

Collaborative Approach to Research Computing and Data Management at Mississippi State University

Mike Navicky, Director of High Performance Computing,
High Performance Computing Collaboratory

Lauren Geiger, Digital Archivist,
Mississippi State University Libraries

RCD (Research Computing and Data) Capabilities Model

- Identifies various approaches and factors for creating and maintaining an RCD program
- Potential uses for the capabilities model:
 - Use it as an in-depth SWOT (Strength, Weakness, Opportunities, Threats) chart
 - Comparison to peer institutions' capabilities
 - Collaboration potentials with other institutions

RCD (Research Computing and Data) Capabilities Model

- **Summary page of Capabilities Model results**

Facing Area <i>(click the "+" to the left of each to expand areas within)</i>	Computed Coverage	Domain Support	Domain Weighted Coverage
Researcher Facing Capabilities Coverage	55%	30%	28%
Data Facing Capabilities Coverage	41%	70%	32%
Software Facing Capabilities Coverage	43%	30%	22%
System Facing Capabilities Coverage	67%	30%	34%
Strategy and Policy Facing Capabilities Coverage	57%	30%	29%
Total Organizational Research Computing and Data Co	52%	38%	29%

RCD (Research Computing and Data) Capabilities Model

System Facing Capabilities Coverage	67%	30%	34%
Infrastructure Systems			
<i>Infrastructure Support</i>	68%		
<i>Compute Infrastructure</i>	79%		
<i>Storage Infrastructure</i>	63%		
<i>Network and Data Movement Infrastructure</i>	77%		
<i>Specialized Infrastructure</i>	72%		
<i>Infrastructure Software</i>	75%		
Systems Operations			
<i>Monitoring and Measurement</i>	64%		
<i>Change Mngmnt, version control, administration, and ticketing</i>	36%		
<i>Documentation</i>	56%		
<i>Planning</i>	61%		
Systems Security and Compliance			
<i>Best security practices for open environments</i>	83%		

RCD (Research Computing and Data) Capabilities Model

Data Facing Capabilities Coverage	41%	70%	32%
<i>Data Creation</i>	56%		
<i>Data Discovery and Collection</i>	39%		
<i>Data Analysis</i>	17%		
<i>Data Visualization</i>	11%		
<i>Research Data Curation, Storage, Backup, and Transfer</i>	58%		
<i>Research Data Policy Compliance</i>	62%		
<i>Data Security/Sensitive Data Support</i>	41%		

Research Team

- High Performance Computing
- MSU Libraries
- IT Services
- Office of Research Security and Compliance
- Researchers

Research Computing and Data Strengths

- Strengths of Library
 - Research data management
 - Description
 - Visualization
 - Service oriented work
 - Well connected on campus
- Strengths of HPC
 - Data storage
 - Computational capability and knowledge
 - Systems

Research Computing and Data Challenges

- Challenges

- Siloed departments
- Limited access to software, storage, and compute.
- Limited staff to teach/support researchers
- Recognizing the value of an RCD

- Approach

- Build relationships focused on RCD
- Discover researcher needs
 - Educate and collaborate with researchers on potential uses
- Build systems and support staff capabilities
- Communicate with senior leaders, faculty, and researchers about long-term value of RCD

CC* Area 7 Planning Grant

- Establish a plan to assess the specific needs and wants of our campus' stakeholders
- Main Goals
 - Better Understand Researcher Computing Needs
 - Create Cyberinfrastructure Plan
 - Broaden Research Computing Support

Future Objectives

- Better direct use of our resources from data gathered by the CC* Grant.
- Create a 'storefront' at the library to assist faculty and researchers leverage their research compute and data.
- Democratize research compute and data management on our campus.

Montana State University's Research Alliance: Intentional Library and University Partnerships for Research and Data Services

Doralyn Rossmann and Jason Clark

Montana State University Library
December 2023

Outline

- Assessment of Need (Doralyn)
- Partnership Formation (Doralyn)
- Projects and Instructional Collaborations (Jason)

Assessment of Need - Research Alliance

The Research Computing and Data Capabilities Model (RCD CM) was developed by a diverse group of institutions with a range of support models, in a collaboration among Internet2, CaRCC, and EDUCAUSE.

