Facilitating RCD Community Success
September 19, 2024

Dana Brunson, Internet2, Research Engagement
John Hicks, Internet2, Research Engagement
Timothy Middelkoop, Internet2, Research Engagement
The Internet2 Research Engagement Team
Why is supporting research important to Internet2?

It’s part of our mission.

- Research is the “R” in “R&E”. It’s part of our DNA.
- We have a strategic initiative to “Improve focus on research”.

It’s valuable to our community.

- It supports the research enterprise at our community institutions and organizations.
- Research is a core part of Internet2 value proposition.

It drives the demand for Internet2 services.

- Research drives the need for innovation in technologies that Internet2 develops.
- Research drives adoption of technologies that Internet2 provides.
- Research drives development of these technologies.
The Research Computing & Data Landscape is complex
The role of RCD+ Professionals:

- For research that relies on technology, support from a highly skilled workforce is now required.
- **This workforce is an emerging profession:** The Research Computing and Data (RCD+) professional community.

Institutions cannot realize their full research potential without a strong RCD workforce.

The Challenge for Researchers:

- Research and scholarship is now reliant on computing and data infrastructure, services, and support across most disciplines.
- Research is more frequently **conducted by widely dispersed communities** that are sharing **larger, more complex data sets** and using **more sophisticated tools**.
- A rapidly increasing level of knowledge and skill is required to make effective use of research computing and data infrastructure and services.
- Many researchers simply cannot keep up.
What we know today

Researchers struggle to take advantage of ever-evolving RCD technology
- They have no time to become technology experts
- They rely heavily on RCD professionals for tech services, support and facilitation.

Institutions need support and guidance
- To boot-strap, grow, and sustain their programs
- To explain to leadership how RCD is critical to research
- To recruit, retain, and develop RCD professional staff

RCD professionals are clamoring for a shared voice
- To advocate for the profession
- For broader understanding of their roles
- For training, mentoring, and professional development resources.
Research Engagement at Internet2

**Internet2 Facilitation for RCD/Research Community**
- Support RCD/Research to effectively leverage Internet2 offerings.
- Support and collaborate with Internet2 colleagues in on key projects that address current and future challenges.
- Be a conduit for information flow between RCD and Internet2.

**RCD Community Facilitation** – Engaging, coordinating, and leading the community to address common challenges (via NSF funded RCD Nexus, CaRCC, MS-CC, ERN, etc)
- Best practices clearinghouse
- Workforce development focus
- Institutional RCD Capabilities assessment, impact measurement, strategic communication
- International RCD cooperation/collaboration

**Advocate for RCD/Research Communities** between institutional administrators, national funding agencies, professional organizations, and the private sector
Research Engagement at Internet2

Internet2 Facilitation for RCD/Research Community
- Support RCD/Research to effectively leverage Internet2 offerings.
- Support and collaborate with Internet2 colleagues in on key projects that address current and future challenges.
- Be a conduit for information flow between RCD and Internet2.

RCD Community Facilitation – Engaging, coordinating, and leading the community to address common challenges (via NSF funded RCD Nexus, CaRCC, MS-CC, ERN, etc)
- Best practices clearinghouse
- Workforce development focus
- Institutional RCD Capabilities assessment, impact measurement, strategic communication
- International RCD cooperation/collaboration

Advocate for RCD/Research Communities between institutional administrators, national funding agencies, professional organizations, and the private sector
Community Needs Drive Our Priorities

Sources of Community Input:
- 2017 CaRCC Stakeholder Mapping Survey
- 2020 NSF CI Workforce Workshop
- 2021 RCD Workforce Survey
- 2021 RCD Career Arcs Survey
- 2021 Missing Millions Report
- 2021 MS-CC Stakeholder Mapping
- 2021 RCD Professionals Stakeholder Pulse Survey
- 2022 NSF ACCI Workforce Working Group Report
- 2022 CaRCC Strategy retreat
- CaRCC & RCD Nexus Advisory Committees
- RCD Nexus Day @PEARC22
- 2023 CaRCC Strategy Retreat
- RCD Nexus Day @PEARC23

Key Findings: Institutions and RCD+ professionals are clamoring for curated resources and a shared voice to:
- Help recruit, retain, and develop RCD+ professional staff
- Advocate for the profession and broaden understanding of their roles
- Provide training, mentoring, and professional development resources
- Better understand the RCD+ landscape and their own institutional strengths
- Increase diversity, equity, and inclusion
Who is CaRCC?
Many contributors - https://carcc.org/our-thanks

