intelligence (AI)-based technologies unfolds over the next decade. The scenarios are intended to offer insight into how the numerous uncertainties surrounding AI technologies and their deployment may play out and to explore some possible implications. At this session, after framing the scenario process, objectives, and potential uses of the scenarios, members of the Task Force will introduce the four draft scenarios, comment on the key uncertainties that the group has tried to capture in the portfolio of scenarios, and seek feedback from the audience around the plausibility and consistency of the scenarios.

<https://www.arl.org/news/association-of-research-libraries-and-coalition-for-networked-information-appoint-joint-task-force-on-scenario-planning-for-ai-and-machine-learning-futures/>

4:25–5:10 pm
California C

3.2 Reimagine Descriptive Infrastructure: Dreaming and Enacting Change

Merrilee Proffitt (OCLC)

Camille Callison (University of the Fraser Valley and National Indigenous Knowledge & Language Alliance (NIKLA), Tāltān Nation)

Stacy Allison-Cassin (Dalhousie University and National Indigenous Knowledge & Language Alliance (NIKLA), Métis Nation of Ontario)

Reimagine Descriptive Workflows (RDW) was a community-informed project funded by the Andrew W. Mellon Foundation that took place in 2021 and was structured to surface issues and identify opportunities to effect lasting change in descriptive practices. Among the goals set by the participants were to: acknowledge a need to change the current system; identify opportunities to engage in collaborative problem-solving; and develop concrete approaches to enable reimagined descriptive metadata practices. A report that documents the convening and its findings, *Reimagine Descriptive Workflows: A Community-informed Agenda for Reparative and Inclusive Descriptive Practice*, reflects that "power and bias in collections is hard coded from the beginning of the descriptive workflow process," and that powerful naming and labeling systems, which include content standards and data communication formats, "can create systemic imbalances beyond the inherent problems of labeling and description." The project briefing will focus on a

number of efforts that have taken place between 2022 and 2024 that go beyond addressing individual subject headings and that have been focused on the difficult work of repairing existing systems and workflows (such as those provided by OCLC) and creating new systems (such as the Mellon-funded, Indigenous-led National Indigenous Knowledge and Language Alliance Respectful Terminology Platform Project) in order to disrupt a cycle of harm in library descriptive practices.

<https://www.nikla-ancla.com/>

<https://www.nikla-ancla.com/respectful-terminology>

<https://www.oclc.org/research/areas/community-catalysts/reimagine-descriptive-workflows.html>

<https://www.oclc.org/research/publications/2022/reimagine-descriptive-workflows.html>

4:25–5:10 pm
Santa Fe

3.3 Information Infrastructure to Address Societal Grand Challenges

Donald Waters (Coalition for Networked Information)

In late 2022, the Coalition for Networked Information (CNI) launched the CNI Senior Scholars Program. The program is designed to facilitate efforts to explore broad research questions that align with CNI's program and interests. The inaugural CNI Senior Scholar, Donald Waters, is studying the information infrastructure that universities need to best address climate change as a grand challenge of immense public interest. Relying on background literature and extensive interviews with scholars and information experts at ten selected universities, Waters is now preparing a final report. The briefing session will present and invite discussion of his findings, beginning with a definition of the "wicked" nature of climate change as a grand challenge and an outline of mission-driven efforts of universities to scale up and better integrate faculty research on this complex issue. He will suggest that libraries and information technology organizations could contribute substantially to this scaling and integration by: (a) focusing on the requirements of interdisciplinary research; (b) supporting such research in specific university centers and institutes; (c) enhancing support for faculty collection and analysis of data; and (d) facilitating faculty engagement with local communities.