- <https://carcc.org/rcdcm/>
- [2020 RCD CM Community Data report](#)

[Benchmark Report: Montana State University RCD CM](#)

1.2. Structure of the RCD CM Community Dataset

The Community Dataset structure mimics that of the RCD Capabilities Model to facilitate benchmarking analyses by institutions that completed the RCD CM Assessment Tool. The Model recognizes different roles that staff and faculty fill in supporting Research Computing and Data with names that reflect who or what each role is facing (i.e., focused on), noting that a given individual may fill roles in multiple facings.

1. **Researcher Facing Roles.** Includes research computing and data staffing, outreach, and advanced support, as well as support in the management of the research lifecycle. Example roles include: Research IT User Support, Research Facilitator, CI engineer.
2. **Data Facing Roles.** Includes data creation; data discovery and collection; data analysis and visualization; research data curation, storage, backup, preservation, and transfer; and research data policy compliance. Example roles include: Research Data Management specialist, Data Librarian, Data Scientist
3. **Software Facing Roles.** Includes software package management, research software development, research software optimization or troubleshooting, workflow engineering, containers and cloud computing, securing access to software, and software associated with physical specimens. Example roles include: Research Software Engineer, Research Computing support.
4. **Systems Facing Roles.** Includes infrastructure systems, systems operations, and systems security and compliance. Example roles include: HPC systems engineer, Storage Engineer, Network specialist.
5. **Strategy- and Policy Facing Roles.** Includes institutional alignment, culture for research support, funding, and partnerships and engagement with external communities. Example roles include: Research IT leadership.



Research Alliance - www.montana.edu/research-alliance

Researcher Alliance

- Idea was initially floated 10+ years ago - library as a hub
- Identify appropriate partners through discussions with administrators leading those units.
- Secure funding and space allocation
- Consulting work with Rebecca Bryant from OCLC - research development partnership.

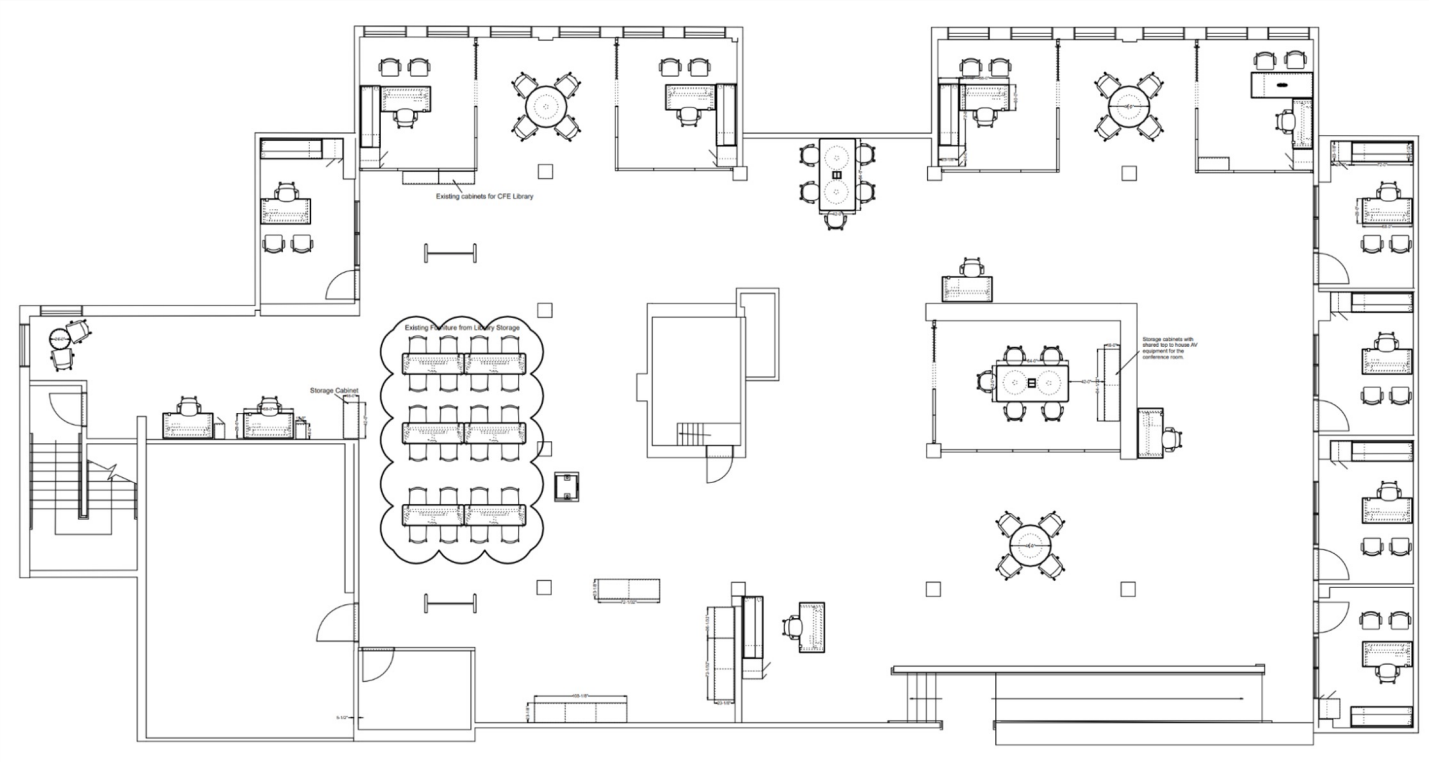
Partnership Formation- Research Alliance

