



## Annual Report for May 15, 2021-April 30, 2022

### RCD Nexus

# NSF CI CoE: Demo Pilot: Advancing Research Computing and Data: Strategic Tools, Practices, and Professional Development

[NSF Grant OAC-2100003](#)

May 15, 2021-April 30, 2023

For Public Distribution

RCD Nexus Team

Dana Brunson (Internet2)	Overall program leadership and guidance
Patrick Schmitz (Semper Cogito)	Enhancing/supporting the RCD Capabilities Model, portal development oversight
Claire Mizumoto (UC San Diego)	Professional development of staff
Scott Yockel (Harvard)	RCD professionalization (job families, staffing survey)
Thomas Cheatham (U of Utah)	Involving students in RCD/workforce development
Daphne McCanse (CaRCC)	Communications coordination and facilitation

## About the RCD Nexus

The RCD Nexus is funded by NSF's Office of Advanced Cyberinfrastructure as an NSF Cyberinfrastructure Center of Excellence (CI CoE) demo pilot. It is a Research Computing and Data Resource and Career Center that creates tools, practices, and professional development resources to support individuals and institutions.

The main areas of emphasis for the RCD Nexus include:

- A more robust and sustainable implementation of the RCD Capabilities Model and a new Community Dataset portal
- Conducting an RCD Professional Staffing survey and sharing the results
- Advancing the adoption of the HR Job Family Framework
- Developing a Career Arcs resource for RCD professionals
- Curating leading practices for staff professional development, and for involving students in RCD/workforce development

For information about the RCD Nexus, please visit the project website: <http://rcd-nexus.org>.

## About This Report

This report represents the first project year of NSF grant OAC-210003 from May 15, 2021 to April 30, 2022.

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.

This work is made available under the terms of the Creative Commons Attribution 4.0 Unported License. Please visit the following URL for details:

<https://creativecommons.org/licenses/by/4.0/>.

Please cite this report as:

Brunson, Dana; Schmitz, Patrick; Mizumoto, Claire; Yockel, Scott; Cheatham, Thomas and McCause, Daphne. (2022) RCD Nexus Annual Report for Project Year 1. April 2022. Zenodo. <https://doi.org/10.5281/zenodo.6502662>

For updates to this report and other reports from the RCD Nexus, please visit <http://rcd-nexus.org>.

## Executive Summary

The RCD Nexus is a Research Computing and Data Resource and Career Center that creates tools, practices, and professional development resources to support individuals and institutions. As the need for RCD professionals continues to grow rapidly, so does the need for training, reference data, HR tools, and professional connection to support these professionals. However, the role of RCD professionals is poorly understood compared with other IT professionals and the field is only just starting to be recognized as its own unique profession. RCD Nexus strives to be a central resource for current data, leading practices, professional development, and connection within the RCD community. Significant accomplishments during the reporting period are summarized below.

### **RCD Workforce Survey**

From June - September 2021 the RCD Nexus team conducted a survey of the broad RCD workforce at institutions across the US. Survey questions focused on:

- Educational background, individual characteristics, and area of RCD focus
- Level of satisfaction with current position
- How included they feel in RCD field
- Job responsibilities and compensation

There were nearly 600 valid survey responses and early observations include:

- There are differences in the feeling of belonging between historically dominant identities and respondents from historically excluded groups.
- A large portion of respondents feel their work is valued by those they work with, but relatively few felt it was valued by their institution.
- Overall, survey results support prior qualitative assessments of the challenges to recruit and retain talent, develop career pathways, and grow the profession.

### **2021 Capabilities Model Report**

In summer of 2021, the project team completed the 2021 RCD Capabilities Model community dataset campaign, gathered the data into a new analytic tool set, and prepared the community data report. A shortened version of the 2020 analysis was published (and won an award for Best Paper) at the PEARC'21 conference in July, 2021.

## **IEEE Computer and Science Engineering Article**

In September 2021, the RCD Nexus team published a journal article entitled “Advancing the Workforce That Supports Computationally and Data Intensive Research” in the IEEE journal Computer Science and Engineering. <https://ieeexplore.ieee.org/document/9492830>.

## **Career Arcs Survey**

The project team completed a survey about the career paths or “arcs” of RCD professionals from November 2021 - January 2022. Survey questions focused on:

- General characteristics of RCD professionals and their entry into the field
- Factors related to RCD career entry, advancement, and satisfaction
- Factors related to changing jobs and/or leaving RCD for other domains

The survey received 225 responses (159 complete) and early observations include:

- Top ranking answers to “What advancement means” were recognition, salary, and influence on organizational strategy
- Top factors motivating job changes were more meaningful contribution, professional development, and joining a more innovative organization
- Women ranked recognition and impact as much more important than salary, while men ranked them more equally.

In the coming year, RCD Nexus will build on the solid foundation it has built in Year One of the award period and will focus on expanding the RCD Nexus website to provide more resources for RCD professionals and HR directors. The project team will also be participating in and hosting numerous educational and networking events for RCD professionals, which are detailed in the body of this report.

