Development of a Profession: A Research Computing and Data Community

Scott Yockel
University Research Computing Officer
Harvard

Campus Research Computing Consortium

an organization of dedicated professionals developing, advocating for, and advancing campus research cyberinfrastructure (CI) and associated professions

think ***research computing & data*** **(RC&D)** infrastructure, including human infrastructure

Activities include:

Professional Community ("People Network")

- in the form of year-round virtual conference with focused tracks; many calls on YouTube

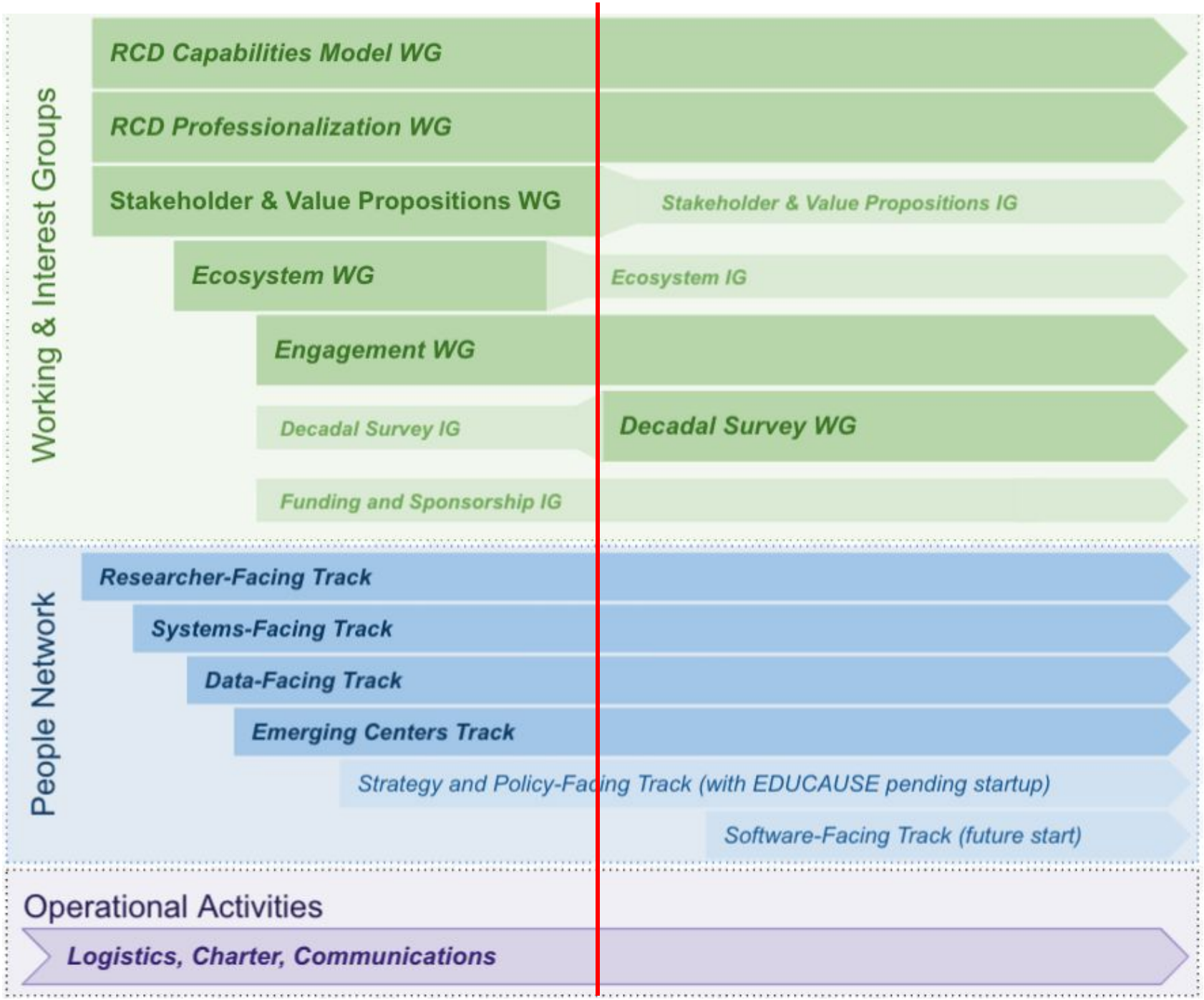
Working Groups with specific objectives/products, including:

- Professionalization and Workforce Development
- Developing a common Capabilities Model for Research Computing and Data
  - Assessment Tool for campus research CI, covered in the previous CNI presentation
- Research IT Ecosystem
  - 2018 workshop, survey and catalog research CI entities, submitted PEARC paper

Interest Groups for discussion, which may lead to objectives/working groups

CaRCC
Campus Research Computing Consortium

https://carcc.org/how-to-join/



CaRCC
Campus Research Computing Consortium

Principles & Charter

Inclusive, transparent, trustworthy, collaborative, consensus driven

- CaRCC supported activities are community-owned (and ‘open’)
- CaRCC facilitates and spans boundaries
- CaRCC welcomes partners and collaborations with shared credit & attribution

Examples of Collaborators / Partners:

Campus Champions

Why Professionalize Research Computing and Data?

To address several national, organizational, and occupational challenges:

➔ National shortage of Research Computing & Data personnel
- Individuals do not understand RCD *as a career*
  - No formal career path structure
  - Little awareness of career potential

➔ High employee turnover
- RCD groups struggle to recruit and retain staff
  - Where to find people?
  - How to train them?
  - How to develop and promote them?

Why Professionalize Research Computing and Data?