- Office of Research Development (ORD)
- Center for Faculty Excellence (CFE)
- Undergraduate Scholars Program (USP)
- Research Cyberinfrastructure (RCI)
- MSU Library
 - Research Optimization, Analytics, and Data Services (ROADS)

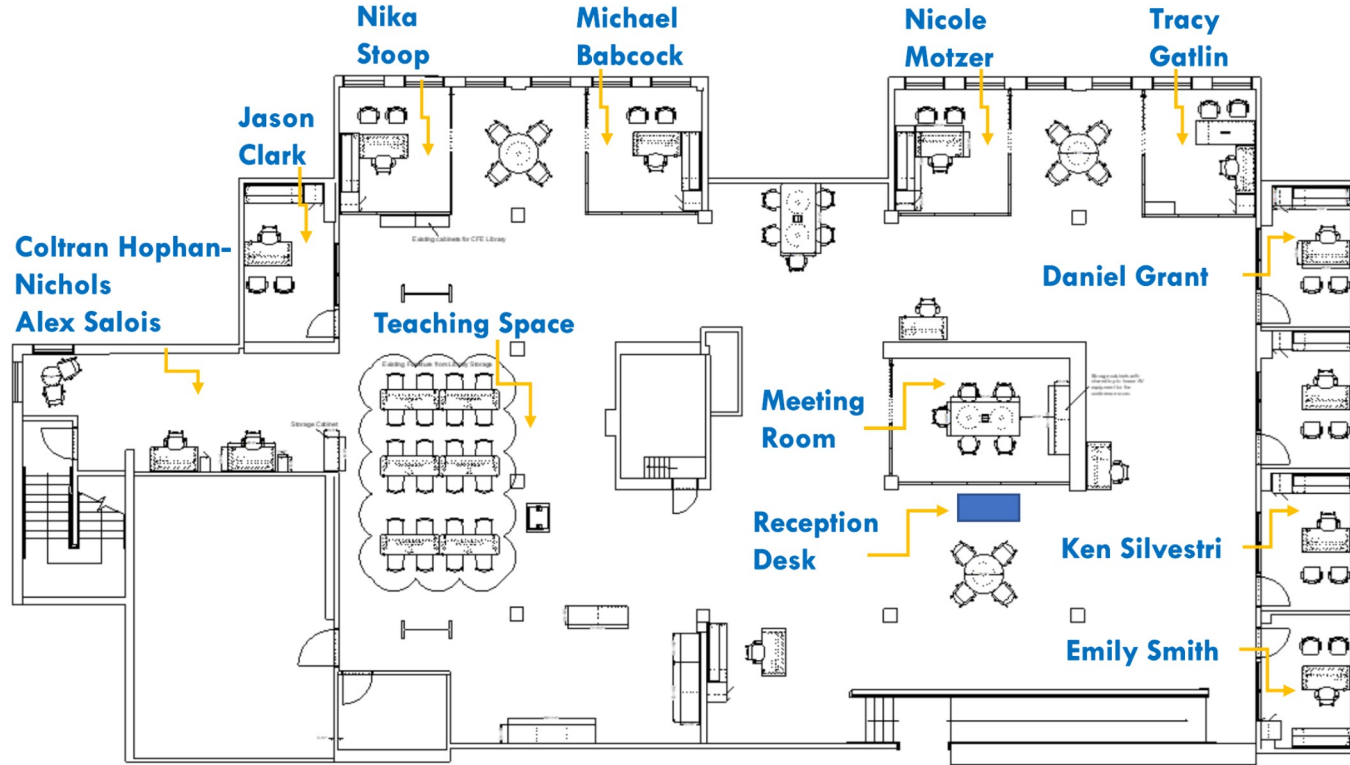
Researcher Alliance

The Research Alliance brings together expertise from across MSU to help researchers achieve their goals. We make grant proposals more competitive, make your research more visible, provide data management and publishing support, and help you translate your research into the classroom.

Partnership: Spaces (3rd Floor diagram)



Partnership: Spaces (3rd Floor Diagram - Labeled)



Researcher Alliance

Uniting research expertise in a single space, simplifying faculty and student access to services that will increase their research impact.

Supporting successful research from idea to achievement.

Your one-stop shop to increase your research impact.

- Agreeing to a Memorandum of Understanding (MOU) (linked here)
- Move from virtual to in-person partnership
- Defining services and communications
- Naming, signage
- Communication channels: Email list, reservable spaces
- For the users, but also for the partners

Projects + Collaborations

Assessment of Research Support and Data Needs: RCD CM

Ranking for: Montana State University			All Data (41)
<i>Facing Area</i> (click the "+" to the left of each to expand)	Coverage	Tier	Average Coverage
Researcher Facing Capabilities	57%	II	53%
<i>Research Computing and Data Staffing</i>	68%	II	60%
<i>Research Computing and Data Outreach (Initial Contact)</i>	42%	III	46%
<i>Research Computing and Data Advanced Support</i>	56%	III	53%