## Table of Contents

<b>About RCD Nexus</b>	<b>1</b>
<b>About this Report</b>	<b>1</b>
<b>Executive Summary</b>	<b>2</b>
<b>1 Program Administration</b>	<b>5</b>
<b>2 Developing Career Arcs</b>	<b>5</b>
<b>3 RCD Capabilities Model</b>	<b>7</b>
<b>4 RCD Staff Resources</b>	<b>10</b>
<b>5 Student Workforce Development</b>	<b>12</b>
<b>6 Job Family Matrix Adoption</b>	<b>13</b>
<b>7 RCD Workforce Survey</b>	<b>15</b>
<b>8 Looking Forward</b>	<b>16</b>
<b>9 Publications and Presentations</b>	<b>15</b>
<b>10 Committees and Working Groups</b>	<b>19</b>

# 1 Program Administration

RCD Nexus consists of a set of related but distinct projects or workstreams, each managed by one of the Principal Investigators (PIs). In order to keep those projects in sync with one another, and to enable identification of cross-project risks and opportunities, we conducted a series of program administration activities during the reporting period, including:

- Established an RCD Nexus Advisory Committee and held a strategy and planning meeting in November 2021
- Planned and executed regular Advisory Committee and leadership team meetings
- Managed project and engagement planning activities which reflect plans and progress against milestones
- Liaised with key collaborators and other community organization and project leads in order to establish and solidify partnerships
- Managed sub-award contracts and oversaw spending
- Established program templates, policies, and procedures
- Posted presentations and other resources to the website and Zonodo for community use. (See section 9.)
- Disseminated program news and upcoming event information
- Shared program updates via presentations at community events
- Hired and onboarded communications coordinator, Daphne McCanse, who joined the project team in March, 2022
- Planned first RCD Nexus Day as a co-located event scheduled for July 10, 2022 preceding PEARC'22 with focused workshops on Student Workforce Programs; Staff Onboarding and workforce development practices; and defining an RCD Capabilities Model Essentials version for smaller institutions and emerging RCD programs

## 2 Developing Career Arcs

RCD Nexus is working to develop a model of Career Arcs for RCD Professionals in order to explain RCD career options and help existing RCD professionals pursue professional development and advancement.

Our key objectives are to address the following set of challenges, as identified in a series of workshops and meetings:

- Managers need better resources and the ability to offer defined career paths, so their staff are not forced to seek growth opportunities outside of RCD.
- Many RCD professionals say they see no clear path for career advancement; this is especially so for data-centric and researcher-facing roles.

- There is a national shortage of RCD (CI) personnel and high employee turnover, with significant training needed to orient new employees.
- There is no real talent pipeline seeking to join the RCD field. Students who are studying in relevant programs do not have an easy way to learn about or enter the field.
- RCD personnel experience precarious employment with few opportunities for advancement.

We began work to develop a framework for describing the roles and career arcs of RCD professionals and distributed content through engagement with the community. Specific activities included:

- Gathering example career-arc narratives from existing resources, such as Women in HPC (WHPC)
- Designing and running an online survey of factors that influence career choices among RCD professionals
- Conducting a series of in-depth interviews of RCD professionals to better understand their choices, motivations, and the factors that enable success

We gathered a series of example career arc resources and analyzed their approaches and components to inform our framework designs. We also designed and conducted the online survey of factors that influence RCD career decisions. This work was supported through considerable outreach to a range of RCD communities, and included a Birds-of-a-Feather session at the supercomputing conference (SC21) in November. The survey data was analyzed and written up as a paper that was accepted as a full paper to the PEARC'22 conference in summer 2022 (See section 9, Chaudhry, S. et al). We developed the interview protocol for the next step, and are currently working with Institutional Review Board (IRB) groups at several Universities to get approval, after which we will begin interviews. We currently have eight volunteers to be interviewed, and will be soliciting more this spring and summer.

Additionally, we formalized a plan to complete further analyses from the survey data and share the results on the RCD Nexus website. We will conduct the interviews, synthesize them into personas, and write up narratives for each persona. Each will be enriched with anonymized quotes, to make them personal without identifying real individuals. We will also develop a visualization of career arcs and pathways. We will either hold a workshop or a series of community outreach sessions to get feedback on the materials we develop and gather additional use-cases and data to supplement the personas and visualization. We will then produce a companion paper to the visualization with recommended practices for hiring managers for both recruitment as well as retention of RCD professional staff. Much of this activity will be conducted in collaboration with the Staff and Student workforce development activities.

### 3 RCD Capabilities Model

The goal of the RCD Capabilities Model (RCD CM) workstream is to develop a more robust and sustainable implementation of the RCD Capabilities Model and an associated data-exploration portal that eliminates the need to create a host of individual reports by hand.

