



# Identifying high-performance computing needs at member institutions in a regional consortium

Becky Yeager\*

Shelley Knuth

Alana Romanella

becky.yeager@colorado.edu

shelley.knuth@colorado.edu

alana.romanella@colorado.edu

University of Colorado, Boulder

Boulder, Colorado, USA

Chris Reidy

chrisreidy@arizona.edu

University of Arizona

Tucson, Arizona, USA

Ben Nickell

ben.nickell@inl.gov

Idaho National Laboratory

Idaho Falls, Idaho, USA

Jason Yalim\*

yalim@asu.edu

Arizona State University

Tempe, Arizona, USA

## ABSTRACT

The Rocky Mountain Advanced Computing Consortium (RMACC) is a collaborative organization of academic, government, and professional research institutions located throughout nine inter-mountain states. RMACC's mission is to facilitate the efficient adoption of high-performance computing (HPC) resources within the region. During the Summer of 2023, the Executive Committee of RMACC conducted thirty-minute interviews with the majority of its thirty-four member institutions. As a result of these interviews, several opportunities and challenges in regional HPC practices were identified. Our findings indicate the importance of strong connections between institutions to enhance collaboration and support Research Computing regionally. Specific themes indicate increased knowledge sharing, additional networking opportunities, and opportunities for workforce development to aid in technical staff retention and student growth. A need for a variety of training was specifically outlined, as well as additional support for systems professionals who may be the only (or one of few) supporting a cluster at their institution.

## CCS CONCEPTS

• **Human-centered computing** → *Collaborative and social computing*; • **Social and professional topics** → *Computing / technology policy*; **Computing organizations**.

## KEYWORDS

HPC Consortia, Democratization, Workforce Development

## ACM Reference Format:

Becky Yeager, Shelley Knuth, Alana Romanella, Ben Nickell, Chris Reidy, and Jason Yalim. 2024. Identifying high-performance computing needs at member institutions in a regional consortium. In *Practice and Experience in Advanced Research Computing (PEARC '24)*, July 21–25, 2024, Providence, RI, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3626203.3670602>

## 1 INTRODUCTION AND BACKGROUND

The Rocky Mountain Advanced Computing Consortium (RMACC) was founded in 2011 by Dr. Thomas Hauser, then Director of Research Computing at The University of Colorado (CU) Boulder [18]. In the Fall of 2010 leadership at Front Range educational institutions and National Labs in Colorado and Wyoming met to establish a committee to pursue regional HPC collaboration opportunities. RMACC was officially established in 2014 with expansion out to Utah, Montana, Idaho and New Mexico, and grew from the success of the original Front Range Consortium for Research Computing (FRCRC) which comprised Colorado State University (CSU), CU Boulder, CO Sch. of Mines, U. of Wyoming, the National Center for Atmospheric Research (NCAR), the National Renewable Energy Laboratory (NREL), and the National Oceanic and Atmospheric Administration (NOAA). In 2018 RMACC expanded to include schools in Nevada, Washington, and Arizona.

RMACC has been successful in acquiring funding to support under-resourced regional institutions and expand opportunities for individuals. At the time of this writing, there have been six direct National Science Foundation (NSF) awards directly supporting RMACC totaling just over \$6 million, with NSF numbers: 1532235, 1532236, 1659425, 1925766, and including two currently active NSF awards, 2201538 & 2322260. The first two grants supported RMACC's Summit supercomputer, which provided 85-million core-hours per year to RMACC members [1]. The third resulted in the creation of the RMACC "Cyberteam" in 2017 to provide cyberinfrastructure support for members [16, 17].