<https://www.cni.org/events/cni-workshops/senior-scholars-program/information-infrastructure-to-address-societal-grand-challenges>

<https://www.cni.org/topics/ci/information-infrastructure-to-address-societal-grand-challenges>

9:00–9:45 am
California A&B

4.1 Navigating the New Era: The Impact of Generative Artificial Intelligence on Information Discovery and Literacy

Emily Singley (Elsevier)

Leo Lo (University of New Mexico)

Elias Tzoc (Clemson University)

Generative artificial intelligence (GenAI) is revolutionizing traditional search methods, leading to significant changes in information discovery and literacy. The panel will explore the shift from keyword-based searches to conversational, AI-driven queries and their impact on scholarly information-seeking behaviors. Discussions will include an analysis of a university's GenAI pilot project and its varied effects across different academic disciplines. Additionally, insights from a leading academic publisher will highlight the development of AI-enhanced platforms and the role of user feedback. The session will conclude with strategies for integrating GenAI into information literacy frameworks, enhancing librarian skills, and exploring new collaborative initiatives, providing a crucial dialogue for libraries and content providers in the context of GenAI's rapidly changing landscape.

9:00–9:45 am
California C

4.2 Opening Collections of Marginalized Voices through Crowdfunding and Crowdsourcing

Michael Levine-Clark (University of Denver)

Rhonda Manzanares (Colorado State University - Pueblo)

Jasmine Wilson (Reveal Digital)

Academic and cultural institutions around the world are stewards of valuable but often hidden primary source and special collections content that can be of great use for teaching, learning, and research. More and more faculty want to teach with primary source/special collections, and more and more students and early career researchers need to conduct their research using these same materials. For the past decade, Reveal Digital has been offering a library crowdfunding (Fund2Open) and crowdsourcing approach to publishing open access scholarly collections of primary source material focused on social movements, hidden voices, and marginalized communities. Collections such as Independent Voices, Student Activism, American Prison Newspapers, and others have been built (funded and sourced) by librarians for a broad range of potential users while being hosted on JSTOR and preserved in Portico. The presentation will include discussion about some important lessons learned from trying to develop and deliver collections in this model, and why this model is an appealing funding option as well as a useful sourcing alternative with regard to collections that have traditionally been marginalized or not seen as commercially viable.

<https://about.jstor.org/revealdigital/>

9:00–9:45 am
Santa Fe

4.3 Linked Data in Production: Moving Beyond Ontologies

David Newbury (Getty)

Over the past six years, Getty has been engaged in a project to transform and unify its complex digital infrastructure for cultural heritage information. One of the project's core goals was to provide validation of the impact and value of the use of linked data throughout this process. With museum, archival, media, and vocabularies in production and others underway, the session will share some of the practical implications (and pitfalls) of this work—particularly as it relates to interoperability, discovery, staffing, stakeholder engagement, and complexity management. The session will also share examples of how other organizations can streamline their own, similar work going forward.

<http://getty.edu/art/collection/>

<http://getty.edu/research/collections/>

<http://vocab.getty.edu>

<https://www.getty.edu/projects/remodeling-getty-provenance-index/>

10:00–10:30 am
California A&B

5.1 Cloud Labs and Self-Driving Laboratories Update and Futures

Sayed Choudhury (Carnegie Mellon University)

Keith Webster (Carnegie Mellon University)

Cliff Lynch (Coalition for Networked Information)

The National Science Foundation Technology, Innovation, and Partnerships (TIP) Directorate recently funded two workshops, one at Carnegie Mellon University and one at North Carolina State University, focused on automated science facilities encompassing cloud labs and self-driving laboratories. The session will provide an overview of observations, lessons learned, challenges, and opportunities toward the development of a national network of such research facilities, with an emphasis on open science.

<https://cloudlab.cmu.edu/>

10:00–10:30 am
California C

5.2 So You Made an Institutional Strategy, Now What? A Canadian Approach to Research Data Management Strategy Implementation

Alison Hitchens (University of Waterloo)

Caroline Hyslop (University of Ottawa)

All Canadian research institutions eligible for federal Tri-Agency funding have created and published an institutional strategy for research data management (RDM). In their work to comply with the Tri-Agency requirement for a strategy, Canadian universities reached out to one another, largely through networks of academic libraries, to share approaches, discuss challenges, and brainstorm solutions. These conversations revealed an appetite for a more coordinated approach and a desire amongst academic institutions to bring together libraries, research offices, and information technology (IT) to discuss common themes and find ways to move forward together for

9:00–9:45 am
Santa Fe

4.3 Linked Data in Production: Moving Beyond Ontologies

David Newbury (Getty)

Over the past six years, Getty has been engaged in a project to transform and unify its complex digital infrastructure for cultural heritage information. One of the project's core goals was to provide validation of the impact and value of the use of linked data throughout this process. With museum, archival, media, and vocabularies in production and others underway, the session will share some of the practical implications (and pitfalls) of this work—particularly as it relates to interoperability, discovery, staffing, stakeholder engagement, and complexity management. The session will also share examples of how other organizations can streamline their own, similar work going forward.