We have two broad objectives for the RCD CM workstream:

- Implement a 2.0 version of the RCD Capabilities Model
- Expand consulting and support services to ensure a broad range of institutions can use the model and incorporate it into strategic planning work for their respective RCD programs

In the first year of activity we had to continue our support for the prototype implementation of the RCD CM assessment tool and the community dataset that we developed from contributed assessments. At the same time, we were designing and beginning implementation of the new version of the assessment tool and data exploration portal.

Highlights of our continuing work with the prototype implementation include:

- Published an analysis of the 2020 community dataset at PEARC'21 and awarded Best Paper in the Workforce Development, Training, Diversity, & Education track
- Conducted a workshop at PEARC'21 on strategic planning for RCD programs and published an associated report on what the community would need from a national resource to support strategic planning. (See section 9 Schmitz, P. et al, "Towards a National Best Practices Resource...")
- Completed the 2021 RCD CM community dataset campaign, including office hours support, individual outreach and support to institutions conducting assessments, and regular communications to the community to promote the effort
- Gathered the resulting data into a new analytic tool set based on Jupyter notebooks. (This was required when the increase in data overwhelmed the prototype analytics workflow in Google Sheets. It also ensures that the analysis workflow is repeatable and more easily understood by others.)
- Prepared a 2020/21 community data report that expanded the analysis we completed in 2020, considered changes in the patterns seen, and more carefully explored factors that contributed to specific patterns in the data. (See Section 9.)



- Leveraged the 2020/21 combined dataset to support the EPSCoR CI Workshop project<sup>1</sup> with data and analysis. Published the supporting analysis as a Tech Report

Our planned Implementation of Version 2.0 of the RCD Capabilities model includes:

- A more robust survey-based Assessment Tool
- A community database of assessment data
- A web application supporting data exploration and peer benchmarking
- A reporting functionality that allows users to generate summary reports that clearly communicate results to stakeholders

The core areas of work to re-implement the Assessment Tool on a new platform include:

- Updating the Introduction and Guide to usage for the 2.0 implementation and coordinating it with the help resources integrated into the survey itself
- Analyzing alternatives to improve the help functions that link the assessment tool to the Guide to usage, and designing a model for expanding upon the available support text
- Repurposing resources to improve the help text linked directly in the assessment tool. Because we are not porting the assessment tool to a very different platform (see below), the planned work to update the Introduction and Guide is much reduced.

We also completed work on assessment of new platform environment alternatives. We analyzed a number of survey platforms and reconsidered some of our original strategy. Several users have noted that the collaborative features of the current tool are important to their work and said a single user model on a survey platform would be much less useful. The conclusion of our platform analysis was that the survey tools we had considered lack necessary functionality to support key features of the assessment tool (leveraging input from the community on the functionality). As such, the current tool will be retained through 2022 and a very similar user experience will be implemented on the new data exploration platform.

Although we must add new tasks to the plan to implement the assessment tool on the new platform, we do not anticipate that this will involve significantly more resources than were planned for porting to a survey platform. This work has been prototyped on the new platform, but the user-experience for this has not yet been determined and will depend upon the user-centered design work we are undertaking.

### **Platform development**

We identified and hired a developer for the core platform work. Working with him, we have completed implementation of the core technology stack that supports the data exploration portal. We developed the data model for the RCD CM data, and loaded the data from the 2020

---

<sup>1</sup> NSF (OIA) Award 2033483, Award 2033519, and Award 2033514, Collaborative Research: “Building Research Cyberinfrastructure in EPSCoR Jurisdictions: Assessment, Planning and Partnerships”

and 2021 community datasets. We developed the low-level RESTful Data APIs that provide access to this data and developed the authentication and authorization framework to support federated login and role management for access to the data.

We encountered challenges identifying a candidate for the user-centered design domain expert. After many attempts, we believe we have identified a partnership that will allow us to proceed. The user experience and user interface for the portal has been delayed for this reason, but we should be resuming the work soon.

During the reporting period, we have:

- Identified a framework for the broader RCD Nexus web site that will host data exploration portals as well as resource collections, news, and other elements.
- Secured the domain names for the RCD Nexus resource and career center, and created a basic site. (Further work on this depends upon the user-centered design work we are undertaking.)
- Drafted elements of the information model, design preferences, and personas that will inform the user-centered design process for the RCD Nexus portal.
- Conducted a recruitment, identified, and onboarded a developer to build the RCD CM data exploration portal infrastructure. This infrastructure work will be generalized to support other data exploration applications in addition to the RCD CM.
- Analyzed platform alternatives, identified core stack elements and developed the backbone framework for the portal.
- Designed the data model and ported existing 2020/2021 data into a new data framework.
- Imported the assessment tool questions and structure, and ported the computational logic for capabilities coverage values to the new data framework.

As part of ongoing efforts to assess the usability of the RCD CM tools, we learned many smaller institutions and emerging programs were somewhat overwhelmed by the scope of the current RCD CM assessment tool. In response, we resolved to create an Essentials version of the model.