Despite the benefits that they provide, regional HPC Consortia are not universal. In 2009, Gropp & Snir [11] called for the creation of an HPC consortium to improve the efficiency and software quality of HPC systems, discussing unaddressed gaps in an



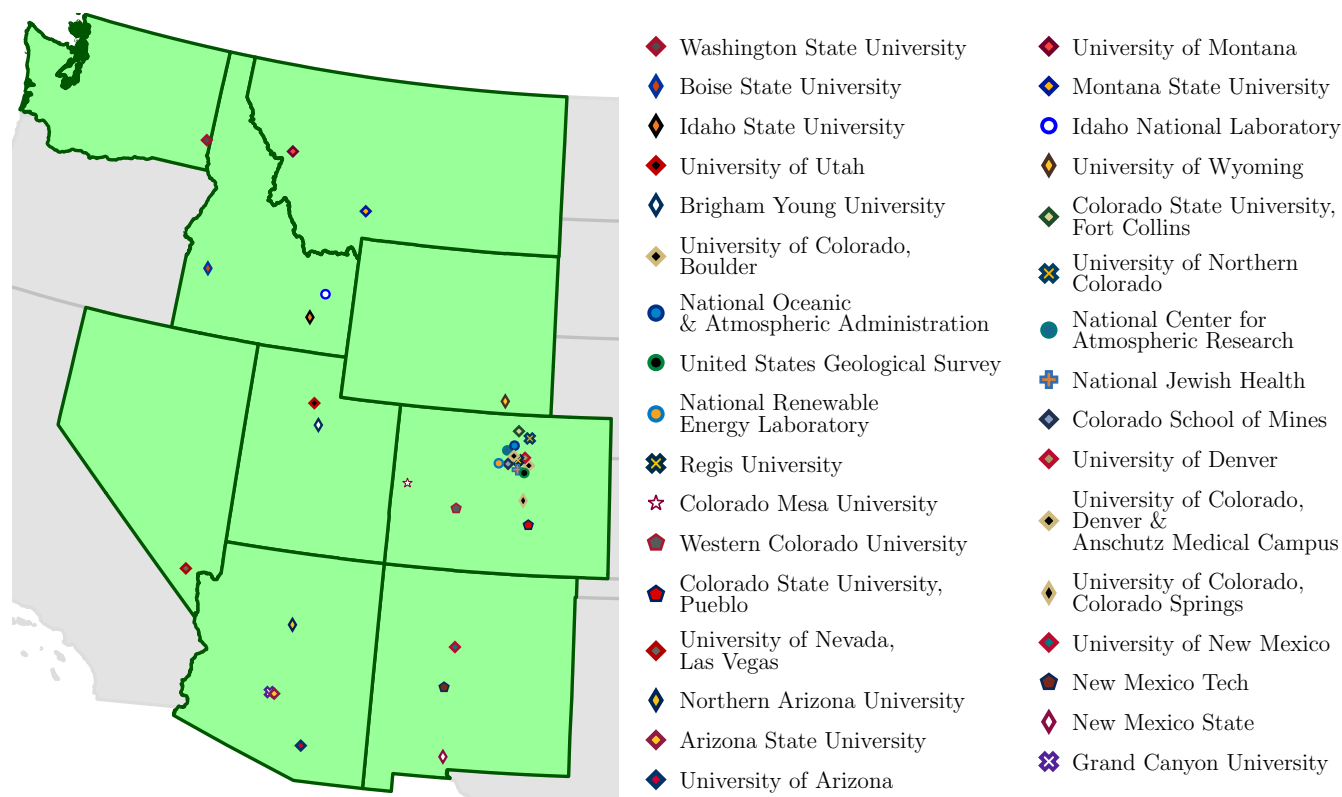
This work is licensed under a Creative Commons Attribution International 4.0 License.

PEARC '24, July 21–25, 2024, Providence, RI, USA

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0419-2/24/07

<https://doi.org/10.1145/3626203.3670602>



**Figure 1: Map of the western continental United States with RMACC’s thirty-four members’ host states highlighted in green. Markers indicate the locations of participating RMACC institutions. Marker types indicate institution type, i.e., universities that are R1 (square diamonds), R2 (rhombuses), D/PU (filled-X), M2 or M3 (pentagons), else (stars); the hospital (filled plus); the government entities (circles). As illustrated, the five Denver, three Boulder, and two Golden locations in Colorado are shifted slightly from respective city centers to improve visibility. Note: University of Colorado Denver and Anschutz Medical Campus are distinct RMACC members.**