<http://getty.edu/art/collection/>

<http://getty.edu/research/collections/>

<http://vocab.getty.edu>

<https://www.getty.edu/projects/remodeling-getty-provenance-index/>

10:00–10:30 am
California A&B

5.1 Cloud Labs and Self-Driving Laboratories Update and Futures

Sayed Choudhury (Carnegie Mellon University)

Keith Webster (Carnegie Mellon University)

Cliff Lynch (Coalition for Networked Information)

The National Science Foundation Technology, Innovation, and Partnerships (TIP) Directorate recently funded two workshops, one at Carnegie Mellon University and one at North Carolina State University, focused on automated science facilities encompassing cloud labs and self-driving laboratories. The session will provide an overview of observations, lessons learned, challenges, and opportunities toward the development of a national network of such research facilities, with an emphasis on open science.

<https://cloudlab.cmu.edu/>

10:00–10:30 am
California C

5.2 So You Made an Institutional Strategy, Now What? A Canadian Approach to Research Data Management Strategy Implementation

Alison Hitchens (University of Waterloo)

Caroline Hyslop (University of Ottawa)

All Canadian research institutions eligible for federal Tri-Agency funding have created and published an institutional strategy for research data management (RDM). In their work to comply with the Tri-Agency requirement for a strategy, Canadian universities reached out to one another, largely through networks of academic libraries, to share approaches, discuss challenges, and brainstorm solutions. These conversations revealed an appetite for a more coordinated approach and a desire amongst academic institutions to bring together libraries, research offices, and information technology (IT) to discuss common themes and find ways to move forward together for

the benefit of the Canadian research community. In September 2023, the University of Waterloo hosted a national workshop, inviting participants from across Canada to discuss the next steps in RDM strategy implementation. Importantly, every participating institution sent three representatives: one from each of their library, research office, and IT departments. The result was a rich and dynamic discussion of the challenges and potential for excellence in RDM across the country. The session will describe the workshop, its purpose, and structure and will explore the challenges and next steps for RDM strategy implementation.

<https://uwaterloo.ca/research-data-management-community-workshop/>
<https://science.gc.ca/site/science/en/interagency-research-funding/policies-and-guidelines/research-data-management/published-institutional-research-data-management-strategies>

10:00–10:30 am
Santa Fe

5.3 Combining Micropublications into a Sustainable Back End and an Integrated Reading Environment

Gregory Crane (Tufts University)

The session will report on long-term work that aims to integrate as many different categories of data as possible about cultural spaces. While the focus is Greco-Roman culture, the methods are much more general. Where the initial challenge lay in creating materials, the availability of Creative Commons and other open licenses and openly licensed materials has fostered the emergence of a vigorous ecosystem of open data research and research publications. Efforts increasingly concentrate on the challenges of working with data from this growing network of interacting projects. Publication classes include: linguistic annotations; named entity classification and linking with associated visualizations (such as dynamically generated maps); analysis of metrical form; word and phrase level alignments between source texts and translations; links between textual and linguistic datasets and regions of interest from manuscripts and other text-bearing objects; natural language annotations that can address particular strings in particular versions of a canonical work. Challenges include managing individual and group credits on openly licensed, machine-actionable publications as new projects upload and modify prior work. In creating a prototype focused on Homeric Epic alone, researchers wished to draw upon data from more than 15 projects and were integrating data based on content that they had published. This approach is a new and much more decentralized form of intellectual production that blends human and automatically produced content and provides opportunities for students and members of the public to contribute to research. While the challenges are daunting, the decentralized approach has the potential to open up and invigorate academia, particularly the humanities. The researchers welcome feedback and opportunities for collaboration.

