# Assessing the Landscape of Research Computing and Data Support

## The 2020 RCD Capabilities Model Community Dataset

Patrick Schmitz
patrick@sempercogito.com
Semper Cogito Consulting
Berkeley, CA, USA

## ABSTRACT

We describe the first Research Computing and Data Capabilities Model Community Dataset, aggregating the assessments of 41 Higher Education Institutions. These assessments were completed using the 1.0 version of the Research Computing and Data Capabilities Model (RCD CM), over a period of several months in the Spring and Summer of 2020. The RCD CM allows organizations to self-evaluate across a range of RCD services and capabilities for supporting research, leveraging a shared vocabulary to describe RCD support. The Model supports a range of stakeholders and provides structured input to guide strategic planning and enable benchmarking relative to peer institutions. This Community Dataset provides insight into the current state of support for RCD across the community and in a number of key sub-communities. The dataset shows stark differences between Public and Private institutions, between institutions with a larger and smaller share of national funding, etc. In many cases, the patterns in the data confirm common perceptions about support across the community, but the dataset provides a quantitative baseline for understanding current RCD support, as well as granular insights to groups of institutions that are seeking to collaborate on shared solutions and strategies to advance RCD support. Over time, longitudinal data will provide additional insight into trends, and a means of evaluating the impact of programs designed to increase RCD support.

## CCS CONCEPTS

• **Social and professional topics** → **System management**; • **General and reference** → **Evaluation**; *Computing standards, RFCs and guidelines.*

## KEYWORDS

Research Computing and Data, Research IT, Cyberinfrastructure, Capabilities Model, Assessment, Strategic Planning

## 1 INTRODUCTION

Research is increasingly dependent upon Cyberinfrastructure (CI), from instruments and sensors to Research Computing and Data (RCD) infrastructure and services. RCD is being used in new domains and is expanding beyond High Performance Computing (HPC) into secure computing; big data management; AI/machine learning; and into heterogeneous compute models, edge computing, and cloud-based computing. The rapid evolution and diversification of RCD poses significant challenges to academic institutions as they try to effectively assess and plan for the necessary resources required to keep pace with the growing needs of researchers. These challenges are especially acute for smaller and emerging RCD support organizations, which often lack experience supporting RCD and have limited resources to develop an analysis framework for strategic planning. Many institutions would like to assess their capabilities in comparison to peers, and have traditionally lacked a framework (much less the data) to do so. The Research Computing and Data Capabilities Model (RCD CM) [7] allows organizations to self-evaluate across a range of RCD services and capabilities for supporting research, leveraging a shared vocabulary to describe RCD support. The Model supports a range of stakeholders and provides structured input to guide strategic planning and enable benchmarking relative to peer institutions.

Forty-one institutions completed assessments using the RCD CM and contributed these to the 2020 Community Dataset. These institutions represent 28 states, a mix of public and private, R1, R2, and other Carnegie Classifications, and also include a number of key institutional demographic sub-communities. The resulting dataset provides important insights into the state of support for RCD, at both a summary as well as a granular level. The Dataset also clearly shows the different levels of RCD support among certain sub-communities. In many cases, the patterns in the data confirm common perceptions about support across the community and particularly about relative levels of support (and gaps) among sub-segments of the community. While these conclusions may be unsurprising to some, it is important to provide quantitative data so that we have a baseline for understanding RCD support broadly and in these sub-communities. In several cases the data made clear that differences among certain groups are even more profound than

many may have expected and this allows RCD leadership and others to refine their understanding of which areas of RCD support may merit attention.

In the next sections we provide some background on the Capabilities Model; a description of the structure of the Community Dataset; visualization and analysis of significant patterns and themes for capabilities coverage; and a description of the priorities identified by institutions. We close with conclusions and future work.