- ➔ Occupational distinction from
  - IT (i.e. Administrative/Enterprise IT)
  - Researchers in other disciplines
- ➔ Precarious employment and careers

See Berente, Nicholas, et al., Professionalization in Cyberinfrastructure (February 15, 2017): http://dx.doi.org/10.2139/ssrn.3138592

CaRCC
Campus Research Computing Consortium

RCD Professionalization Timeline

Scott Yockel | June 10, 2021

CaRCC
Campus Research Computing Consortium

NSF

RCN Meetings

RC Directors and Organizational Sociologist deliberate on the "CI Profession"

CI Professionalization Workshop

30 Institutions gathered to create the Job Elements and Career Guide, validating the "-facings" concept, outlining their roles/responibilities.

RCD Workforce Survey

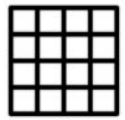
8 RCD professionals met over 5 months to create a national RCD survey, which will create the first large data set on the composition of the RCD workforce, a resource useful for individuals & hiring managers alike.

2016 | 2017 | 2018 | 2019 | 2020 | 2021

facings term introduced

To describe "who" the CI professionals are, we described what entity the staff faced on a regular basis: researchers, systems, software/data, strategy/policy)

HR Job Family Matrix



6 RCD & 2 HR professionals met over 6 weeks and create a HR framework template with a unique RCD job family to be used by any institution.

Career Arcs

15 RCD professionals began meeting to brainstorm methods to convey themes of Career Arcs from individual interviews.

Research IT Professionalization Process (Cornfeld)

Step 1: Establish a Research IT professional association (CaRCC?)

Step 2: Formalize and disseminate Research IT knowledge base

Step 3: Implement education programs for Research IT professionals and organizational managers

Step 4: Establish academic Research IT programs in universities

Step 5: Create Research IT graduate degree programs

Step 6: Establish local, national, and global Research IT labor market institutions

Step 7: Incubate start-up Research IT professional-services firms

CaRCC
Campus Research Computing Consortium

Outcomes of the RCN Meetings: Nicholas Berente, James Howison, Joel Cutcher-Gershenfeld, John L. King, Stephen R. Barley, John Towns. http://dx.doi.org/10.2139/ssrn.3138592

CaRCC: Professionalization Working Group

Professionalization: the process by which our occupation organizes itself to ensure its practitioners perform their services well and thereby viewed as its own distinct occupation and as an integral part of scholarly research.

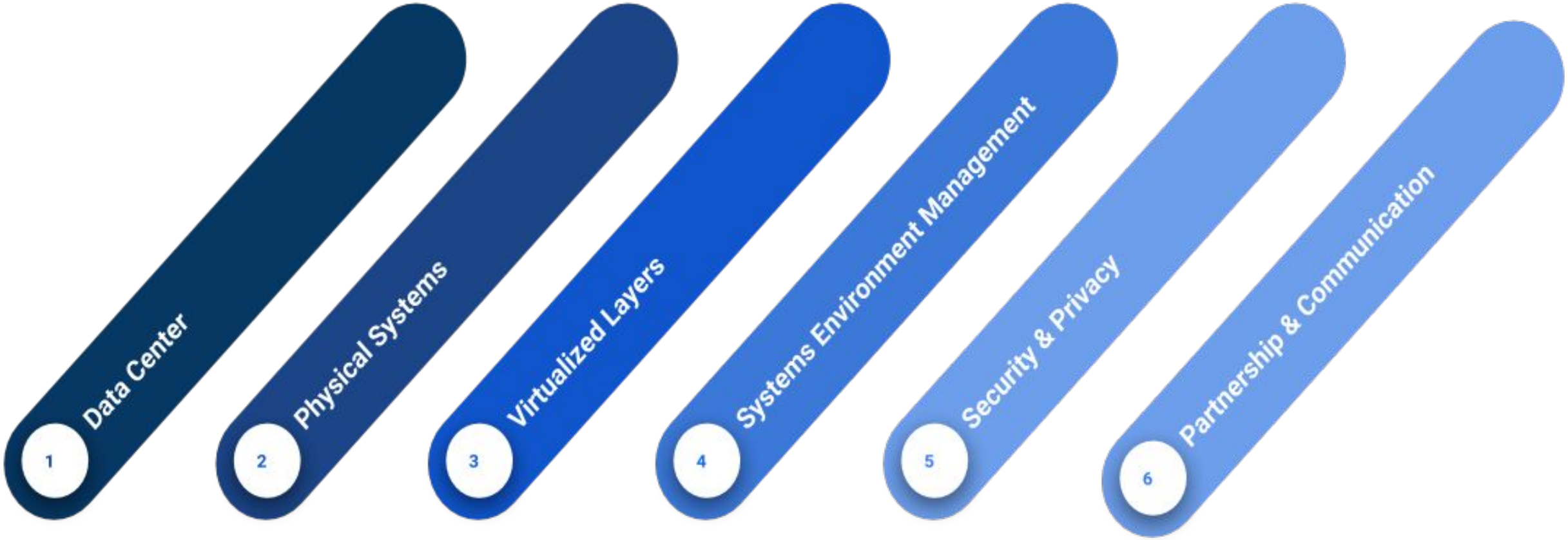
Goal: develop and disseminate frameworks and approaches to institutional leaders in HR, IT, Research, … that elevates RC&D as a distinct and highly valued career path.

Value: attracting, retaining, diversifying, and developing professional talent in RC&D.