<i>Research Computing Management of the Research Lifecycle</i>	73%	I	59%
Data Facing Capabilities	56%	II	49%
<i>Data Creation</i>	70%	II	57%
<i>Data Discovery and Collection</i>	40%	III	51%
<i>Data Analysis</i>	60%	II	49%
<i>Data Visualization</i>	61%	II	48%
<i>Research Data Curation, Storage, Backup, and Transfer</i>	65%	I	49%
<i>Research Data Policy Compliance</i>	69%	I	50%
<i>Data Security/Sensitive Data Support</i>	23%	III	43%

Assessment of Data Need: RCD CM

Data Discovery and Collection	40%
Do researchers have access to data discovery consulting , i.e., to help them identify appropriate data repositories (on campus, in domains, and more generally)? <i>Note: this may come from Research Computing and Data staff, library staff, or other partners.</i>	70%
Do researchers have access to expertise about common Terms of Service for frequently crawled websites/data repositories and best practices guidance? E.g.,: i. library or other staff with knowledge about common Terms of Service for frequently crawled websites/data repositories and best practices guidance ? ii. library or other staff with skills and capacity to inform policies and educate researchers on data use agreements (DUAs) ?	10%
Do researchers have access to software supporting data collection (i.e., for data crawling, scraping, gathering, etc.)?	40%
Do researchers have access to resources (e.g., staff) to develop software supporting data discovery and collection ? E.g.,: i. resources to develop software for collection (crawling/scraping/etc.). ii. resources to develop user interfaces or do web development to collect and interact with data with appropriate security protocols and policies.	40%