## 2 THE RESEARCH COMPUTING AND DATA CAPABILITIES MODEL

The initial version of the RCD Capabilities Model was developed as a collaboration among the Campus Research Computing Consortium (CaRCC) [1], Internet2 [5], and EDUCAUSE [2], with support from the National Science Foundation (NSF OAC-1620695) and from many volunteers who provided input and review from a diverse set of universities (large and small, public and private) and related organizations. The 1.0 version became publicly available in January 2020, and has been downloaded by over 125 institutions across 44 states and 2 Canadian Provinces, including both public and private institutions, a range of Carnegie classifications (including over half of all R1 institutions), and many EPSCoR-eligible[1] and minority-serving institutions. The RCD CM Assessment Tool presents roughly 150 capabilities (in the form of questions) structured around the five facings that are increasingly used as a means of characterizing the roles of people who support RCD: Researcher Facing, Data Facing, Software Facing, System Facing, and Strategy and Policy Facing.



**Figure 1: Excerpt from the Researcher Facing sheet**

The initial version of the Assessment Tool is implemented as a spreadsheet workbook, with one sheet per facing. On each of these sheets there is a list of questions that represent key aspects or factors associated with supporting Research Computing and Data;

the questions are grouped into themes both for general usability of the tool, and as well to provide a more granular summary of capabilities (an excerpt is presented in Figure 1). For each row, an assessment team will answer the question from three perspectives or lenses, and the answers are combined to produce a numerical coverage value for that aspect in the model:

(1) How widely is support deployed at the institution (from *No support* to *Deployment institution-wide*)
(2) What is the level of multi-institutional collaboration (e.g., *Exploring collaboration*, *Leading a collaboration*, etc.)
(3) What is the Service Operating Level (E.g., *No service*, *Lights-On Only*, etc.)

The calculated coverage values are combined to produce a summary coverage value for thematic groupings within each facing, and are then aggregated into a coverage value for each facing. In addition to the facings sheets, the Assessment Tool also presents a summary sheet that rolls up the assessment results into a single page for use in presentation to leadership. As an assessment team works through the tool, they may also identify specific aspects as an area of priority in their institutional planning and mark these as either Medium Priority or High Priority (these do not contribute to the coverage values and are just for local strategic planning work). For details see [7] and [4].

## 3 THE 2020 RCD CM COMMUNITY DATASET

The Community Dataset[2] structure mimics that of the RCD Capabilities Model in order to facilitate benchmarking analyses by institutions that completed the RCD CM Assessment Tool. This includes the primary structure around the five facings, as well as the grouping of questions into topics under each facing. The Community Dataset aggregates coverage values for all the contributing institutions from the question level up through the summary values for the facing.
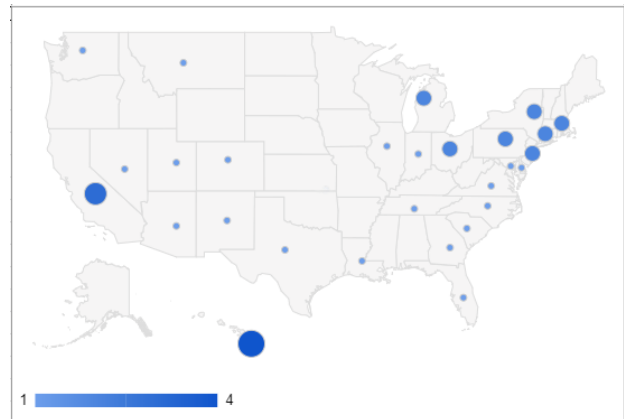


**Figure 2: Contributing Institutions by U.S. State**

As institutions request a copy of the RCD CM Assessment Tool, they provide some basic institutional demographic information; this allows us to understand how representative the data is, and

| Facing Area (click the "+" to the left of each to expand) | Coverage | Tier | All Data (41) Average Coverage | Public (31) Average Coverage | Tier among Publics | Tier among Privates | Tier among R1s | Tier among R2s |
|---|---|---|---|---|---|---|---|---|
| **Researcher Facing Capabilities** | 52% | III | 53% | 52% | II | IV | III | III |
| Research Computing and Data Staffing | 61% | III | 60% | 58% | II | IV | III | III |
| Research Computing and Data Outreach (Initial Contact) | 39% | III | 46% | 44% | III | IV | III | III |
| Research Computing and Data Advanced Support | 60% | II | 53% | 52% | II | III | III | I |
| Research Computing Management of the Research Lifecycle | 55% | III | 59% | 58% | III | III | III | I |
| **Data Facing Capabilities** | 44% | III | 49% | 47% | III | IV | III | II |
| **Software Facing Capabilities** | 49% | II | 46% | 45% | II | III | II | III |
| **System Facing Capabilities** | 60% | II | 56% | 57% | II | III | II | III |
| **Strategy and Policy Facing Capabilities** | 35% | IV | 56% | 53% | IV | IV | IV | III |
| **Total Organizational RCD Coverage** | 48% | III | 52% | 51% | III | IV | III | III |