Jay Alameda, University of Illinois, Urbana-Champaign
Troy Baer, Ohio Supercomputer Center
Gladys Karina Andino Bautista, Purdue University
Venice Bayrd, Montana State University
John Blaas, National Center for Atmospheric Research
Justin Booth, Michigan State University
Tara Borland, University of Alaska
Jim Bottum, Clemson University
Geoff Boushey, UC San Francisco
Carrie Brown, Penn State University
Dana Brunson, Internet2
Katia Bulekova, Boston University
Cyd Burrows-Schilling, UC San Diego
Zach Byerly, Louisiana State University
Laura Carriere, NASA
Sarvani Chadalapaka, UC Merced
Dhruva Chakravorty, Texas A&M
Cathy Chaplin, Internet2
Shafaq Chaudhry, University of Central Florida
Thomas Cheatham, University of Utah
Dirk Colby, Michigan State University
Galen Collier, Rutgers University
Jane Combs, University of Cincinnati
Martin Cuma, University of Utah
Joel Cutcher-Gershenfeld, Brandeis University
Deborah Dent, Jackson State University
F. Alex Feltus, Clemson University
Bob Freeman, Harvard University
Kimberley Grash, University of Chicago
Fred Harris, University of Nevada Reno
Brian Haymore, University of Utah
John Hicks, Internet2
Betsy Hillery, Purdue University
Gwen Jacobs, University of Hawaii
Douglas Jennewein, Arizona State University
Kerk Kee, Texas Tech
Rich Knepper, Cornell University
Shelley Knuth, University of Colorado Boulder
Christina Koch, University of Wisconsin - Madison
Amy Koshoffer, University of Cincinnati
Gail Kroitz, Internet2
Elizabeth Leake, STEM TREK
Jim Leous, Penn State University
Julie Ma, MGHPCC
Tobin Magle, Northwestern University
Christina Maimone, Northwestern University
Sara Mannheimer, Montana State University
Ruth Marinshaw, Stanford University
Deb McCaffrey, University of Michigan
Lauren Michael, University of Wisconsin - Madison
Bill Michener, University of New Mexico
Timothy Middelkoop, Internet2
Claire Mizumoto, UC San Diego
Amy Neeser, UC Berkeley
Janna Nugent, Northwestern University
Barr von Oehsen, Rutgers University
Arman Pazouki, Northwestern University
Sai Pinnepalli, Rutgers University
Sam Porter, University of Maryland
Wirawan Purwanto, Old Dominion University
Annelie Rugg, UC Los Angeles
Patrick Schmitz, Semper Cogito
Erin Shaw, University of Southern California
Jerry Sheehan, San Diego State University
Andy Sherman, Yale University
Kurt Showmaker, Jacksons Labs
Alan Silver, University of Wisconsin - Madison
PRESTON Smith, Purdue University
Tracy Smith, University of Illinois
Dan Stanzione, Texas Advanced Computing Center
Ashley Staufer, Penn State University
Scotty Strachan, University of Nevada Reno
John Towns, NCSA
Pips Veazey, University of Maine
Jason Wells, Bentley University
Karen Wetzel, EDUCAUSE
Jim Wilgenbusch, University of Minnesota
Scott Yockel, Harvard University
Nuyun Zhang, Georgia Institute of Technology
Anita Orendt, University of Utah
... from many institutions
Campus Research Computing Consortium

An organization advancing the frontiers of research by improving the effectiveness of RCD professionals, including their career development and visibility, and their ability to deliver services & resources for researchers.

https://carcc.org/
CaRCC's People Network = Inclusive Community

A year-round virtual conference to share/leverage expertise in various facings (Tracks)

**Data-Facing Track:**
Discussions of analysis, curation, preservation, compliance, movement

**Researcher-Facing Track:**
Topics include consulting, bridging, facilitation, outreach, training

**Software-Facing Track -> US-RSE:**
Development or contribution to research software, scripts, workflows

**Emerging-Centers Track:**
Issues of interest to newer or emerging RCD service leaders and programs

**Systems-Facing Track:**
Discussions around hardware, networks, security, compliance, data centers

**Strategy/Policy-Facing Track:**
Topics such as institutional alignment, culture, funding, policy, community
CaRCC is one of the largest and fastest growing organizations of RCD professionals.
The People Network: Highlights 2022-2023

Externally

- ~1500 Track Call attendees across 28 Track & joint calls
- >6500 YouTube Recordings views across all Track Calls in the past year
- Launch of the Strategy & Policy-Facing Track, and tie-in with Emerging Centers Track
- First Accessibility Month call series: Community discussion, 3 speakers, and feedback/debrief
- Broadening inclusion for Community Calls announcements & accessibility for Track materials
- Continuing to develop and build new collaborations with peer communities.