Assessment of Research Support Need: RCD CM

Research Computing and Data Outreach (Initial Contact)		42%
Is there an institutional practice to proactively reach out to researchers, new faculty, or prospective faculty (for example during interviewing or new faculty onboarding processes) to explain research support services and help with computing beyond the desktop?		50%
Do researcher-facing staff have the skills and capacity to effectively engage in community outreach and broader impacts?		40%
Are researchers made aware of research computing and data related resources ? E.g.: i. intra-campus resources (e.g., support, training, engineering, central IT services, library services, related centers or institutes) ii. cross-institution, regional, national, and/or international entities that comprise the larger ecosystem of Research Computing and Data (e.g., ACI-REF, Campus Champions, Research Software Engineers, CASC, CaRCC, CI Engineers, PEARC)?		60%
Does your institution have a process to assess researcher awareness, satisfaction, and engagement related to Research Computing and Data services and support?		0%
Does your institution have a process to assess the impact of research computing and data support?		30%
Does your institution have marketing/communication resources (staff) with the skills and capacity to help publicize and explain research support services?		70%

Research Optimization, Analytics, and Data Services (ROADS)

- Open Education Theory and Practice
- Scholarly Communication and Publishing
- Data Science Applications
- Data Management and Data Curation
- Research Optimization and Analytics

Source: [ROADS Proposal](#)

- **Open Planning and Proposal Development** – finding funding opportunities, proposal development, submitting a proposal, individual development plans for retention, tenure and promotion
- **Training and Leadership** – technology, equipment and tools, data management and visualization, research integrity, managing teams and budgets, grants management, collaborating across disciplines, outreach and community engagement
- **Publication/Presentation** – writing groups; where to publish; outreach/research impact; creating retention, tenure and promotion dossiers
- **Preserve/Disseminate** – Data sharing, data curation, open access repositories
- **Measuring Impact/Putting into Practice** – evaluation, outreach and community engagement
- **New Ideas** – connecting faculty and students across disciplines, networking and ideation sessions, idea development

Activities (Research Alliance) - Shared Projects between Partners + Library

- Integration of Data Services with Research CyberInfrastructure
 - AI example: Using GPUs to pilot custom LLMs (dataset and software)
- Research Community Analysis with Center for Faculty Excellence (CFE)
 - Network Analysis of our Researchers (software)
- Teaching and Instruction with CFE
 - R and Python Workshops, Generative AI guidelines, AI Prompt Engineering, Telling your Research Story, etc.
- Researcher Visibility project with Office of Research Development
- Student mentorship and fellowships with Undergraduate Scholars Program

Research Analytics (Library) + Center for Faculty Excellence (Mentoring Matches)

Sara Mannheimer

Published Works (12):

- [RDAP15 Summit Report: Introduction](#)

 Sara Mannheimer

- [Ready, engage! outreach for library data services](#)

 Sara Mannheimer

- [A balancing act: The ideal and the realistic in developing Dryad's preservation policy](#)

 Sara Mannheimer

► Show non-MSU authors

- [A multi-disciplinary perspective on emergent and future innovations in peer review](#)

 Sara Mannheimer

► Show non-MSU authors

- [Qualitative Data Sharing: Data Repositories and Academic Libraries as Key Partners in Addressing Challenges](#)

 Sara Mannheimer

► Show non-MSU authors

► Show other published works

Author Profile

► Author Details

Topic Cluster:

[Data mining](#)
[Data science](#)
[Database](#)
[Internet privacy](#)
[Knowledge management](#)
[Law](#)
[Operating system](#)
[Programming language](#)
[World Wide Web](#)