**Figure 3: Screenshot of Benchmarking report tool (edited for simplicity)**

also to filter the data according to broad subgroups including public and private institutions, Carnegie classifications, EPSCoR eligibility and minority-serving status, etc.

- 28 states are represented in the data (see Figure 2); these states include roughly 80% of the R1 and R2 Universities in the US (as reported by Carnegie[3]), and so should be fairly representative.
- Roughly ¾ of the contributing institutions are public, and ¼ are private. Carnegie reports that about 70% (185 of 266) of R1 and R2 institutions are public, and 30% (80 of 266) are private, so the proportions in our data are comparable to the broader US mix.
- Just under 20% of our reporting institutions are designated as minority serving. This is slightly higher than the proportion of R1 and R2 institutions that Carnegie lists as minority serving (42 of 266, or about 16%). However, we note that our Dataset includes no Historically or Predominantly Black Colleges and Universities (HBCU/PBIs) or American Indian-serving (Tribal Colleges and Universities/TCU) institutions (a gap that we hope to remedy in future years).
- Somewhat over 24% of our reporting institutions are in EPSCoR eligible states. This is slightly higher than the 22% (58 out of 266) of R1 and R2 institutions Carnegie reports in these states, although only 9 of the 25 EPSCoR states are represented.

Since the release of the 2020 report a few months ago, it has been viewed or downloaded by 200 unique individuals. We do not have data on the readers' associated institutions, but we note that there are only about 260 R1 and R2 institutions in the U.S. and so this seems to reflect a high level of interest. Among the contributing institutions, 75% reported that access to the detailed report made them more likely to complete and contribute an assessment in the future, with the remainder being just as likely to do so.

## 3.1 Supporting Institutional Benchmarking

Many of the institutions indicated a strong interest in being able to benchmark their assessments relative to the community and also to demographic slices. In part due to resource constraints, and in an effort to make the benchmarking reports follow the same presentation model as the assessment tool, we created a benchmarking report that has the same structure and layout as the assessment spreadsheet. The tool queries content from the set of completed assessments and averages values for each row. Filtering on institutional demographic information produces the averages for community subsets, such as Public institutions, R1s, R2s, etc. The report for a given institution shows their capabilities coverage in the first column, and then has columns for the entire community and for the community subsets.

The Capabilities Model presents coverage as a percentage of possible capability on a scale that is intended to be highly aspirational, and most institutions will likely be in a mid range of coverage percentage for many areas. However, some institutions reported a degree of consternation at the coverage value *percentages* and the possible comparison to *grades* in a course. This, despite the fact that in contrast to the aspirational scale for capabilities coverage, course grades are intended to measure against a reasonable expectation of full comprehension and are not at all comparable. After several iterations and feedback from the community, we settled on a model that indicates in which of four *tiers* an institution's coverage sits, allowing them to see their comparative capabilities coverage. The Tiers are just four equal divisions of the institutions (computed for

each demographic slice). We considered a model in which Tier 1 was lowest, in part to allow for upward adjustment of the scale upwards (e.g., adding a new top Tier 5) as institutions develop broader coverage over time. However, this would alter the definition of the tiers and undermine benchmarking evaluation over time (e.g., to track the progress of one's institution in response to strategic investment in priority areas). We settled on a model in which the Tiers are labelled from I to IV (highest to lowest) and are color-coded from blue (Tier I) to red (Tier IV). Figure 3 illustrates the summary page of the resulting tool (edited to highlight the indication of an institution's RCD coverage Tier relative to several sub-communities). A full institutional report provides benchmarking comparisons for each topic across the five facings in the model.