The project was funded most recently by the Mellon Foundation, Harvard University's Center for Hellenic Studies, and the National Endowment for the Humanities.

<https://sites.tufts.edu/perseusupdates/2024/01/16/towards-a-new-perseus-update/>
<https://www.perseus.tufts.edu/hopper/>
<https://scaife.perseus.org/>
<https://beyond-translation.perseus.org/reader/urn:cts:greekLit:tlg0012.tlg001.perseus-grc2:1.1-1.7>

11 am–12 pm
California A&B

6.1 Academic Applications of Artificial Intelligence: Building Collaboration Among Libraries, IT, Faculty, and Students at San Diego State University

Scott Walter, EJ Sobo, and Abir Mohamed (San Diego State University)

Over the past decade, the University Library and the Division of Information Technology's Instructional Technology Services have developed a successful record of partnership promoting student-facing innovation in areas including the design of technology-enhanced classrooms, open educational resources, online learning, and digital scholarship. Building on these successful models, and engaging project-specific partners from the Center for Teaching and Learning and the Division of Student Affairs and Campus Diversity, San Diego State University (SDSU) has launched a suite of data-informed initiatives around the academic applications of artificial intelligence (AI), including a faculty learning community, survey of emergent student practices, and microcredential program aimed at helping instructional faculty and librarians to apply generative AI technology efficiently, effectively, and ethically in face-to-face and online instruction, and as part of student success services offered in the library and through campus tutoring programs. The panel will present data from the largest current survey of student use of AI technologies in the classroom (7,000+ respondents), as well as data drawn from the pilot offering of the Academic Applications of AI (AAAI) Microcredential (spring 2024). It will also explore how student engagement and professional development for SDSU staff and faculty may be employed in scaling up a system-wide solution for these emergent issues across California State University, the largest and most diverse public, four-year, higher education system in the United States.

<https://its.sdsu.edu/ai>

11 am–12 pm
California C

6.2 Handling Academic Copyright and Artificial Intelligence Research Questions as the Law Develops

*Jonathan Band (Counsel to the Library Copyright Alliance)
Timothy Vollmer (University of California, Berkeley)*

The United States Copyright Office and courts in many United States jurisdictions are struggling to address complex copyright issues related to the use of generative artificial intelligence (AI). Meanwhile, academic research using generative AI is proliferating at a fast pace and researchers still require legal guidance on which sources they may use, how they can train AI legally, and whether the reproduction of source material will be considered infringing. The session will include a discussion of current perspectives on copyright and generative AI in academic research.

11 am–12 pm
Santa Fe

6.3 Navigating Generative Artificial Intelligence: Early Findings and Implications for Research, Teaching, and Learning

Beth LaPensee & Kevin Guthrie (ITHAKA)

Starting in mid-2023, ITHAKA began investing in and engaging directly with generative artificial intelligence (AI) in two broad areas: a generative AI research tool on the JSTOR platform and a collaborative research project led by Ithaka S+R. These technologies are so crucial to our futures that working directly with them to learn about their impact, both positive and negative, is extremely important.

The presentation will share early findings that illustrate the impact and potential of generative AI-powered research based on what JSTOR users are expecting from the tool, how their behavior is changing, and implications for changes in the nature of their work. The findings will be contextualized with the cross-institutional learning and landscape-level research being conducted by Ithaka S+R. By pairing data on user behavior with insights from faculty and campus leaders, the session will share early signals about how this technology-enabled evolution is beginning to take shape.

<https://www.jstor.org/generative-ai-faq>

1:15–2:00 pm
California A&B

7.1 Making Research Data Publicly Accessible: Estimates of Institutional & Researcher Expenses

Cynthia Hudson Vitale (Association of Research Libraries)

Jennifer Moore (Washington University in St. Louis)

Jake Carlson (University of Buffalo)

Alicia Hofelich Mohr (University of Minnesota)

Institutional expenses for public access to research data continue to be of significant importance for university administration, including the research office, campus information technology, the libraries, and other institutes and offices on academic campuses. Articulating these expenses is critical to effectively advocate and plan for compliance with federal policies for data management and sharing. The panel will report on recent mixed-methods research conducted on the institutional expenses for these policies from a researcher and administrator perspective.