2018 CI Professionalization Workshop Goals

- Learn from leading practices regarding Research IT job definitions and career paths
- Develop a flexible framework to organize Research IT hiring, career development, retention and other aspects of HR in the ecosystem
- Apply the framework for Research IT work:
  - Researcher-facing roles
  - Systems-facing roles
  - Software- & Data-facing roles
  - Strategy/Policy-facing roles

Systems-facing spectrum



CaRCC

Researcher-facing roles

Providing possible solutions to facilitate and/or transform research

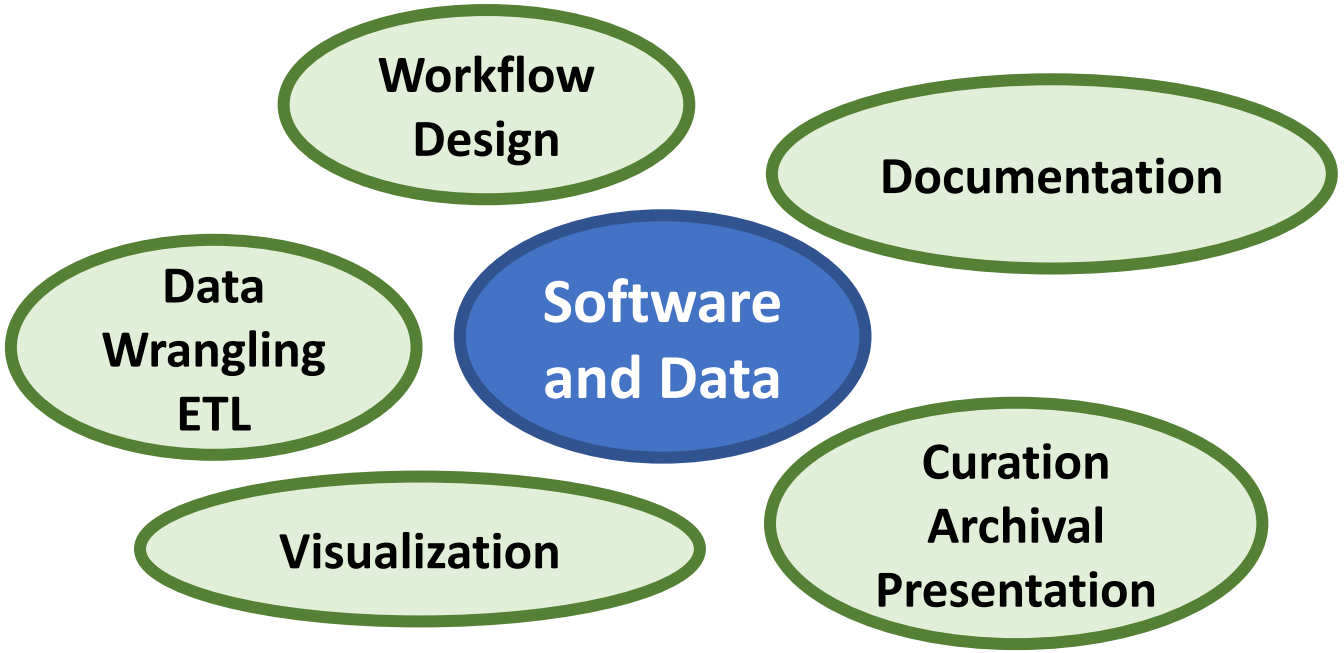
Engaging researchers to co-create and co-learn relevant advanced computing and data capabilities

Facilitating collaboration and coordination with other people in the technology/research landscape

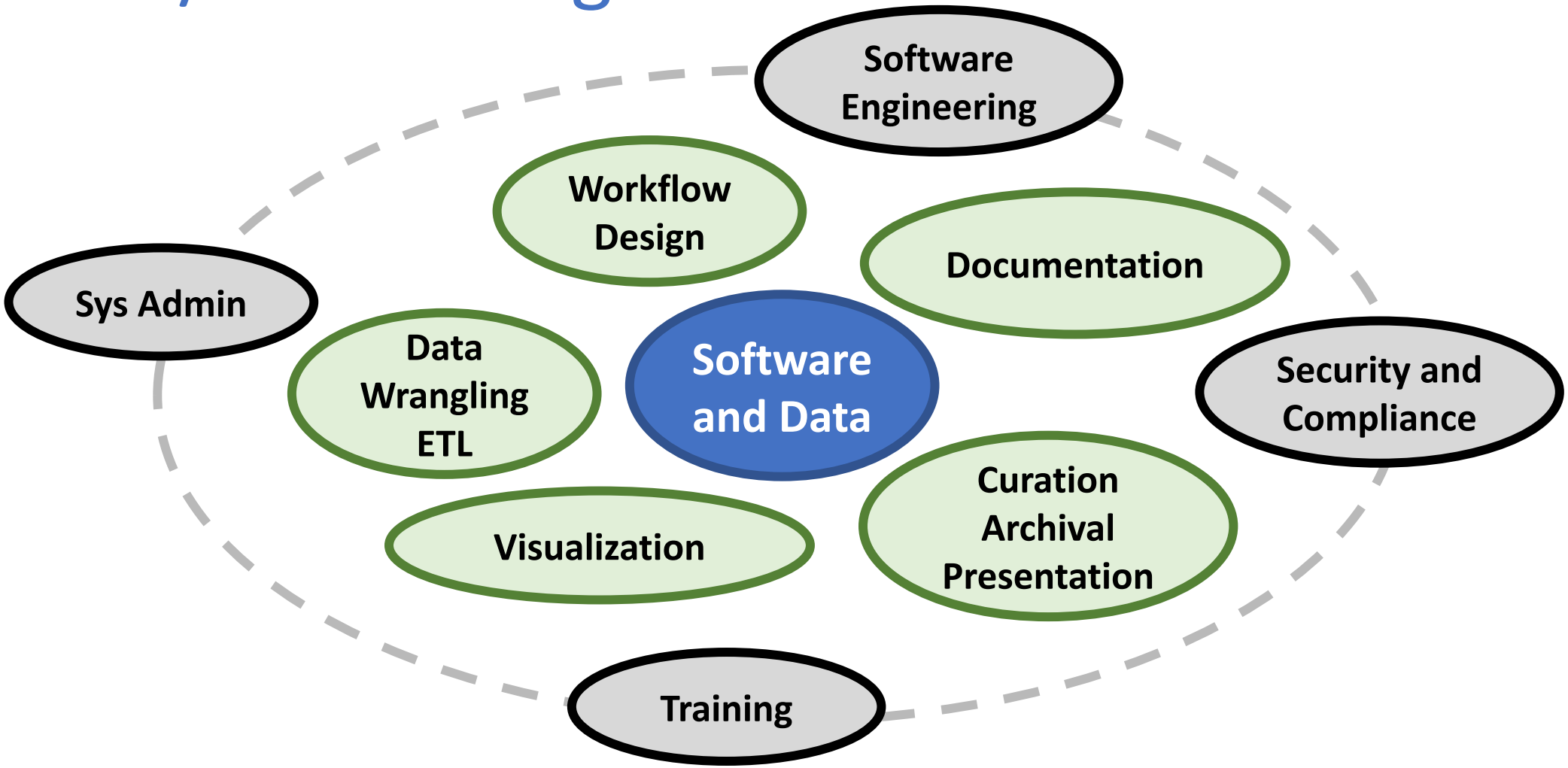
Providing regular communications to the campus or external research community

Actively solving problems & enhancing learning with full awareness of the local and larger research computing and data landscape.