Among the contributing institutions, over 80% reported that access to their individualized benchmarking report made them more likely to complete and contribute an assessment in the future, with the remainder being just as likely to do so. Just over half reported that access to the raw data would make them more likely to complete and contribute an assessment in the future; while there is clear interest in the underlying data, the reports and analysis seem to provide value to more institutions. When asked how they plan to use their individualized benchmarking report:

- 100% plan to use it as input to strategic planning
- Nearly 80% plan to use it to help justify funding of their campus RCD program
- Over a third (36%) will use it to support a grant proposal

Others reported plans to develop roadmaps, to help in analyzing a broad survey of researchers at their institution, and to support explorations of potential collaborative efforts.

## 3.2 Aggregating Institutional Priorities

We developed two tools to aggregate the areas marked as priorities by institutions. One follows the structure of the Assessment Tool itself, however the relative sparsity of the data (institutions did not consistently mark items for priority), and the lack of a good mathematical model to summarize such sparse, inconsistent data by area limits the utility of this as a method of understanding patterns. A second tool simply filters the data to select the top ten individual areas of priority across the entire RCD Capabilities Model (with filters for different demographic slices). While the first tool does give some broad indication of where institutions planned to prioritize attention and resources, the second tool does a better job of pulling out specifics. Results of our analysis are summarized in Section 5.

## 4 SIGNIFICANT THEMES FOR CAPABILITIES COVERAGE

One of the most striking aspects of the Community Dataset is the significant variation in the data. For the data as a whole and for many of the subsets in the data (selecting a facing, a theme, etc. and by different demographic slice), the standard deviation is often a very large proportion of the mean value. In a few cases, we had to cap the minimum values in visualizations graphs to 0 because error bars (at one standard deviation) extended below the associated axis.

The scatter graph in Figure 4 illustrates the range and variation of assessed RCD capabilities coverage for the institutions represented in the 2020 Dataset. Each vertical stripe represents a given
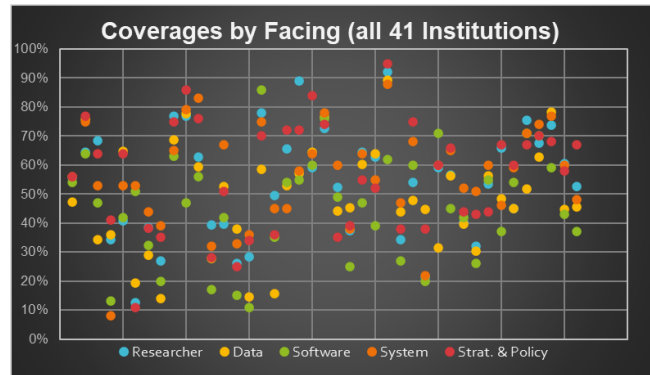


**Figure 4: Scatter graph of capabilities coverage by facing for all 41 institutions**

institution (in no particular order) with 5 colored dots indicating the summary coverage for the 5 facings. Several features are worth noting in the scatter graph visualization:

(1) The coverage values are literally all over the map, from very low to very high.
(2) Only a few institutions have coverage values that are consistent across facings. Most have fairly different levels of coverage in each facing or at least one facing for which coverage is quite different.
(3) There is little commonality to the relative ranking of facing coverage across institutions. I.e., different institutions have strengths and weaknesses in different areas.

The last point may be of interest for sub-communities, regional groups, and other potential collaborators, as it points to potential opportunities for collaboration where partners are likely to have complementary areas of strength, with the potential to share leading practices in different areas.
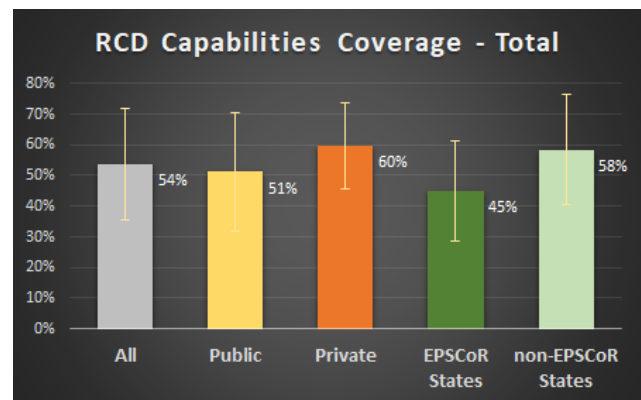