“[HPC] performance pyramid.” The authors specifically wanted an institutional-sharing structure more focused than the existing nationwide consortia (like TeraGrid) that could facilitate operational and end-user practice sharing, as well as increase collaboration and increase software homogeneity across centers. There are a number of U.S.-scale consortia, like Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS), which replaced the Extreme Science and Engineering Discovery Environment (XSEDE), which had replaced TeraGrid. ACCESS is most comparable to Compute Canada or the international PaRtnership for Advanced Computing in Europe (PRACE) [3]. In the United States, the National HPC Consortium was founded for fighting COVID-19 [5, 13], demonstrating how networks of HPC centers can improve the state of the art and decrease the time to science or, especially in the case of COVID, help direct policy. There are a number of regional networks that are more comparable to RMACC. These include but are not limited to SWEETER, the NE Cyberteam, the Great Plains Network, and the ECEP [2, 4, 6–10, 12, 19, 20]. In order to operate efficiently, consortia member needs must be regularly assessed.

At the time of this writing, RMACC has thirty-four member institutions. Figure 1 illustrates where these member institutions are with markers that reflect the 2021 Carnegie classifications [15] of academic members, or if a member is a hospital or Federal Government entity. Twenty-eight of RMACC’s members are academic, and by the shorthand defined in the 2021 “Basic Classification” Carnegie classifications [14]: fourteen<sup>1</sup> are R1, seven are R2, three are D/PU, two are M2, one is M3, and one is ranked as a Baccalaureate College with Diverse Fields. Five of RMACC’s members are Government entities: NREL, NOAA, USGS, NCAR/UCAR, and INL. One of RMACC’s members is a research hospital.

RMACC’s member institutions are often represented by their expert HPC facilitators and researchers. To determine consortia members’ needs, one-on-one, video-conferenced interviews with those representatives were organized by RMACC’s Executive Committee. The results from those interviews and the resulting discussion of the Executive Committee are qualitatively reported by this short paper.

<sup>1</sup>The University of Colorado Anschutz Medical Campus is considered a distinct RMACC member, but is combined with the University of Colorado Denver in the Carnegie classifications.

## 1.1 RMACC Structure and Benefits

RMACC is a free collaborative network for computational researchers and HPC facilitators to draw from and contribute to. The organization maintains an enterprise Slack channel as a forum for member communication. Funding typically comes from sponsorship and registration for an annual HPC Symposium which is historically hosted in May by CU Boulder. In 2023 it was hosted by Arizona State University (ASU); the host is planned to start rotating annually among member sites by 2025. The HPC Symposium has typically been a two-and-a-half day event, primarily focused on encouraging presentations, discussions, and collaborations between system administrators, facilitators, researchers, and industry sponsors. Attendees often will present site-updates and discuss operational challenges and opportunities with one another. The Symposium also encourages student participation by hosting a Student Poster competition. The winner is awarded full financial support to attend a larger global conference, typically SuperComputing.

An annual system administrator meetup is organized by RMACC. It has typically been hosted in August–October at a rotating host site, e.g., it was hosted by CSU in Fort Collins in 2023 over two-and-a-half days in late September.

RMACC is led by an Executive Committee, which is currently comprised of six members. Community-wide elections select the Chair, Vice-Chair, Executive Director, and three Board members by simple majority. Elections take place every two to three years, the timing of which balances the effort involved in conducting elections and maintaining effective community representation. The Executive Committee meets at least monthly to discuss and act upon the operation and needs of the Consortium.

RMACC has three major subcommittees that focus on systems, users, and Diversity, Equity, and Inclusion (DEI) topics. The three subcommittee meetings occur at least monthly and are open to the broader community. Ad hoc subcommittees will form and disband to organize and support the annual HPC Symposium and system administrator meetup.

## 2 MEMBER INTERVIEWS

Over August–October of 2023, Executive Director Becky Yeager met with representatives from thirty-one of the thirty-four RMACC members. The Executive Director was typically joined by one other RMACC Executive Committee member to help facilitate the interview process. Member institution representatives were typically HPC center leaders, system administrators, or facilitators (see table 1) — those who are the primary contacts for RMACC communications and often participate in RMACC's events. Each member was asked the same set of questions, which are highlighted in figure 2.

Additionally, two interviews were conducted for the RMACC DEI subcommittee by Executive Director Becky Yeager. These interviews were conducted in a one-hour meeting and covered the same questions as the member interviews with time to delve deeper into the answers and more complex questions.

