

RCD Nexus Day 2024 Final Report

Dana Brunson
Internet 2
dbrunson@internet2.edu

Betsy Hillery
Purdue University
eahillery@purdue.edu

Lauren Michael
Internet2
lmichael@internet2.edu

John Hicks
Internet2
jhicks@internet2.edu

Daphne McCanse
CaRCC/Internet2
daphne@carcc.org

Timothy Middelkoop
Internet2
tmiddelkoop@internet2.edu

Patrick Schmitz
Semper Cogito Consulting
patrick@sempercogito.edu

With support from a Cyberinfrastructure Center of Excellence Pilot grant from the National Science Foundation, [RCD Nexus Day 2024](#) was held on July 21, 2024 in Providence Rhode Island, as an official co-located event of the [PEARC24 Conference](#). This workshop hosted 92 attendees from 48 institutions, including nineteen (19) from non-R1 higher education institutions, eleven (11) from non-higher education organizations, nine (9) from [EPSCoR](#) jurisdictions, six (6) from Hispanic Serving Institutions, three (3) from Historically Black Colleges and Universities, one (1) from a Tribal College, and one (1) from another type of Minority Serving Institution. Fourteen (14) attendees received travel grants that allowed them to attend not only RCD Nexus Day, but also the entire PEARC24 conference, with preference given to attendees from the categories listed above. Dana Brunson of Internet2, John Hicks of Internet2 and Daphne McCanse of CaRCC and Internet2 (contract) served as general workshop organizers. The day began with a short general session before dividing into the two workshops described in detail below.

Workshop 1: What We're Facing as Research Computing and Data (RCD) Professionals

Workshop Facilitators

General Facilitators:

Lauren Michael, independent consultant and Internet2 (contract); CaRCC Logistics co-Chair, People Network co-Coordinator, RCD Professionalization Working Group, Capabilities Model Working Group, Data-Facing Track Steering Committee

Timothy Middelkoop, Internet2 Research Engagement; CaRCC Staff Workforce Development Interest Group co-chair, Logistics

Embedded Breakout Facilitators:

Justin Booth, Michigan State University; CaRCC Researcher-Facing Track co-Coordinator

Patrick Clemins, University of Vermont; CaRCC Emerging Centers Track co-Coordinator

Bob Freeman, Harvard Business School (retired); CaRCC People Network co-Coordinator, Logistics, Metrics Committee

Forough Ghahramani, NJ Edge; CaRCC Capabilities Model Focused Tools Committee co-Chair

Brian Haymore, University of Utah; CaRCC Systems-Facing Track co-Coordinator, RCD Student Workforce Interest Group co-Chair

Scotty Strachan, Nevada System of Higher Education; CaRCC EPSCoR CI Interest Group co-Chair

Stefan Robila, Montclair State University; CaRCC RCD Professionalization Working Group

Introduction

The “What We’re Facing as RCD Professionals” workshop leveraged a proven CaRCC crowdsourcing approach to generate ideas for potential future community-driven work by leveraging in-person brainstorming to answer several key questions for the RCD profession: *What are we still facing as RCD professionals? What additional resources or community discussions would be supportive in light of existing resources?*

To this end, the workshop began with a crowd-sourced list of preferred resources for RCD professional development, identified and evaluated for specific relevance within breakout groups aligning around different career perspectives: Pre-Career and New Professionals, Established/Experienced Professionals, Organization/Leader Perspectives, Community- and Collaboration-Builders. With this resource review having generated awareness and discussion around existing resources, the second half of the day leveraged the same perspective alignment to discuss remaining professional challenges, and to generate new ideas for community-developed resources and discussion topics. This themed brainstorming was repeated with participants moving into [Facing](#)-specific breakout groups in a final brainstorming session, to inform potential facing-specific work and discussion topics for [People Network Tracks](#), and beyond. Results from each breakout group were presented to the entire room and later shared with attendees of the “Data” workshop, inviting interested individuals to sign up for potential downstream work.