1:15–2:00 pm
California C

7.2 National Research Platform: Open Cyberinfrastructure for Research

Mahidhar Tatineni (University of California, San Diego)

Research teams are increasingly both multi-disciplinary and multi-institutional, and their workflows incorporate large datasets from many different locations. The National Research Platform (NRP) enables interoperable federations of compute and data resources from multiple projects and provides an open cyberinfrastructure solution for research. The talk will cover the innovations and additions to NRP in the past year. The rise of artificial intelligence (AI) in research has redefined academic infrastructure needs. Examples illustrating NRP use for AI research and education will be discussed.

The session builds on two previous CNI briefings (“Towards a High-Performance National Research Platform Enabling Digital Research” and “Towards an Open Global Cyberinfrastructure Enabling Digital Research”) and in addition to current examples will provide information on forthcoming developments on NRP.

1:15–2:00 pm
Santa Fe

7.3 Recommitment and Recalibration: HathiTrust's Strategic Vision

Mike Furlough (HathiTrust)

Claire Stewart (University of Illinois Urbana-Champaign)

HathiTrust's new Strategic Vision is the result of a year-long project to closely evaluate its accomplishments and potential through the eyes of its members, governance bodies, and staff, as well as non-members and peer organizations. With the guidance of a Strategic Visioning Task Force, the team reviewed HathiTrust's place in the current research, educational, socio-political, technological, and environmental landscapes to consider how it can respond to current challenges. The result is an ambitious agenda of work, focused on HathiTrust's unique strengths and potential, and based on a purposeful recommitment to its foundational mission. As HathiTrust continues to develop transformational uses of the world's largest collection of digitized library collections, it also seeks to strengthen its overall ecosystem to address challenges. The project briefing will overview how the organization is positioning itself to achieve its core strategic objectives—increasing access to and use of an expanded, diversified, and enhanced collection—by further expanding lawful access to in-copyright works, scaling the identification of public domain works, and integrating computational methods to the corpus for user access and text and metadata improvement. The briefing will also include discussion of the methods and findings as well as some reflections on the organization's growth and development.

<https://www.hathitrust.org/about/mission-history/strategic-visioning/>

2:15–2:45 pm
California A&B

8.1 A New Approach to Data-Intensive Research Support: Computational Methods and Data at Yale University Library

Rebecca Dikow (Yale University)

Computational Methods and Data is a new department within the Research and Learning portfolio at Yale University Library. The department was formed in response to the high demand for data-intensive research support and the need to break down disciplinary silos to provide high-quality data services. Researchers across disciplines and at every academic level are navigating ever-rising expectations to gather, analyze, visualize, and publish datasets of increasing size and complexity. While users may have ready-access to artificial intelligence (AI) tools, the cloud, and open datasets, how to navigate these responsibly in the context of a research project is not at all straightforward. The department is composed of staff in the areas of Digital Humanities, Geospatial Services, Research Data Management, and Statistical Support Services, which were previously separate initiatives at the Library. The support modalities provided by staff in the department include research consultations, workshops, dataset acquisition, documentation, access to compute and storage,

formal instruction in university courses, and collaboration. The department also serves as a touchpoint to and collaborator with many other parts of the university, including academic departments, research centers, and core facilities. The present state, examples of current projects, and plans for growth for the new department will be detailed.

2:15–2:45 pm
California C

8.2 3D Digital Herbarium & 3D Exhibits4Learning

Cyril Oberlander & AJ Bealum (California State Polytechnic University, Humboldt)

The California State Polytechnic University, Humboldt Library groundbreaking project 3DHerbarium.org version one was released open source on January 25, 2024. This immersive 3D web-based interface with augmented and virtual reality views enhances botanical education for all ages, connecting botany students, faculty, citizen scientists, and the community to learn and share knowledge. The goal of version two is to extend the platform for any discipline, from archeology to zoology, and to further advance the research in creating layered animated models that enable learners to explore specimens from the whole body to internal microscopic structures.