Interview notes were collected in a shared Executive Committee document and further discussed by the Executive Committee during regular meetings. The qualitative analysis and inference from those discussions form the basis of this report.

| role                  | #    |
|-----------------------|------|
| HPC Director or CIO   | 14+1 |
| System Administrators | 14   |
| User Facilitators     | 8    |
| Multiple Roles        | 6+1  |
| Faculty               | 5    |
| Other Role            | 4    |
| Researcher            | 2    |
| Student               | 1    |
| Total                 | 54+2 |

**Table 1: Interviewee roles from the thirty-one conducted member interviews. The additional two interviewees were from a set of two interviews conducted for the RMACC DEI subcommittee by Executive Director Becky Yeager.**

- (1) Are the contacts in today's discussion appropriate?
- (2) Which individuals should be added to the recipient list for RMACC emails?
- (3) What is your or your institution's familiarity with RMACC offerings? That is, with the:
  - (a) RMACC HPC Symposium,
  - (b) system administrator meetup,
  - (c) three subcommittees on system administration, user support, and DEI,
  - (d) available shared resources,
  - (e) trainings,
  - (f) and the ACCESS Affinity group?
- (4) What is your or your institution's current involvement with RMACC?
- (5) What would enhance your institution's level of participation?
- (6) What specific expectations do you or your institution have from RMACC? What would you or your institution like to see offered in the future?

**Figure 2: List of interview questions, which were asked to each participating member.**

## 2.1 Interview Findings

Nineteen member institutions reported having on-premise computational resources, with eleven members relying on RMACC or National (ACCESS) resources for research computing needs.

There were a number of overarching themes and requests. The interviewees all communicated a need for more:

- (1) training for the users of member HPC systems as well as for the faculty and staff supporting those users,

- (2) knowledge sharing between RMACC partners about topics of interest to their staff, e.g., effective grant writing, or their users, e.g., application affinity groups,
- (3) networking opportunities and chances to gather with their regional peers to discuss topics of interest,
- (4) hardware resource sharing within the region,
- (5) and an improved ability to discover available resources both regionally and nationally.

As a result of the interviews, several immediate actions were taken. To improve inter-institutional transparency and facilitate collaboration, a Google spreadsheet was shared among member institutions. This spreadsheet had two subsheets that members were asked to fill out. The first sheet was for self-reporting site information, collaboration goals, and RMACC-sharable resources (e.g., the University of Colorado Boulder's Alpine). The second sheet collected information on member institutions' flagship system, storage options, personnel, and service offerings, such as cloud or secure compute.

To help facilitate outreach, networking, and educational opportunities for system administrators, the existing sysadmin group will be brought more formally under the RMACC umbrella. This includes ensuring that the main HPC symposium has a dedicated sysadmin track for its entire duration. Additionally, the group's monthly meeting will be more formalized with pre-cultivated conversation topics that encourage those facilitation goals. The other RMACC communities will also adopt this structure. Additionally, opportunities for the community to gather at external conferences, i.e., SuperComputing and PEARC, will be organized.

To address member interests, a more regular cadence of community conversations will be developed, on topics including grants, ACCESS, and available resources for RMACC members. Opportunities for member lightning talks will be provided.

A committee was created to develop a list of more "generic" trainings that are not institution or machine-specific and can be offered remotely or asynchronously to enable attendance by the RMACC community. One of the synchronous goals is to create a calendar of shared trainings on the RMACC website, as well as a collated list of shared training resources on the RMACC web site. This list includes documentation, training videos, and other available resources from RMACC members. This list will be made available to the ACCESS Knowledge Base.

A plan is in place to meet with smaller RMACC member institutions to learn about their needs regarding the use of the RMACC partition on the Alpine Supercomputer and develop specific trainings from the University of Colorado Boulder that meet their specific needs and questions.

Finally, in part due the interviews conducted by RMACC, the ACCESS project has put out a call to fund community-led workshops on workforce development.

### 3 DISCUSSION AND NEXT STEPS

There is a training gap on how to administer new systems. Under-resourced institutions would especially benefit from better information sharing on how to obtain funding for new systems, as well as how to train personnel, set-up, administer, and support users

on said system. Topics range from, for example, working with a scheduler or debugging MPI on multi-node systems.