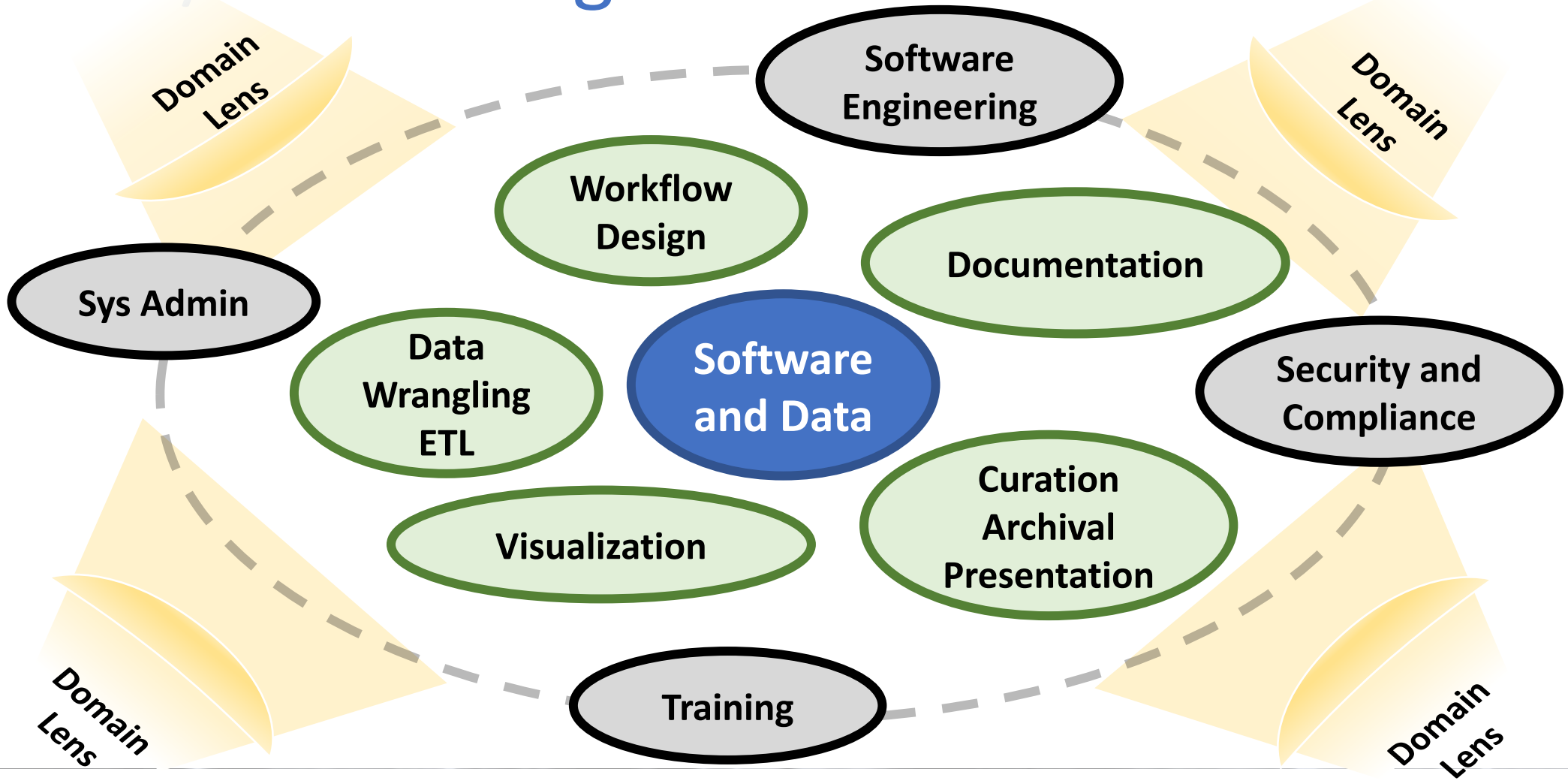
Software/Data facing roles



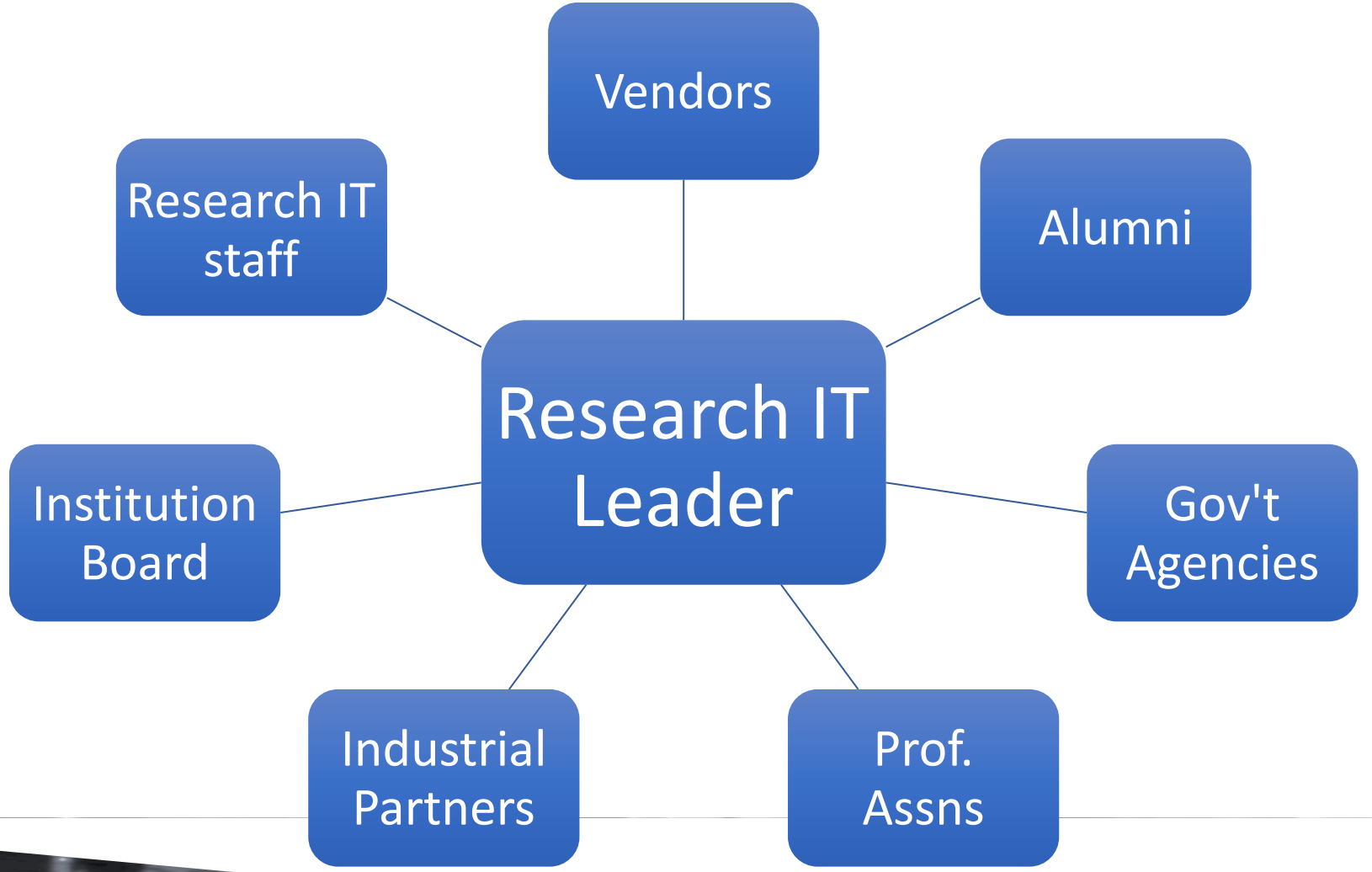
Software/Data facing roles



Software/Data facing roles



Strategy and Policy Facing Network



2018 CI Professionalization Workshop Goals

- Learn from leading practices regarding Research IT job definitions and career paths
- Develop a flexible framework to organize Research IT hiring, career development, retention and other aspects of HR in the ecosystem
- Apply the framework for Research IT work
- Anticipate potential complications and disconnects when implementing the framework across diverse campuses
- Specify next steps in the utilization of the framework and, as a result, the further professionalization of Research IT work

Common Themes from CaRCC workshop

Co-Creation (partnering with researchers): Research computing and data professionals are co-creating methods and software models; Collaborative process, very different from delivery of traditional IT and software services

Career Paths are incomplete in most organizations; creating challenges for recruiting, developing and retaining these professionals.

Digital: The exponential growth of digital technologies underlies work; accelerating change in the work due to changes in hardware, software, systems, and the nature of the data itself.