Internally

- 28 Steering Committee members across all five Tracks
  - Helpful in developing & coordinating topics & presenters, in addition to bringing in new ideas
  - Mentoring potential future Track Chairs.
- More discussion and coordination for cross-track, themed, & plenary calls
- More joint calls to continue to build bridges / break down barriers amount groups & interest areas
Join the mailing list to hear about more opportunities in CaRCC and RCD Nexus:
RCD Nexus - Supporting the RCD Workforce to Advance Research

An NSF-supported CI Center of Excellence Pilot

May 2021 - April 2024 (Award 210003)

Dana Brunson, PI - Internet2 - Executive Director for Research Engagement
Scott Yockel, co-PI - Harvard - University Research Computing Officer
Claire Mizumoto, co-PI - UC San Diego - Director, Research IT Services
Thomas Cheatham, co-PI - University of Utah - Professor of Pharmacy & Director of RC/CHPC
Patrick Schmitz, co-PI - Semper Cogito
Daphne McCanse, Communications and Engagement Lead
RCD Nexus Pilot: Initial Priorities

• Conduct an RCD Workforce Staffing survey
• Advance the adoption of RCD Job Family Framework
• Develop a Career Arcs resource for RCD professionals
• Curate leading practices for Staff Professional Development
• Curate leading practices for Involving Students in the RCD Workforce
• Develop an Enhanced RCD Capabilities Model Assessment Tool and Community Dataset Portal
CaRCC Working Groups

Activities that connect Research Computing and Data (RCD) professionals & organizations to increase knowledge sharing & enable research innovation

RCD Professionalization
Develop and disseminate frameworks, models, and approaches that can be used by leaders in HR, IT, Research, etc. to support and elevate RCD roles.

RCD Career Arcs
Creates resources to explain RCD career options, help existing RCD professionals explore professional development and advancement, and aid in recruitment.

RCD Capabilities Model
Allows institutions to assess their RCD support, identify potential areas for improvement, and understand how the broader community views RCD support.

Engagement
Coordinate with working groups to increase visibility & participation in CaRCC, esp. with underserved & underrepresented persons and organizations.

https://carcc.org/working-groups/
Staff & Student Workforce Development

Leveraging a strong community, addressing current need, and looking to the future

Supporting RCD Professionals
Those doing the job now, and those aspiring to become an RCD professional.

Supporting Student Programs
Growing the workforce, share successful student involvement practices

Supporting Institutions
Develop their programs in support of RCD staff and student workforces

Monthly Virtual Sessions
Alternates between student and staff focus areas. All are welcome to participate.

https://carcc.org/working-groups/
Have you ever wanted to know…?

• How well is my institution supporting computationally- and data-intensive research and how can we get a comprehensive view of our support?
• What is my institution not thinking about or missing that the community has identified as significant?
• How can my institution (and my group) identify potential areas for improvement?
Why is there a RCD Capabilities Model!

• To assess support for computationally- and data-intensive research with a structured model
• To identify gaps and potential areas for improvement
• To understand how the broader community views Research Computing and Data support
• To gather baseline data about our community, to inform policy-makers, funders, etc.
## The RCD CM Assessment tool