The 3D Digital Herbarium advanced the team's understanding of 3D photogrammetry methods for modeling thin complex shapes, such as leaves and plants. Additionally, the team discovered the need for a similar tool in a variety of disciplines and is evaluating various research interests to develop the roadmap for version two, called 3D Exhibits4Learning, an open-source software platform and documentation for other institutions to develop 3D models and to create their own environments.

<https://sites.google.com/vt.edu/lib-collab-grant/>

2:15–2:45 pm
Santa Fe

8.3 Unveiling Whale Wisdom: Digitizing the Patagonian Right Whale Dataset

Harish Maringanti & Matt Brunsvik (University of Utah)

While academic librarians typically focus on readily available, electronic research outputs, a treasure trove of valuable data lies hidden in researchers' filing cabinets. This legacy data, often stored in print and slides spanning decades of work, remains largely inaccessible to the wider research community. This presents a significant challenge, particularly in fields like biodiversity, where historical information is crucial for analyzing trends, adaptations, and long-term relationships. At the University of Utah, a collaboration between the Marriott Library, the School of Biological Sciences, and Ocean Alliance resulted in a Council on Library and Information Resources (CLIR)-funded project that helped to bring a hidden collection into the digital age. The project resulted in digitizing a unique historical dataset: the Patagonian Right Whale Program data. Since 1971, researchers have conducted annual aerial surveys of the whale population, capturing over 70,000 slides. This project brought the entire collection—slides, negatives, aerial maps, and field notes—into the digital age. Not only were they digitized, but metadata was extracted and mapped from an old Microsoft Access database. The session will share project challenges and opportunities and explain how it will benefit researchers worldwide.

CALENDAR OF KEY MEETINGS

IS&T's Archiving 2024, Washington, DC–April 9–12, 2024

Open Repositories, Gothenburg, Sweden–June 3–6, 2024

CNI 2024 Fall Membership Meeting, Washington, DC–December 9–10, 2024

Joint Conference on Digital Libraries, 2024, TBA

Designing Libraries 2024, North Carolina State University, Fall 2024, TBA

CNI 2025 Spring Membership Meeting, Milwaukee, WI–April 7–8, 2025

CNI 2025 Fall Membership Meeting, Washington, DC–December 11–12, 2025

International Digital Curation Conference, 2025, TBA

For more information on CNI and its programs visit www.cni.org

IMAGE CREDITS

Front (left to right; top to bottom)

Dragonfly, Pear, Carnation, and Insect

1561–1562; illumination added 1591–1596

Credit: Joris Hoefnagel (Flemish/ Hungarian, 1542–1600) and Georg Bocskay (Hungarian, died 1575)

Courtesy: Getty Museum

CC0 1.0 Universal

RNA Polymerase II

Credit: David Bushnell, Ken Westover and Roger Kornberg, Stanford University

NIH funding from: National Institute of General Medical Sciences (NIGMS)

Courtesy: National Institutes of Health

CC BY-NC 2.0 DEED

Researchers precisely control movement of defects within active liquid crystals

Research supported by U.S. National Science Foundation grant DMR 2011854

Credit: Alexprague

Courtesy: National Science Foundation

Public Domain

Sunset (California Scenery)

Credit: Albert Bierstadt, 1864, published by L. Prang & Co.

Courtesy of Boston Public Library via Digital Commonwealth

Double Portrait Herm of Aristotle and Plato

Credit: Unknown artist/maker, late 2nd century A.D.

Courtesy: Getty Museum

Public Domain

Back:

Frontier Supercomputer

The system is the first to achieve the level of computing performance known as exascale, a threshold of a quintillion calculations per second.

Courtesy: U.S. Department of Energy, Oak Ridge National Laboratory

CC BY-NC 2.0 DEED

All images were slightly modified for design purposes.

Descriptions are from image sources unless otherwise noted.



cni

Coalition for Networked Information

21 DUPONT CIRCLE, SUITE 800
WASHINGTON, DC, 20036
202-296-5098
www.cni.org