We have made progress on this in the following ways:

- Developed a set of principles that we believe should guide the design of the Essentials version
- Developed a set of straw proposals for how we might define the Essentials version
- Conducted statistical analyses of the community data to explore how smaller institutions assessed their programs in comparison to larger programs and institutions
- Recruited two cohorts of individuals at candidate institutions to join focus groups that review the current model and the principles and approaches to an Essentials version

Their feedback will inform the design of a workshop to be held as part of the RCD Nexus Day in coordination with PEARC'22.

### **Engagement with AIHEC**

We are using a small amount of our consulting resources in an engagement with the American Indian Higher Education Consortium (AIHEC, <http://www.aihec.org/>) in support of their Cyberinfrastructure Strategic Planning (CISP) Initiative. The AIHEC Cyberinfrastructure (CI) Team advances the STEM and education programs at the nation's 37 Tribal College and Universities (TCUs) by implementing a comprehensive CI capacity-building strategy focusing on the colleges' STEM faculty and CI support staff. With support from external partners and regional institutions, this comprehensive CI strategy focuses on CI training, planning, and community-building involving both STEM faculty and TCU IT organizations, providing the resources, technical assistance and national network to advance participating TCUs toward CI-readiness and CI-enabled STEM research and education programs. We are engaged on the AIHEC Cyberinfrastructure Strategic Planning (CISP) Initiative, "Developing a technology roadmap at the TCUs to support their academic, research, business, and public service missions."

### **Office hours**

As part of support for the RCD CM we have established regular office hours to provide support for institutions completing the Capabilities Model and for individuals needing more specific help. Discussions among individuals from participating institutions proved to be supportive and beneficial to both institutions and to the RCD CM Working Group with everything from approaches to campus participation to complete the survey, clarification on specific questions, and improvements for future versions of the Capabilities Model.

Office hours, from the start of the reporting period, have had 84 participants in 13 sessions. Additional email follow-up addressed directly to working group members and to the [capmodel-discuss@carcc.org](mailto:capmodel-discuss@carcc.org) mailing list have also resulted from office hours discussions.

Members of the RCD CM working group have also met individually with representatives from institutions requesting specific attention to their questions about the Model and the submission process, as well as a discussion about strategies for engagement and best practices to complete the Model.

The RCD CM Working Group will continue to host regular office hours, staffed by at least one individual from the working group at each session. They will continue to be provided remotely via Zoom and regularly advertised to the RCD CM participating institutions, general RCD community mailing lists, and the CarCC and RCD Nexus websites and in relevant blog posts and newsletters.

## 4 RCD Staff Resources

RCD Nexus is working to develop Staff Resources to support RCD professionals and institutions that need research computing and data practitioners to support research activities. The key objectives are to address the challenges identified in a series of discussions, call for position papers and program experiences documentation, and a workshop.

These include:

- Identifying leading practices and common elements among staff workforce development programs
- Propagating various models to support workforce development for staff RCD practitioners (as well as undergraduate and graduate students in collaboration with the RCD Nexus Student Resource working group)
- Employing a “train the trainers” model to build a community of trainers that, in turn, helps educate other RCD professionals on workforce development topics

The materials gathered will provide the basis for an analysis to identify gaps and needs for additional training programs that would facilitate the transition from other positions into RCD roles, and, for example, developing existing RCD staff into RCD leadership roles.

During the reporting period, we began forming a steering committee to guide progress, planned for a workshop, and identified partners in the staff RCD workforce development space. Our efforts have included gathering interested individuals to partner in discussions and efforts surrounding staff workforce issues. We assembled a group of interested individuals with diverse institutional representation and are using the CaRCC interest group framework to organize and set clear objectives for the group. We are establishing that framework and meeting to discuss and document the priorities and deliverables.

There were 38 participants in the introductory meeting and 19 initial participants for the RCD Staff Resources workstream interest group. The interested participants represent 17 different US and Canadian institutions. We will continue to grow this community of practice and develop partnerships with other RCD community activity groups who have related interest and group scope.

We have also been doing the necessary preparation to solicit and distribute completed position papers focusing on RCD staff workforce issues. Setting up and administering the RCD staff workforce interest group and gathering RCD community individuals has all been part of the groundwork necessary to foster and grow this community of interest.

Planning is in progress for a workshop to discuss leading practices, models, and success stories – as well as challenges, gaps and further needs – in staff RCD workforce development. The workshop is scheduled on RCD Nexus Day on July 10, 2022, preceding the PEARC’22 annual conference and co-located near the conference site for convenience to participate. It will be a half-day workshop, aligned closely with the RCD Student Workforce workstream efforts. The subsequent workshop report will follow.

## 5 Student Workforce Development

The focus of the Student RCD Workforce aspect of the project is to explore existing models for student RCD workforce development at both the undergraduate and graduate levels. Our main goal is to create a guide and other resources that share leading practices and models from organizations who are already doing this successfully.