Key takeaways

A number of themes and resource development priorities emerged from within the varying perspectives throughout the day. Multiple groups indicated ideas that could be achieved via a catalog of existing resources that could leverage the list generated during the workshop. Both the *Organization/Leader* and *Strategy-and-Policy-Facing* groups identified opportunities to share templates and stories for RCD strategic initiatives, including initial and ongoing RCD strategic plans and governance structures, as well as a need for peer support (also echoed by the other career stages) and consulting. Opportunities for coordination between existing CaRCC efforts included discussion topics around software implementation on RCD systems (for software-and systems-facing collaboration) and people-centric user experiences. Strategic issues and

requirements for facilitation learning around artificial intelligence (AI) and machine learning (ML) use cases were also popular.

Future work and next steps

Items identified as opportunities for experience-sharing will be carried forward as discussion topics within the most appropriate People Network Tracks and CaRCC Interest Groups. Ideas indicating the potential for new professional and organizational development resources will inform the work of potential new and some existing efforts within and beyond CaRCC support, including opportunities for collaboration with other efforts across the broader RCD community. We plan for one of these resources to be a new online tool reflecting and growing the [list of resources generated during the first part of the workshop](#); the [Staff and Student Workforce Development Interest Group](#) determined next steps during their September 8th call. Additionally, an Interest Group on AI Facilitation and a Working Group on developing AI Facilitation Materials are both being launched in late September.

Feedback on the workshop

A post-workshop survey was almost entirely positive.. Attendees indicated that they enjoyed the workshop approach to discussions - not only because they learned about resources they had not heard of and were motivated to get involved with [next steps](#) – but also because they enjoyed the opportunity to network meaningfully with other attendees via the breakout groups, with multiple opportunities to shuffle. One attendee commented that they liked the “great networking opportunities and great discussions” and went on to say “It was helpful and useful to see how many of the gaps/challenges I/we are seeing at our institution are the same for so many others and to discuss possible solutions.” Constructive feedback was individualized and largely focused on workshop logistics, with no clear patterns emerging.

Workshop 2: Data Governance and Working with Data

Workshop facilitators

Betsy Hillery, Purdue University, CaRCC Student co-Chair, Mentoring Working Group co-Chair

Patrick Schmitz, Semper Cogito, CaRCC

Debra McCaffrey, Arizona State University, CaRCC Data Facing Track co-Chair

Venice Bayrd, Montana State University, CaRCC Data Facing Track co-Chair

Susan Ivey, North Carolina State University

Jason Simms, Swarthmore College

Clark Gaylord, The George Washington University

Introduction

In the evolving landscape of research data management, ensuring that data is managed effectively and remains secure, accessible, and FAIR (Findable, Accessible, Interoperable, and Reusable) is paramount. The "Data Governance and Working with Data" workshop brought together a diverse group of professionals from across the RCD spectrum to tackle these pressing challenges.

This workshop provided a unique opportunity for participants to explore the current state of the art in research data lifecycle management. Our goal was to collaboratively outline a future resource that could support secure and effective research data lifecycles across institutions of varying sizes and data maturity levels.

To facilitate meaningful dialogue and ensure that all perspectives were considered, we organized the workshop into breakout groups categorized by emerging centers and mature centers in terms of data management and governance. This setup allowed for focused discussions that addressed specific needs and experiences, fostering a collaborative environment where both large and small institutions could share insights and solutions.

As we move forward, the insights gained from this workshop will play a crucial role in shaping strategies and resources aimed at enhancing data governance practices and supporting researchers in managing their data more effectively.

Key takeaways

Community-Oriented Resources

- **Consistency is Key:** A structured and consistent process for data governance is essential.
- **Recurring Questions:** Frequently asked questions highlight the need for more accessible, centralized resources.
- **Delphi Survey as a Tool?:** Exploring the use of Delphi surveys may help in gathering community insights.
- **Challenges with Public Data:** Securing public data remains a significant challenge. Although there is a demand for this data, participation in surveys to collect it is often low.
- **Preparation is Crucial:** Gathering preliminary data before workshops could enhance discussions.
- **RDAP Knowledge Gap:** Many participants are unfamiliar with RDAP (Research Data Alliance Projects), indicating a need for broader awareness.