<table>
<thead>
<tr>
<th>Area</th>
<th>Questions to consider for: Example University</th>
<th>Help</th>
<th>Deployment at Institution</th>
<th>Multi-Institutional Collaboration</th>
<th>Service Operating Level</th>
<th>Computed Value</th>
<th>Local Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Storage Infrastructure</strong></td>
<td>Do researchers have access to active data storage services (a.k.a., “scratch” storage, often a parallel filesystem) sufficient for HTC/HP’C?</td>
<td>Deployment/support institution-wide - 5</td>
<td>No existing multi-institutional collaboration - 1</td>
<td>Basic/Economy - 4</td>
<td>79%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Do researchers have access to sufficient storage to support researchers’ data intensive computing needs?</td>
<td>Deployment/support for parts of the institution - 4</td>
<td>No existing multi-institutional collaboration - 1</td>
<td>Basic/Economy - 4</td>
<td>68%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Do researchers have access to mechanisms for isolated and secure support for storage of sensitive/secure data?</td>
<td>Deployment/support institution-wide - 5</td>
<td>No existing multi-institutional collaboration - 1</td>
<td>Basic/Economy - 4</td>
<td>79%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Do researchers have access to policies and technologies that facilitate management and wide access to data? E.g.: i. Automated tiering and data migration. ii. Security/compliance management support for sensitive/controlled data that require special access/export controls</td>
<td>Tracking potential use - 2</td>
<td>Exploring multi-institutional collaboration - 2</td>
<td>Lights on Only - 3</td>
<td>36%</td>
<td>Med Prio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Do researchers have access to data archival and preservation services (e.g. tape, cloud)?</td>
<td>Deployment/support for parts of the institution - 4</td>
<td>Exploring multi-institutional collaboration - 2</td>
<td>Substantial Risk of Failure(s) - 2</td>
<td>48%</td>
<td>Mod Prio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Do researchers have access to a place to store final research data to address institutional policy and/or funding agency requirements?</td>
<td>Deployment/support for parts of the institution - 4</td>
<td>Sustaining multi-institutional collaboration - 4</td>
<td>Lights on Only - 3</td>
<td>66%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Do researchers have access to support for collaborative data grids and/or data repositories (national/local-commercial: e.g. OSN, iRODS, etc)?</td>
<td>Tracking potential use - 2</td>
<td>Exploring multi-institutional collaboration - 2</td>
<td>No existing service or support - 1</td>
<td>12%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **Network and Data Movement Infrastructure** | 71% |
| Do researchers have access to a high performance network that supports research within campus? | Deployment/support institution-wide - 5                      | Sustaining multi-institutional collaboration - 4                     | Basic/Economy - 4            | 92%          |                |
| Do researchers have access to a Science DMZ (a means to securely enable high performance inter-campus data flows that bypass campus firewalls)? | Deployment/support institution-wide - 5                      | Sustaining multi-institutional collaboration - 4                     | Basic/Economy - 4            | 92%          |                |
Who is using it?

- 155 Unique institutions, as of Feb 2023

**Pie Chart 1:**
- Private: 20.7%
- Public: 79.3%
- EPSCoR: 27.1%
- Non-EPSCoR: 72.9%

**Pie Chart 2:**
- Not a minority-serving institution: 120
- Hispanic-Serving Institutions (HSI): 3
- Minority-Serving Institution (MSI): 12
- Historically or Predominantly Black Colleges and Universities (HBCU/PBI): 20

Map of the United States with institutions marked.
Who is using it?

Intended Use

- Benchmarking of current service offerings: 82%
- Strategic planning: 87%
- Understanding common practices: 62%
- Other: 7%

160 Unique institutions, as of Dec 2022
RCD Capabilities Model *Focused Tools*

- Define a *new* set of tools for first-step and more focused assessments at smaller and emerging programs
- Developing in engagement with community partners (focus groups, workshops, interviews, etc.)
  - EPSCoR, GPN, NJEdge, I-Light, SOX, and other communities
  - Smaller institutions (non R1/R2) with significant research activity
- Current thinking: two new resources
  - Conversation template/script as first introduction to ideas
  - “Chart your own path” model for quicker/focused assessments
Discussion

Research Engagement at Internet2
dbrunson@internet2.edu
jhicks@internet2.edu
tmiddelkoop@internet2.edu
Acknowledgements

Portions of this material are based upon work supported by the National Science Foundation under the following grants:

Grant No. OAC-1440783, “A Model for Advanced Cyberinfrastructure Research and Education Facilitators”

Grant No. OAC-1548562, “XSEDE 2.0: Integrating, Enabling and Enhancing National Cyberinfrastructure with Expanding Community Involvement”

Grant No. OAC-1620695, “RCN: Advancing Research and Education Through a National Network of Campus Cyberinfrastructure, Infrastructures – The CaRC Consortium”

Grant No. OAC-2036534, “Building the research innovation workforce: a workshop to identify new insights and directions to advance the research computing community.”

Grant No. OIA-2033483, 2033514, & 2033519, “Building Research Cyberinfrastructure in EPSCoR Jurisdictions: Assessment, Planning and Partnerships”

Grant No. OAC-2100003, “CI CoE: Demo Pilot: Advancing Research Computing and Data: Strategic Tools, Practices, and Professional Development” (RCD Nexus)