MSU Network Forecast:

[Scott D. Young](#) (3 collaborations)
[Doralyn Rossmann](#) (1 collaboration)
[David Swedman](#) (1 collaboration)
[Justin D. Shanks](#) (1 collaboration)
[Jason A. Clark](#) (1 collaboration)
[Stacey A. Hancock](#) (1 collaboration)
[Allison Theobald](#) (1 collaboration)
[Matthew Regan](#) (1 collaboration)
[Jason A. Clark](#) (1 collaboration)

▼ External Network Forecast:

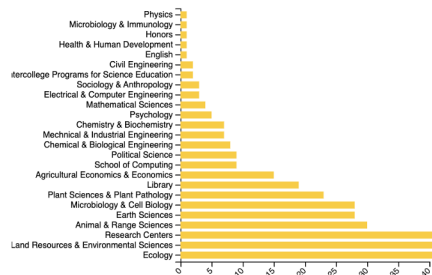
[Ayoung Yoon](#) - University of North Carolina at Chapel Hill (1 collaboration)
[Shea Swauger](#) - University of Colorado Denver (1 collaboration)
[Josmel Pacheco-Mendoza](#) - Universidad San Ignacio de Loyola (1 collaboration)
[Cameron Neylon](#) - Curtin University (1 collaboration)

Library offering:

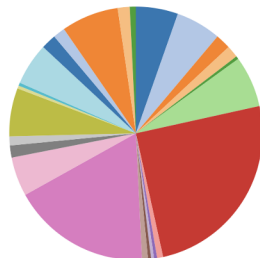
- Training data
- Topic clustering
- Network analysis
- External network forecasting

Researcher Visibility (Library) + Office of Research Development (Data Outreach)

Contributors by Department



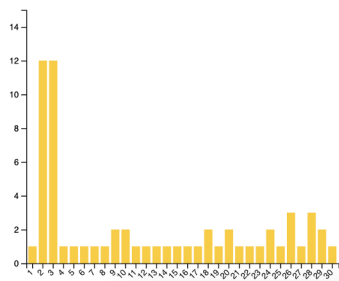
Total Datasets by Department



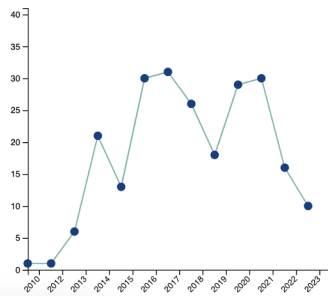
Library offering:

- Dataset Collection
- Metadata
- Visualization

Total Datasets by External Partners



Total Datasets by Year



Cyberinfrastructures + Data Science Consults



- Intentional work to connect
- Beyond co-location = concerted collaboration
- Social Interoperability (Rebecca Bryant, OCLC)
 - Events: Open House, Research Development Day
 - Data: Researcher Needs Assessment 2023
 - Monthly meetings with shared leadership
 - Weekly check-ins or informal visits
 - Tea time + Potlucks

[See: Social Interoperability at MSU \(From OCLC\)](#)

Contact Us

Doralyn Rossmann
Dean, Montana State University
Library
doralyn@montana.edu
doralyn.org

Jason A. Clark
Head, Research Optimization,
Analytics, and Data Services,
MSU Library
jaclark@montana.edu
www.jasonclark.info

Michael Navicky
Director of High Performance
Computing,
Mississippi State University
navicky@hpc.msstate.edu

Lauren Geiger
Digital Archivist,
Mississippi State University Libraries
lgeiger@library.msstate.edu

Discussion?

Questions?

Facilitated Discussion: Leading Questions, Open Questions from Audience

- How do you see building interdisciplinary support expertise on your campus?
- Where do you see challenges or opportunities here?
- Are you seeing support at the administration level for these types of collaborations/ideas?
- How might you take these ideas to other corners of your administration or faculty?
- Is there an opportunity at your institution for developing this type of ecosystem for research computing and data?