Status: Work of research computing & data professionals generally held in high regard by faculty with whom they work; important status and power differences between these professionals and principle investigators that are part of a larger "two-tier" culture in most university settings.

Terminology: Work centered on "cyberinfrastructure for research" and touches on many related domains, including "data science" and "high performance computing." This work is distinct from, but connected to the work of "information technology" professionals.

RCD Professionalization Timeline

Scott Yockel | June 10, 2021

CaRCC
Campus Research Computing Consortium

NSF

RCN Meetings

RC Directors and Organizational Sociologist deliberate on the "CI Profession"

CI Professionalization Workshop

30 Institutions gathered to create the Job Elements and Career Guide, validating the "-facings" concept, outlining their roles/responibilities.

RCD Workforce Survey

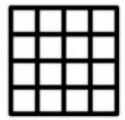
8 RCD professionals met over 5 months to create a national RCD survey, which will create the first large data set on the composition of the RCD workforce, a resource useful for individuals & hiring managers alike.

2016 | 2017 | 2018 | 2019 | 2020 | 2021

facings term introduced

To describe "who" the CI professionals are, we described what entity the staff faced on a regular basis: researchers, systems, software/data, strategy/policy)

HR Job Family Matrix



6 RCD & 2 HR professionals met over 6 weeks and create a HR framework template with a unique RCD job family to be used by any institution.