Technical Tools

- **Institutional Culture Influence:** Institutional culture heavily influences how data governance and technical tools are adopted.

- **Institutional Demographics:** Understanding the institutional context, or "demographics," is vital for selecting appropriate tools and processes.
- **Middleware and Metadata:** The choice of middleware for Research Data Management (RDM) and effective metadata collection are crucial for success.
- **Storage Services:** Identifying and managing storage services is a common challenge.
- **Tool Proliferation:** Many institutions use multiple tools for the same task, leading to inefficiencies. Governing the number of tools is essential.
- **Institutional Tools Wishlist:** There is a demand for a more curated and focused set of institutional tools.

People, Processes, and Governance

- **Institutional Buy-In:** Achieving institutional buy-in is a common challenge across organizations.
 - **Example Narratives:** Sharing successful examples can help in gaining support.
 - **Metrics and Algorithms:** Utilizing metrics and algorithms can aid in storytelling and demonstrating the value of data governance.
 - **Communication Plans:** Effective communication strategies, especially regarding data findability, are crucial for success.

Future work and next steps

As we look to the future, the insights and discussions from the "Data Governance and Working with Data" workshop at Nexus Day provide a strong foundation for our next steps. Our focus will be on developing resources and strategies across three key areas: community-oriented resources, technical tools, and people, processes, and governance. Here's how we plan to move forward:

Community-Oriented Resources

To support the broader research data community, we will:

- **Investigate and Develop Artifacts:** We aim to create practical tools and documents that can help establish consistent data management processes. This includes but is not limited to:
 - **Consistent Processes:** Establishing standardized methodologies for data governance and management that can be widely adopted.
 - **Usage of Delphi Survey:** Employing this technique to gather expert consensus and refine best practices in data governance.
 - **Making Public Datasets Discoverable:** Enhancing visibility and accessibility of public datasets to improve their usability.
 - **Community Engagement:** Collaborating with organizations like RDAP (Research Data Access and Preservation) to align our efforts with ongoing community initiatives and resources.

Technical Tools

In the realm of technical solutions, our efforts will focus on:

- **Investigating and Developing Artifacts:** To improve technical infrastructure and support effective data management, we will:
 - **Middleware for RDM (Research Data Management):** Exploring and developing middleware solutions that facilitate metadata collection and management.
 - **Storage Strategies:** Examining and recommending storage solutions that meet the diverse needs across the RCD landscape.
 - **Governing Tools:** Evaluating and enhancing existing tools to support data governance and management practices.

People, Processes, and Governance

To address the human and organizational aspects of data governance, we will:

- **Investigate and Develop Artifacts:** We will create resources that address critical aspects of institutional data governance, including:
 - **Institutional Buy-In:** Crafting example narratives, metric algorithms, and "data stories" to demonstrate the value of data governance and secure institutional support.
 - **Data Classifications:** Defining and standardizing data classifications to improve data handling and protection.
 - **Governance Frameworks:** Developing clear models for what effective data governance should look like, including policy formulation and prioritization.
 - **Best Practices:** Compiling and disseminating best practices to guide institutions in their data governance efforts.

Feedback on the workshop

Survey feedback was overwhelmingly positive, with attendees citing the ability to learn from their peers as the most valuable part of the workshop. One attendee listed the best thing about the workshop as being “the energy and willingness from the community to share and attack issues we are facing,” while another described the workshop as “well organized, collaborative, excellent prompts and opportunities to engage.” Constructive feedback was minimal, with the most common suggestion being to include even more time and ways for attendees to mingle and get to know one another.

RCD Nexus is supported by the National Science Foundation under grant OAC-2100003. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.