**Figure 5: Total RCD coverage by key demographics**

Figure 5 presents the summary capabilities coverage for different demographic slices. Note the significant difference between public and private institutions and between institutions in EPSCoR-eligible states vs. those in other states. While each of these might well conform to expectations, the data clearly shows that popular
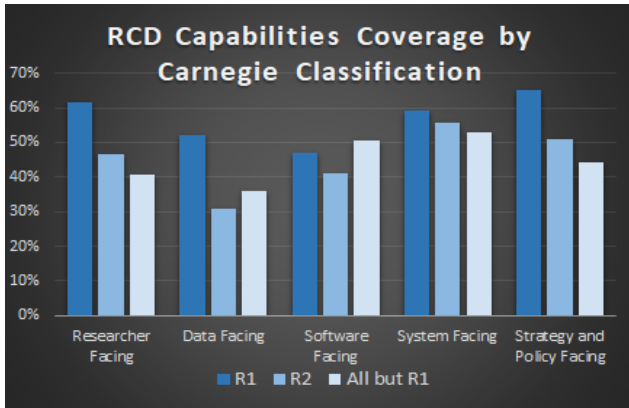
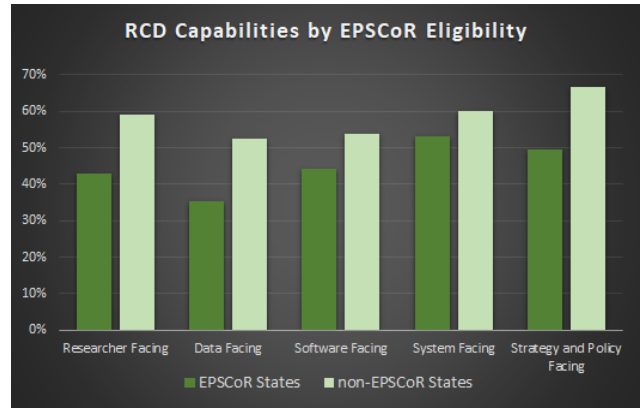**Figure 6: Median Capabilities coverage across facings by Carnegie Classification**

conceptions are borne out in experience and that these differences are often fairly substantial (although because of the relatively large variance in the data only the EPSCoR distinction is statistically significant). Looking deeper into the facings, certain areas show wide gaps between public and private institutions:

- In the Data Facing themes for *Data Analysis* and *Data Visualization*, private institutions average 25% to 30% higher coverage than public institutions, and 20% higher coverage for *Data Security/Sensitive Data Support*.
- Private institutions average 20% to 25% higher coverage of capabilities in the Strategy and Policy Facing themes of *Institutional Culture for Research Support; Funding*; and *Partnerships / Engagement with External Communities*. The median values for private institutions in these same themes are from 25% to 44% higher than for public institutions.

In Figure 6 we can see significant variation in certain facing areas when we filter the data by the Carnegie Classification of the institutions. Institutions are actually fairly comparable in the System Facing and Software capabilities, perhaps reflecting the longer traditions and understanding of requirements and good practices for systems definition, administration, and maintenance, and of software management. However, there is considerable variation in capabilities coverage in Researcher Facing, Data Facing, and Strategy and Policy Facing areas, where roles and good practices have more recently emerged and/or are rapidly evolving.

A similar pattern emerges for the EPSCoR vs. non-EPSCoR institutions (Figure 7), but with even starker differences. There is again somewhat closer parity in System Facing capabilities, but the EPSCoR-eligible institutions experience wide gaps across Researcher Facing, Data Facing, and Strategy and Policy Facing areas. The data also show gaps in the areas of RCD Staffing and RCD Outreach, in all areas of Data Facing capability, and especially in *Data Security/Sensitive Data Support* where EPSCoR-eligible institutions have a median value roughly one-fourth that of other institutions (Figure 9). EPSCoR-eligible institutions also show much lower assessed coverage values in the areas of *Institutional Culture for Research Support* and *Diversity, Equity, and Inclusion*.