Career Arcs

15 RCD professionals began meeting to brainstorm methods to convey themes of Career Arcs from individual interviews.

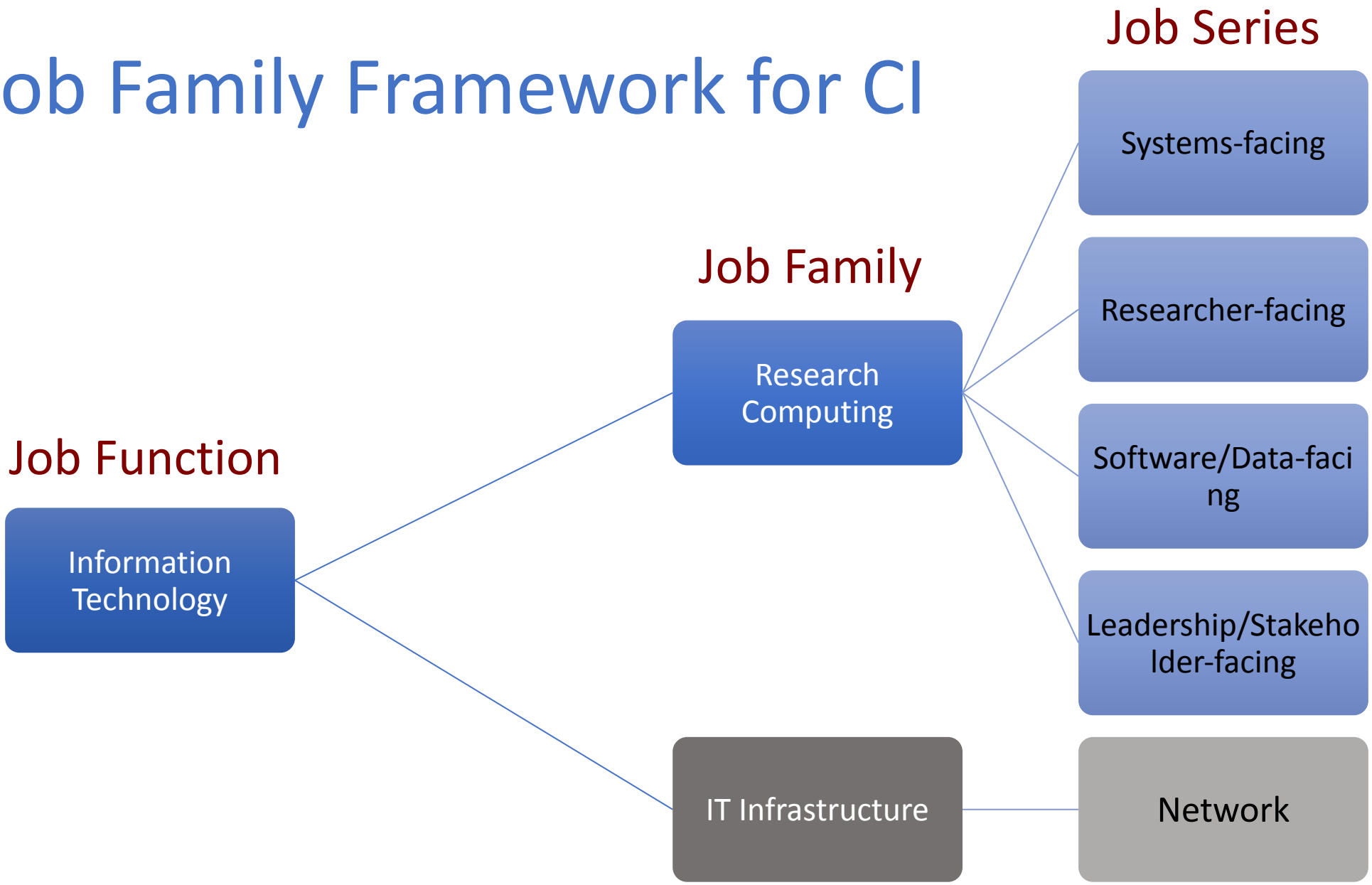
2019 HR Job Family Framework - Timeline

- 6 weeks - 1 hr meetings – April 8th – May 13th
  - Week 1: Agree upon framework used by Stanford and Harvard IT.
  - Week 2-5: skeleton draft (mid-level) --> email draft --> review on call --> flush out series --> finalize series --> …
  - Week 6: Review supporting document
- Sent out for review to all participants of 2018 CI Prof. Workshop
- Addressed and/or incorporated comments June-July

Scott Yockel (Harvard)	Wayne Gilmore (Boston U)
Melissa Lucius / Nicole Breen (Harvard HR)	Eric Adams (Purdue)
Erik Deumens (U FL)	Brian Balderston (SDSC)
Anna Thrombly (U FL, HR)	Janae Baker (Rutgers)

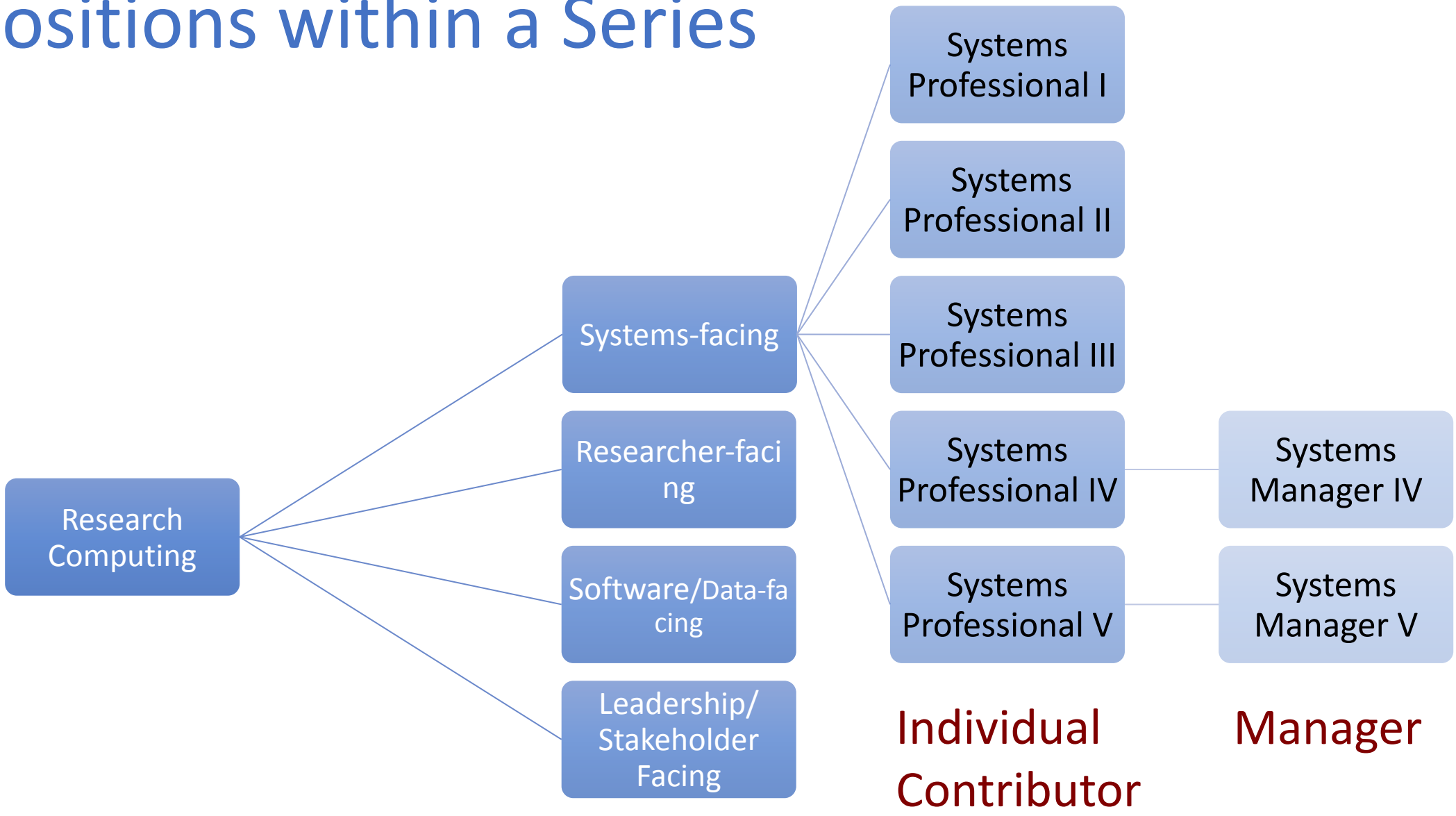
CaRCC
Campus Research Computing Consortium