As a community, we understand the increasing demand for trained RCD professionals. RCD Nexus envisions a model in which a student can get real world experience in using, developing, and optimizing RCD infrastructure and services during the course of obtaining their degree. Not only would this provide extra staffing support for the existing RCD staff or researchers, but the student learns and gains experience beyond their degree, making them more valuable to potential employers. Moreover, this can be a path towards internal recruitment and for universities to fill their own pipeline of required RCD professionals. Among the undergraduate RCD student workforce at Utah, many of our student employees have been selected for competitive internships at companies like Google and Facebook over the summer, and they bring back new knowledge to inform on innovations in the RCD infrastructure, operations, and services when the student returns in the fall.

Like the Staff Resources group, we are working to identify and share a set of models and best practices for Student Workforce development. We are also holding a series of discussions, call for position papers and program experiences documentation, and conducting a workshop.

The RCD Student Resources workstream effort included an initial interest meeting for individuals to self-identify as participants in discussions and efforts revolving around student workforce issues. We assembled an initial group of 14 interested individuals with diverse institutional representation and are using the CaRCC interest group framework to organize and set clear objectives for the group. We are establishing that framework and meeting to discuss and document the priorities and deliverables and fourteen initial participants for the RCD Student Resources workstream interest group.

An RCD Student Workforce Workstream workshop will be co-located at PEARC'22 and held on July 10, 2022 with a report to follow. We will identify leading practices and common elements among student workforce development programs as well as various different models for supporting workforce development for undergraduate and graduate students. We will describe models and leading practices on the associated websites and develop materials to inform the RCD Nexus Resource & Career Center.

The information we collect will be used to identify gaps and needs for additional training programs that would facilitate students transitioning into RCD roles. During the ongoing RCD student workforce interest group meetings and the July 10, 2022 workshop, the group will discuss and develop the organizational structure and categories for materials to be submitted and curated.

## 6 Job Family Matrix Adoption

The goal of the RCD Job Family Matrix effort is to broadly distribute the HR template and provide a support mechanism for institutions who want to roll out RCD specific positions. An essential component of establishing equity across positions is having a formal job classification system that corresponds to the complexity of the role and level of responsibility of positions, as well as job descriptions that are up-to-date with respect to the regular tasks being performed. Without these HR structures, it is difficult to reclassify positions over time and provide opportunities for promotion beyond a traditional management track. This creates a glass ceiling for RCD professionals and leads to difficulty in retaining the most talented staff members.

During the reporting period, we primarily focused on creating the RCD Workforce Survey and thus changed some of our initial timelines for other planned activities. Below we will summarize what was done and plans going forward.

We continued to give presentations to new audiences about the HR Job Family framework (See section 9) and advertised the upcoming workshops and direct campus engagements and began working with the Communications Lead to brainstorm about communication plans for this effort. Downloads of the Job Family have come from 173 unique emails across 117 different institutions. Ninety of these downloads are from the past year.

During the reporting period, we provided our first consulting engagement to Yale research computing and HR. This initial meeting also helped create a rough structure for how these sessions should be organized. This session helped Yale navigate implementation questions such as: (1) How does mobility and establishing a career path work in practice? (2) When would you

use the high-level lead and principal positions? (3) How do these positions overlap with digital humanities and libraries?

We also completed a smaller consultation with Arizona State University as they were building out a new team under the new Associate Director. After this engagement, they adopted the Facilitation job-series (researcher-facing) and this new team is Outreach and Engagement.

Through these engagements when HR is present it has been important to stress the following:

- Enterprise IT mostly provides common IT infrastructure that is a ubiquitous utility for all, such as email, calendaring, financial systems, ordering systems, HR systems, etc.
- RCD Professionals are tasked with co-learning problems and co-creating solutions with researchers, differentiating their work from Enterprise IT.
- We must think about the positions that are needed for an organization to be successful, not the person that is in the position.

The in-person 2022 CASC Annual Meeting was a successful advertisement for future engagements. Many Directors seemed to desire this type of engagement, especially those with upcoming program growth. We are currently in the process of creating a more formal outreach plan and a call for consulting engagements and creating a follow-up survey.

In the coming year, we will:

- Host a virtual workshop in September to share experiences from institutions that have already implemented the new HR Job Families, and help interested institutions learn how to implement it locally.
- Cultivate a cohort of subject matter experts (SMEs) in Fall that can provide support for campuses and provide a feedback mechanism for campuses to evaluate the SMEs support.
- Develop a guidebook for HR and hiring managers who are considering the adoption of RCD Job Family Matrix by December. This guidebook will provide a rich resource for campuses interested in making a change to how they treat RCD related positions. It will be updated regularly to incorporate feedback from the campus engagements.