**Figure 7: Median Capabilities coverage across facings by EPSCoR eligibility**

Minority-serving institutions (Figure 8) show a similar stark pattern of gaps relative to institutions that are not minority-serving. These gaps exist across the spectrum of facings, with significant differences in some of the same areas described for the other demographic comparisons, above. Notably, the median values for themes in the System Facing and Software Facing areas are much closer to parity than the averages are, indicating that some Minority serving institutions are facing even greater gaps than many of their peers (this is echoed by the standard deviations for these Facings, which are a significant proportion of the average values).
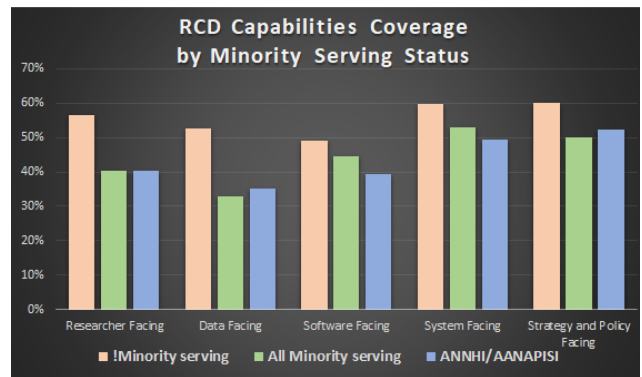


**Figure 8: Median Capabilities coverage across facings by Minority Serving status**

## 5 SIGNIFICANT THEMES FOR PRIORITIES

The full Community Data report [6] includes visualizations of aggregated priorities data, but we note several themes here: for the community as a whole, many of the priorities are in the Researcher Facing and Data Facing areas, and a fair number are in the Strategy and Policy Facing area, with very strong emphasis in:

- *Research Computing and Data Staffing*
- *Research Computing Management of the Research Lifecycle*
- *Data Creation*

**Table 1: Capability areas indicated as priorities for the entire community**

| | |
|---|---|
| 1 | Do researchers have access to introductory user support and training related to the use of research computing and data resources available at local, regional, and national level? |
| 2 | Are Research Computing and Data services funded in a sustainable manner? |
| 3 | Do researchers have access to consulting and expertise to help them identify appropriate data repositories? |
| 4 | Are researchers supported across the full research lifecycle? |
| 5 | Does your Research Computing and Data (RCD) team/group have a strategic plan? |
| 6 | To what extent is there a clear vision, effective guidance, and strategy for the allocation and prioritization of support resources/personnel? |
| 7 | Do researchers have access to advice on research compute and data compliance, security, management, and governance? |
| 8 | Do researchers have access to tools/software that supports data backup, storage, and integrity checking? |
| 9 | Does your institution have research data governance processes in place to establish data policies for research data? |
| 10 | Do researcher-facing staff have the skills and capacity to broadly support researchers across levels (graduate students to PIs) and across domains with information about the use and effectiveness of new technologies? |
| 11 | Do researchers have access to consulting and expertise on data wrangling/manipulation and data analysis? |

...and relatively strong emphasis in these areas as well:

- *Data Analysis*
- *Research Data Curation, Storage, Backup, and Transfer*
- *Research Data Policy Compliance*
- *Software Portability, Containers, and Cloud Computing*
- *Best security practices for open environments*
- *Institutional Alignment*

Across the full community, the most highly prioritized capabilities (i.e., most commonly marked as a priority) are listed in Table 1. The list includes an interesting mix of consulting and engagement support for researchers, funding and strategic planning, and several data governance and compliance items. When we filter for particular sub-communities, there are many overlaps but also some interesting differences:

**Private Institutions**: Only 5 of the top priorities for private institutions are shared with those for the broader community (numbers 1, 2, 6, 9, and 10); the 5 additional priorities (all of which are at the same level) include:

- *Compute and data environments for sensitive/secure data*
- *Support for Software compilation, software good practices*
- *Institutional support for cloud services for research*
- *Staff skills for containers, orchestration*
- *Support for interactive computing (VDI, Jupyter, etc.)*

**R2 Institutions**: The top priorities among R2 institutions vary widely from those of the community as a whole, sharing only numbers 2, 4, 5 and 7 (and even these are ranked somewhat differently). Of the additional priorities, two are common with the top priorities for private institutions:

- *Institutional support for cloud services for research*
- *Support for interactive computing (VDI, Jupyter, etc.)*

Another 5 priorities are distinctive to the R2 institutions, and are listed here in priority order. It is notable that HPC is ranked so high, and that several are associated with research data management:

- *Do researchers have access to a place to store final research data...?*
- *Do researchers have access to high performance (batch) computing (HPC)?*
- *Do researchers have access to policies and technologies that facilitate management and wide access to data?*
- *Is your Research Computing and Data (RCD) strategic plan aligned to campus plans?*
- *Do research funding activities actively integrate the Research Computing and Data (RCD) services group?*

**EPSCoR-eligible** and **Minority-Serving Institutions** had as top priorities a focus on strategy and funding:

1. *Does your Research Computing and Data (RCD) team/group have a strategic plan?*
2. *Are Research Computing and Data services funded in a sustainable manner?*
3. *To what extent is there a clear vision, effective guidance, and strategy for the allocation and prioritization of support resources/personnel?*

## 6 CONCLUSIONS AND LOOKING AHEAD

We have presented an analysis of the first Community Dataset that aggregates 41 institutional assessments using the Research Computing and Data Capabilities Model. The 2020 Community Dataset represents a diverse set of institutions and provides significant insights into the state of support programs for RCD. The Dataset provides an important complement to the model itself, allowing institutions to understand their relationship to the broader community, and providing various entities (including, e.g., funders)
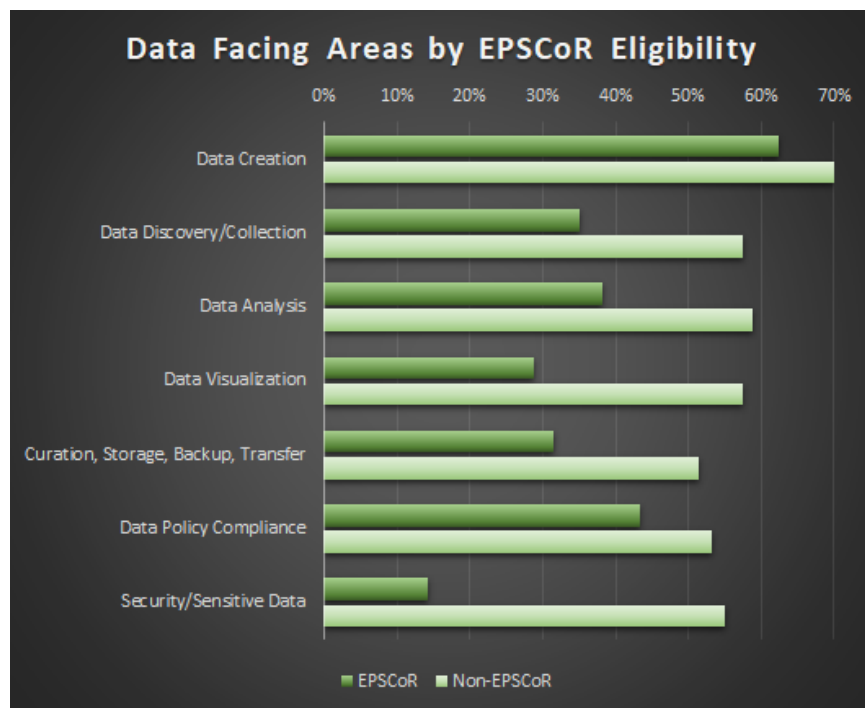
**Figure 9: Median Data Facing Capabilities coverage by EPSCoR eligibility**

with the data to characterize RCD at a fine-grained level, and over time, to follow trends and track the impact of programs designed to advance RCD support. While the data often confirm expected patterns, they also make clear that differences between certain groups within the community are even more profound than many may have expected. The variation of institutional strengths (and weaknesses) across different areas of capability may present opportunities for collaboration, and as the RCD CM model itself provides a structured vocabulary that allows institutions to more clearly communicate about RCD support, the Community Dataset now provides granular insights to groups of institutions that are seeking to collaborate on shared solutions and strategies to advance RCD support. In discussion of these early results with some community members, institutional leaders confirmed that the RCD CM will make it easier to compare programs and practices, to understand where collaborations can be of benefit, and to identify opportunities for joint projects, funding proposals, etc. In a survey of participants, 100% planned to use the results for strategic planning, over 80% planned to use the report and their benchmarking results to justify campus funding for their program, and over a third planned to use the reports to support a grant proposal.