HR Job Family Framework for CI



CaRCC
Campus Research Computing Consortium

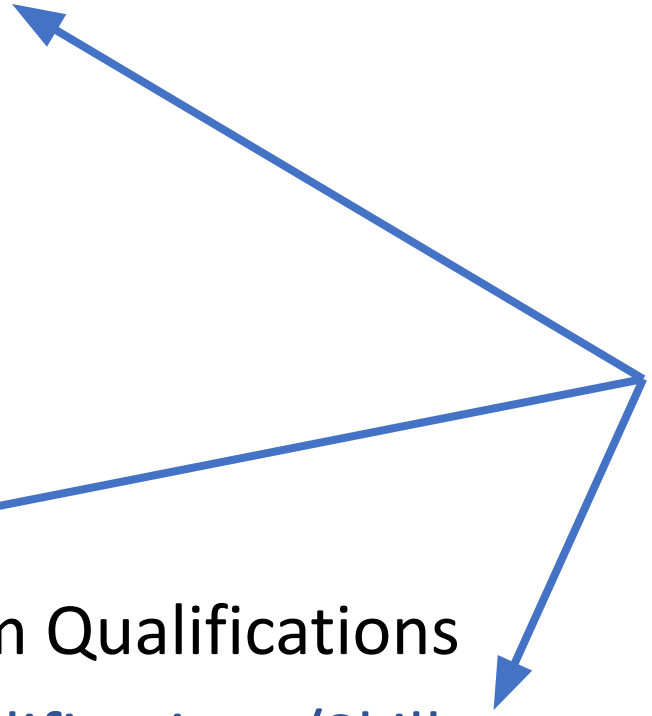
Positions within a Series



CaRCC
Campus Research Computing Consortium

Position Components

- Business Title
- Job Title
- Job Code
- Grade Level
- Job Summary
- Core Duties
- Basic/Minimum Qualifications
- Additional Qualifications/Skills



Parts to consider customizing per specific job

CaRCC
Campus Research Computing Consortium

https://carcc.org/products/ → HR Job Family Matrix

Major Points of Concern

- Non-managerial career paths
  - for those that have the desire to grow into technical leadership roles but do not desire to manage, recruit, and retain professional staff
- Distinctness from Enterprise IT
  - lower level positions have more overlap, but the direction of the series gets much broader at the top, while in Enterprise IT the technical leads generally have a deeper and narrower or more focused scope as they become the expert.
- Roles intentionally left out
  - business operational, research administration, library, finance, purely research
- Lacking in current SysAdmin roles
  - internal development of deployment practices, business operations, and collaborations between different technology entities whether it be data center staff, campus networking/security, and other vendors

Major Points of Concern

- Education vs Experience
  - skills and experiences drive the desired minimum criteria -- there is not a specific required education for each series
  - research-facing series: we have found it crucial for a Facilitator to have a specific academic background with demonstrated research experience as this helps them relate better to the plight of the researcher and the research process.
- Single series for Software/Data-facing?
  - core duties we focused on are relevant to both areas
  - level 4 and 5 are essentially one in the same with a comprehensive set of knowledge and skills.
- How are there multiple Director levels?
  - progression changes with longer timelines and increasing scope of projects and collaborations.
  - position is commensurate with other leadership positions that they regularly engage

Harvard Specific Implementation

- Data Scientist: Research Function or Technology Function?
- Bioinformatics: Research Function or Technology Function?
- Consultants vs Facilitators Job Series
- Level 4 or 5 Individual Contributor push back because too few.
- Management at Level 3 added
- No Leadership/Stakeholder facing
- Identified many position outside of centralized RC groups.
  - Especially RSE & Data Scientists in individual labs.

RCD Professionalization Timeline

Scott Yockel | June 10, 2021

CaRCC
Campus Research Computing Consortium

NSF

RCN Meetings

RC Directors and Organizational Sociologist deliberate on the "CI Profession"

CI Professionalization Workshop

30 Institutions gathered to create the Job Elements and Career Guide, validating the "-facings" concept, outlining their roles/responibilities.