## 7 RCD Workforce Survey

This section covers what the original proposal referred to as “Census of RCD Professionals Across Positions,” however, we have since adopted the name RCD Workforce Survey for this part of the project. Many groups across CaRCC and beyond have grappled with the notion of defining RCD professionals and their progression from a community of practice to an established and

recognized profession. This phase of work is slated to better understand comprehensively the state of RCD staffing across the US by designing, testing, and implementing an RCD Positions survey tool and using it to conduct a national survey. A group of RCD professionals who wanted to contribute to this effort was formed at the NSF Virtual Workshop on the Research Innovation Workforce for Cyberinfrastructure.<sup>2</sup> This study aims to provide the first large data set on the composition of the RCD workforce in terms of demographics, job responsibilities, job types, compensation, and perceptions of the RCD field. The study is limited to the US due to varying data privacy laws and international research requirements.

The main objectives of this part of the project include:

- Establish a working group
- Design, implement and test the RCD Positions survey tool
- Conduct a US national survey of RCD positions to capture specific information on RCD individuals, their current position, and their outlook on RCD as a profession
- Present and publish results on the survey widely with the community
- Anonymize survey data and integrate into the RCD Nexus Community Assessment Data Dashboard

Results from this survey will focus on a number of key areas including:

- Describe the current research computing and data professionals workforce
- Compare this workforce to related academic and technical workforces in terms of demographics and educational background
- Characterize the types of positions in the field and describe work done by different job positions
- Generate data for use by individuals, hiring managers, and human resources professionals on RCD salary ranges, job families, position titles, and responsibility sets

The working group reviewed a number of similar surveys from industry and academia and discussed what themes of questions could provide results that would be most beneficial to individuals and hiring managers. We distilled this into two surveys – one for individuals to take that would dive deeper into their personal background, job-specific information such as title, compensation and grade, the work they perform, and career perceptions – and another survey for managers that focused on institutional data about RCD units. We decided to focus on the first survey and postpone the second until a later time. Because we were collecting personal information that included demographic and compensation information, we needed to create an IRB study. We chose Christina Maimone to be the PI in the study and worked with the Northwestern IRB office to obtain approvals to conduct the study. The survey of individuals

---

<sup>2</sup> [https://www.nsf.gov/awardsearch/showAward?AWD\\_ID=2036534](https://www.nsf.gov/awardsearch/showAward?AWD_ID=2036534)



resulted in 12 questions on background, 12 on role basics, 12 on work being performed, and 10 on job satisfaction. We beta-tested the survey with 20 people from our own RCD community, which resulted in refining questions, providing more context to some questions, and shortening the survey. The survey was launched in June 2021 and closed in September 2021. It was promoted through many community mailing lists, Slack channels, and on social media outlets.

The RCD Workforce survey resulted in nearly 600 unique responses. This was less than our target of 1000, but still the largest responses to any survey of this type that we are aware of. The initial analysis of the survey data was written up and accepted as a full paper to the PEARC'22 conference (See section 9, Maimone, et al). In the coming year, we will continue the working group to review data for future publications. We will also decide whether to launch a follow-on individual study to try to capture wider participation in government labs and industry and also decide whether to launch the second survey that targets managers.

## 8 Looking Forward

In the coming year, the RCD Nexus team will work to build on the solid foundation started in Year One. Major goals in the coming year include launching and promoting the new RCD Nexus Portal as a central resource for the RCD Capabilities Model and Community Dataset, survey results, HR Job Family Framework, Career Arcs information, and other resources. We are working now on in-depth analysis of the Workforce and Career Arcs surveys and will release more detailed reports upon completion. We will be hosting RCD Nexus Day on July 10, 2022 in conjunction with the PEARC'22 conference. That workshop will include sessions on Student Workforce Programs & Staff Onboarding and Workforce Development Practices and Defining an RCD Capabilities Model *Essentials* version for smaller institutions and emerging RCD programs. We are also in the process of creating a comprehensive communication plan to inform the community about the resources that will be available through RCD Nexus.

## 9 Publications and Presentations

### Publications

Chaudhry, S., Pazouki, A., Schmitz, P., Hillery, E., & Kee, K. (2022, July). Understanding Factors that Influence Research Computing and Data Careers. Practice and Experience in Advanced Research Computing (PEARC '22). <https://doi.org/10.1145/3491418.3530292>

Maimone, C., Yockel, S., Alameda, J, Middelkoop, T., Neeser, A., Stauffer, A. (2021). CaRCC Research Computing and Data (RCD) Workforce 2021 Census Questions (Version 1.0 June 2021). Zenodo. <https://doi.org/10.5281/zenodo.5914431>

Maimone, C., Yockel, S., Middelkoop, T., Stauffer, A., & Reidy, C. (2022, July). Characterizing the US Research Computing and Data (RCD) Workforce. Practice and Experience in Advanced Research Computing (PEARC '22) <https://doi.org/10.1145/3491418.3530289>

Schmitz, P., Brunson, D., Mizumoto, C., Jennewein, D., & Strachan, S. (2022). Towards a National Best Practices Resource for Research Computing and Data Strategic Planning (RCDNexus-TR-2022.1) [Technical Report]. <https://doi.org/10.5281/zenodo.6394989>