The 2020 Community Dataset provides an initial snapshot of RCD support, and we will repeat this annually. Over time, longitudinal data will provide additional insight into trends, and a means of evaluating the impact of programs designed to increase RCD support for institutions, for collaborations and sub-communities, and for the community as a whole. At a community level we are seeing interest in understanding the broad state of RCD support and in working collaboratively to advance such support (e.g., a Decadal

Survey currently in planning at CaRCC; https://carcc.org/decadal-survey/). A shared repository of longitudinal assessment data will provide essential baseline data for such an effort, as well as a framework for describing the broad scope of RCD support (and a means of evaluating the impact of the decadal survey itself).

## 6.1 Refining the Assessment Tool and the Data Analysis Platform

The process of completing our analysis of the 2020 RCD Community Dataset has underscored the need for a more robust data analysis platform. The RCD CM Working Group has documented plans to re-implement the RCD CM Assessment Tool on a more robust and functional survey platform. Some of the additional functionality that is under discussion for data portal functionality includes:

- Support for interactive exploration of the data with filtering by broad demographic categories.
- Support to tag institutions with consortia and regional communities and then to filter on these tags.
- Support for geographic filters by regions, states, etc.
- Support to visualize data for a given year and over multiple years (to explore trends and evolution).

We believe that this new functionality will provide additional value to institutions and to the community, and will further motivate contributions of institutional assessment data.

## ACKNOWLEDGMENTS

Advancing Research and Education through a national network of campus research computing infrastructures – The CaRC Consortium").

The Model, the assessment tool, and other associated resources were developed with the generous contributions of time and expertise from the 2018 workshop participants, and the working group members: Alex Feltus, Ana Hunsinger, Cathy Chaplin, Claire Mizumoto, Dana Brunson, Deborah Dent, Doug Jennewein, Gail Krovitz, Galen Collier, Jackie Milhans, James Deaton, Jen Leasure, Jill Gemmill, Jim Bottum, Joe Breen, Joel Cutcher-Gershenfeld, John Hicks, John Moore, Karen Wetzel, Mike Erickson, Patrick Schmitz, Preston Smith, Timothy Middelkoop, and Thomas Cheatham.

In addition, individuals at a number of Universities provided valuable feedback on early versions of the model and assessment tool, for which the working group is very grateful. The 2020 Community Data report upon which this is based benefited from input and review by Claire Mizumoto, Dana Brunson, Doug Jennewein, Galen Collier, Joel Cutcher-Gershenfeld, John Hicks, and Thomas Cheatham.

## REFERENCES

[1] Campus Research Computing Consortium. [n.d.]. . http://carcc.org
[2] EDUCAUSE. [n.d.]. . https://www.educause.edu/
[3] Indiana University Center for Postsecondary Research. [n.d.]. *The Carnegie Classification of Institutions of Higher Education, 2018 edition.* https://carnegieclassifications.iu.edu/
[4] CaRCC RCD Capabilities Model Working Group. 2020. *Capabilities Model Introduction and Guide to Use.* Retrieved May 24, 2021 from https://docs.google.com/document/d/15xiDXMta7AlEvE6IpW4mvadAiW2PPshmBi73AVHTm9g
[5] Internet2. [n.d.]. . https://internet2.edu/community/research-engagement/
[6] Patrick Schmitz. 2020. 2020 RCD CM Community Data report. https://doi.org/10.5281/zenodo.4344057
[7] Patrick Schmitz, Claire Mizumoto, John Hicks, Dana Brunson, Gail Krovitz, James Bottum, Joel Cutcher-Gershenfeld, Karen Wetzel, and Thomas Cheatham. 2020. A Research Computing and Data Capabilities Model for Strategic Decision-Making. In *Practice and Experience in Advanced Research Computing* (Portland, OR, USA) *(PEARC '20).* Association for Computing Machinery, New York, NY, USA, 77–84. https://doi.org/10.1145/3311790.3396643