RCD Workforce Survey

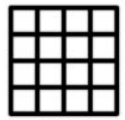
8 RCD professionals met over 5 months to create a national RCD survey, which will create the first large data set on the composition of the RCD workforce, a resource useful for individuals & hiring managers alike.

2016 | 2017 | 2018 | 2019 | 2020 | 2021

facings term introduced

To describe "who" the CI professionals are, we described what entity the staff faced on a regular basis: researchers, systems, software/data, strategy/policy)

HR Job Family Matrix



6 RCD & 2 HR professionals met over 6 weeks and create a HR framework template with a unique RCD job family to be used by any institution.

Career Arcs

15 RCD professionals began meeting to brainstorm methods to convey themes of Career Arcs from individual interviews.

Building the research innovation workforce: a workshop to identify new insights and directions to advance the research computing community.

NSF Org:	OAC Office of Advanced Cyberinfrastructure (OAC)
Initial Amendment Date:	July 14, 2020
Latest Amendment Date:	September 25, 2020
Award Number:	2036534
Award Instrument:	Standard Grant
Program Manager:	Alan Sussman OAC Office of Advanced Cyberinfrastructure (OAC) CSE Direct For Computer & Info Scie & Enginr
Start Date:	July 15, 2020
End Date:	June 30, 2021 (Estimated)
Awarded Amount to Date:	$163,960.00
Investigator(s):	Thomas Hacker tjhacker@purdue.edu (Principal Investigator) Dana Brunson (Co-Principal Investigator) Preston Smith (Co-Principal Investigator) Lisa Arafune (Co-Principal Investigator)

~100 participants

3 webinars (Aug-Sept 2021)

RCD professionals, industry, domains

Recommendations for the RCD Community

- Increase community awareness of ***CI/RCD work as a unique career path*** with an identity distinct from enterprise IT
- A community organization needed to communicate ***HR job families for CI/RCD***, and to develop supporting materials for ***defining career arcs***, identifying good examples of internship and training programs, and dissemination to the community
- ***Educate*** university and non-profit management about the ***role of HPC/CI/RCD professionals*** to communicate their value added to research and researchers
- Organizations need to ***create a compensation, career path, and promotion model*** that ***accommodates the dual research/service roles*** present in the workforce ecosystem
- Bring together the CI community with ***experts in Diversity, Equity, and Inclusion (DEI) and social science*** to study the ***problem of recruiting and sustaining a diverse and inclusive workforce*** and develop a plan of action
- Establish a ***professional organization*** (e.g., like the Campus Research Computing Consortium or within EDUCAUSE) to ***formalizing the profession***, training, etc. to improve the CI workforce ecosystem
- NSF could provide seed funding to incentivize the community to undertake a ***national census of positions, pay grade, cost of living differences, and benefits for the CI/RCD workforce***

2020-2021 RCD Professionals Census

- Bimonthly 1 hr meetings – Dec 2020 - April 2021
  - Dec: Agree upon scope of census, review other similar surveys
  - Jan-Mar: Design survey for individuals
  - Apr: Send out for review (beta tested)
  - May: IRB review & approval
  - Jun-Sept: Survey Open

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.

Scott Yockel (Harvard)	Ashley Stauffer (Penn State)
Timothy Middelkoop (Internet2)	Kimberly Grasch (U Chicago)
Jay Alameda (UIUC)	Amy Neeser (UC Berkeley)
Galen Collier (Rutgers)	Christina Maimone (Northwestern)

CaRCC
Campus Research Computing Consortium

PLEASE TAKE THE SURVEY

tinyurl.com/rcdprof

CaRCC
Campus Research Computing Consortium

What type of data would you like to have?

On Individuals

- Personal Background
- Role: Basics
- Role: Type of work
- Career/Job Perceptions

2021 Elements of a Career Arcs Resource

- Narratives of how (and why) existing RCD staff got into these roles
  - Allow potential staff to see themselves in a story
  - Help groups/managers understand the diversity of career paths into RCD
  - Emergent patterns will identify gaps/opportunities, inform strategic planning, etc.
- Examples of effective recruitment programs
  - Targeted at staff: previous roles, training needed for RCD roles
  - Targeted at students: identifying candidates, making the pitch, effective training
- Next steps for existing RCD staff
  - Paths for advancement: Technical expertise, domain expertise, leadership, etc., etc.
  - Training requirements, sources/providers
- Working with industry partners
  - Staff transitions to and from industry

CaRCC
Campus Research Computing Consortium

Efforts led by Patrick Schmitz

Challenges for RCD Professionals Careers

- Individuals do not understand RCD *as a career*
  - No formal career path structure
  - Little awareness of career potential
- RCD groups struggle to recruit and retain staff
  - Where to find people?
  - How to train them?
  - How to develop and promote them?

Award Abstract # 2100003

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

NSF Org:	OAC Office of Advanced Cyberinfrastructure (OAC)
Awardee:	UNIVERSITY CORPORATION FOR ADVANCED INTERNET DEVELOPMENT
Initial Amendment Date:	May 19, 2021
Latest Amendment Date:	May 19, 2021
Award Number:	2100003
Award Instrument:	Standard Grant
Program Manager:	Alan Sussman alasussm@nsf.gov (703)292-7563 OAC Office of Advanced Cyberinfrastructure (OAC) CSE Direct For Computer & Info Scie & Enginr
Start Date:	May 15, 2021
End Date:	April 30, 2023 (Estimated)
Total Intended Award Amount:	$1,497,243.00
Total Awarded Amount to Date:	$1,497,243.00
Funds Obligated to Date:	FY 2021 = $1,497,243.00
History of Investigator:	Dana Brunson (Principal Investigator) dbrunson@internet2.edu (405)338-5678 Thomas Cheatham (Co-Principal Investigator) Scott Yockel (Co-Principal Investigator) Patrick Schmitz (Co-Principal Investigator) Claire Mizumoto (Co-Principal Investigator)

Discussion and questions

CaRCC
Campus Research Computing Consortium

Appendix