Schmitz, P., Bayrd, V., Strachan, S., & Jacobs, G. (2022). A Baseline of EPSCoR Research CI Capabilities (RCDNexus-TR-2022.2). RCD Nexus. <https://doi.org/10.5281/zenodo.6395204>

Schmitz, P. (2021). Assessing the Landscape of Research Computing and Data Support: The 2020 RCD Capabilities Model Community Dataset. Practice and Experience in Advanced Research Computing, 1–8. <https://doi.org/10.1145/3437359.3465580>

Schmitz, P., Yockel, S., Mizumoto, C., Cheatham, T., & Brunson, D. (2021). Advancing the Workforce That Supports Computationally and Data Intensive Research. Computing in Science Engineering, 23(5), 19–27. <https://doi.org/10.1109/MCSE.2021.3098421>

Schmitz, P. (2022). 2021 RCD CM Community Data report (RCDNexus-TR-2022.3). <https://doi.org/10.5281/zenodo.6502962>

## **Presentations**

Brunson, D. “CaRCC and RCD Nexus Overview”, (2021, September) 2021 NSF Virtual CC\* PI Workshop

Brunson, D. “Overview of RCD Nexus, (2021, July 8) XSEDE Service Providers Forum

Brunson, D., Jennewein, D., Mizumoto, C., Schmitz, P., & Strachan, S. (2021, July 19). Building a Strategic Plan for your Research Computing and Data Program (Workshop). Practice and Experience in Advanced Research Computing (PEARC '21)

Chaudhry, S., Pazouki, A., Schmitz, P., Hillery, E., & Kee, K. (2022, April 14). Understanding Factors that Influence Research Computing and Data Careers. RMACC Women in HPC

Cheatham, T., CaRCC and RCD-Nexus updates. (2021, October) CaRCC and RCD Nexus updates. CASC Fall Meeting

Cheatham, T., Brunson, D., Marinshaw, R., Michael, L., Mizumoto, C., Schmitz, P., Yockel, S. (2021, July 21) Campus Research Computing Consortium (CaRCC) Town Hall. Practice and Experience in Advanced Research Computing (PEARC '21)

Cheatham, T., Marinshaw, R., Michael, L., Mizumoto, C., Schmitz, P., Yockel, S., Brunson, D., (2021, August 18) Campus Research Computing Consortium Addresses Challenges: Supporting Data-Driven Discovery on Campus. I2Online (Internet2 online event).

Cheatham, T. and Yockel, S., (2022 April) CaRCC and RCD Nexus updates. CASC Spring Meeting.

Dent, D., Arafune, L., Brunson, D., Broude Geva, S., Jacobs, G., Cutcher-Gershenfeld, J., Blatecky, A. (2022, April 29). Update from the Learning & Workforce Development Working Group. NSF Advisory Committee for Cyberinfrastructure (ACCI) Spring Meeting.

Hicks, J. and Schmitz, P., (2022, April 8) "RCD Capabilities Model Essentials version Focus Group". (In partnership with Edge.)

Hicks, J. and Schmitz, P., (2022, April 29) "RCD Capabilities Model Essentials version Focus Group". (In partnership with Great Plains Network.)

Mainmone, C., Stauffer, A., (2022, March) Data Professionals and Data Responsibilities in the Research Data and Computing Workforce. Research Data Access and Preservation (RDAP) Summit 2022: Envisioning and Inclusive Data Future.

McCaffrey, D., Schmitz, P., (2022) Presentation to Association of American Medical Colleges Group on Information Resources (AAMC GIR) about CaRCC and the Research Computing & Data Capabilities Model.

Pazouki, A., Freeman, B., Maimone, C., Mizumoto, C., Brunson, D., Jacobs, G., Hicks, J., Michael, L., Schmitz, P., Marinshaw, R., Yockel, S., Cheatham, T., Bayrd, V. (2022, January 11) Campus Research Computing Consortium (CaRCC) Parade. CaRCC People Network Plenary Session.

Schmitz, P., Pazouki, A., Chaudhry, S., (2021, November 16) "Building Bridges toward RCD Professionalization and Careers," Birds of a Feather session at SC21.

Strachan, S., Bayrd, V., Schmitz, P., Brunson, D., Veazey, P., & Jacobs, G. (2021, July 20). Assessing and Communicating Cyberinfrastructure Readiness at EPSCoR and Under-Resourced Institutions (Birds of a Feather session). Practice and Experience in Advanced Research Computing (PEARC '21).

Yockel S., Development of a Profession: A Research Computing and Data Community, (2021, June) University of Oklahoma Virtual Residency Program Workshop.

Yockel, S., RCD Workforce Survey Results Preview. (2021, October) CASC Fall Meeting.

Yockel, S., Grash, K., Mainmone, C., Middelkoop, T., Neeser, A., (2021, July 19) Quantifying the Research Computing and Data Professional Community, (Birds of a Feather session). Practice and Experience in Advanced Research Computing (PEARC '21).

## 10 Committees and Working Groups

### **Advisory Committee**

Linda Akli, Southeastern Universities Research Association

Ewa Deelman, University of Southern California

Joel Cutcher-Gershenfeld, Brandeis University

Sandra Gesing, University of Illinois, Chicago

John Goodhue, Massachusetts Green High Performance Computing Center

Leah Lang, EDUCAUSE

Jackie Milhans, Northwestern University

John Towns, University of Illinois

Von Welch, Indiana University

### **CaRCC Chairs Leadership**

Venice Bayrd, Montana State University

Dana Brunson, Internet2

Thomas Cheatham, University of Utah

Bob Freeman, Harvard Business School

Gwen Jacobs, University of Hawaii

Christina Maimone, Northwestern University

Ruth Marinshaw, Stanford University

Lauren Michael, University of Wisconsin-Madison

Claire Mizumoto, University of California San Diego

Arman Pazouki, Northwestern University

Patrick Schmitz, Semper Cogito

Scott Yockel, Harvard University

### **Capabilities Model Working Group**

Patrick Schmitz, Semper Cogito (co-chair)

Claire Mizumoto, University of California San Diego (co-chair)

Dana Brunson, Internet2 (co-chair)

Thomas Cheatham, University of Utah

Galen Collier, Rutgers University

John Hicks, Internet2

Douglas Jennewein, Arizona State University  
Scotty Strachan, Nevada

**RCD Professionalization Working Group**

Christina Maimone, Northwestern University (co-chair)  
Scott Yockel, Harvard University (co-chair)  
Jay Alameda, University of Illinois Urbana-Champaign  
Kimberly Grasch, University of Chicago  
Tim Middelkoop, Internet2  
Amy Neeser, University of California Berkeley  
Chris Reidy, University of Arizona  
Patrick Schmitz, Semper Cogito  
Ashley Stauffer, Penn State University

**RCD Career Arcs Working Group**

Patrick Schmitz, Semper Cogito (co-chair)  
Arman Pazouki, Northwestern University (co-chair)  
Geoff Boushey, University of California San Francisco  
Laura Carriere, NASA  
Sarvani Chadalapaka, University of California Merced  
Shafaq Chaudhry, University of Central Florida  
Elizabeth Hillery, Purdue University  
Kerk Kee, Texas Tech University  
Elizabeth Leake, STEM TREK  
Kerk Kee, Texas Tech  
Samuel Porter, University of Maryland  
Preston Smith, Purdue

**EPSCoR CI Workshop Working Group**

Gwen Jacobs, University Hawaii (co-chair)  
Venice Bayrd, Montana State University (co-chair)  
Dana Brunson, Internet2  
Zach Byerly, Louisiana State University  
Patrick Clemins, University of Vermont  
Deborah Dent, Jackson State University  
Fred Harris, University of Nevada, Reno  
Doug Jennewein, Arizona State University  
Patrick Schmitz, Semper Cogito  
Scotty Strachan, Nevada System of Higher Education, NevadaNet

Pips Veazey, University of Maine

**RCD Staff Workforce Development Interest Group**

Claire Mizumoto, University of California San Diego (co-chair)  
Thomas Cheatham, University of Utah (co-chair)  
Janae Baker, Rutgers University  
Caughlin Bohn, University of Nebraska  
Michael Colee, University of California Santa Barbara  
Scott Delinger, University of Alberta  
Tony Elam, University of Kentucky  
Bob Freeman, Harvard Business School  
Jeff Gardner, University of British Columbia  
Chuntida Harinritisuk, University of Texas Health Science Center at San Antonio  
Betsy Hillery, Purdue University  
Susan Ivey, North Carolina State University  
Alper Kinaci, Northwestern University  
AJ Lauer, National Center for Atmospheric Research (NCAR)  
Randy Melen, Stanford University  
Jackie Milhans, Northwestern University  
Dylan Perkins, University of Colorado Boulder  
Ashley Stauffer, Penn State  
Scotty Strachan, Nevada System of Higher Education, NevadaNet  
Laura Theademan, Purdue University  
Biru Zhou, McGill University

**RCD Student Workforce Development Interest Group**

Thomas Cheatham, University of Utah (co-chair)  
Claire Mizumoto, University of California San Diego (co-chair)  
Janae Baker, Rutgers University  
Michael Colee, University of California Santa Barbara  
Betsy Hillery, Purdue University  
Alper Kinaci, Northwestern University  
Elizabeth Kinney, University of British Columbia  
AJ Lauer, National Center for Atmospheric Research (NCAR)  
Dylan Perkins, University of Colorado Boulder  
David Reddy, University of South Carolina  
Chris Reidy, University of Arizona  
Scotty Strachan, Nevada System of Higher Ed, NevadaNet  
Jacob Fosso Tande, University of North Carolina Greensboro

Laura Theademan, Purdue University  
Ravi Vadapalli, University of Miami  
Biru Zhou, McGill University