Correspondingly, there are higher-education training gaps for students interested in HPC careers, including a lack of clear degree programs paths or the courses that appropriately foster expertise, like in systems administration or user support. Many of the smaller institutions need workers who have the ability to support users and conduct system administration. This is driven by a general lack of personnel as well as funding for additional hires. Because of the lack of training and workforce development the number of trained administrators does not meet the regional demand and need for trained administrators. These roles are being filled by faculty and staff who do not have the technical expertise needed to fill the gaps, resulting in split time between research, teaching classes, administering and building systems, and supporting others in their research without formal training in those areas.

Training opportunities are further challenged by DEI issues; there are no quick pathways to a successful career. Issues of equity and inclusion in rural areas are a challenge as other career paths are more visible and readily available for students. With a lack of training and workforce development many students are not aware of the opportunities that exist.

Remote areas have additional difficulty in recruiting talent. With a small pool of available HPC-trained workers and open positions at larger R1 institutions available it is difficult to find employees who are interested and willing to live in remote areas. Smaller institutions find it difficult to compete for the available pool of workers because of challenges with pay scales, location, and skill set needed.

Finally, it is difficult to find resources to learn on. While there are many national resources available for researchers and users, there is a general gap in the communication of the existence of those resources and how to access them. Additionally, when training materials do exist, often they are provided at an expert level, making assumptions that new users will have a basic understanding of a command-line interface, which is not always the case. Training to both find the resources and use them, would benefit these schools that do not have their own trainers or user documentation.

Taken together, many institutions are limited in their ability to move up the pyramidal HPC hierarchy of needs, but directed organizational efforts, particularly with funding, can increase the research efficiency and agency of smaller member institutions in ways that may be replicated to the wider national community. While RMACC may not be able to solve all of these issues, it has spurred a number of ideas that are in the early stages of development.

One such effort involves working with the ACCESS project to decentralize their outreach efforts. In part due to these findings, ACCESS is in the late stages of developing a campaign to support individuals' attendance at conferences in return for assistance with developing materials on their resources. This will allow the material provided on ACCESS resources and activities to be provided in community members' own words, which will enhance knowledge sharing. Several RMACC partners are also in the very early stages of developing a plan for a program that will provide both formal and informal training in HPC systems, storage, and network administration. Additionally, several RMACC members have partnered together as part of OAC grant no. 2322260 to form a cohort

of student undergraduates who are working closely with system administrators employed at CU Boulder to engage in hands-on work to design, procure, and deploy new hardware for the Alpine cluster.

## 4 CONCLUSION

Our findings highlight a need within the RMACC community to bolster collaboration and support among member institutions. By implementing strategies such as information exchange opportunities, formalizing sysadmin groups, and organizing regular meetings with curated topics, the RMACC community aims to foster networking, knowledge-sharing, and educational opportunities. As a regional organization, initiatives such as remote training opportunities and tailored support for smaller institutions demonstrate a commitment to inclusivity and addressing the diverse needs of the region while also promoting the regional adoption of HPC/CI. Moving forward, these efforts are poised to strengthen the relationships between RMACC members, promote innovation, and enhance the overall effectiveness of collaborative endeavors within the research computing landscape.

## ACKNOWLEDGMENTS

This work is supported by the National Science Foundation under Grant Numbers 2138286 and 2322260. This research was also supported by Idaho National Laboratory's High Performance Computing located at the Collaborative Computing Center and supported by the Office of Nuclear Energy of the U.S. Department of Energy and the Nuclear Science User Facilities under Contract No. DE-AC07-05ID14517.

## REFERENCES

- [1] Jonathon Anderson, Patrick J. Burns, Daniel Milroy, Peter Ruprecht, Thomas Hauser, and Howard Jay Siegel. 2017. Deploying RMACC Summit: An HPC Resource for the Rocky Mountain Region. In *Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact* (New Orleans, LA, USA) (PEARC '17). Association for Computing Machinery, New York, NY, USA, Article 8, 7 pages. <https://doi.org/10.1145/3093338.3093379>
- [2] Amy W. Apon, Gregory E. Monaco, and Gordon K. Springer. 2006. The Great Plains Network (GPN) middleware test bed. *Scalable Computing: Practice and Experience* 7, 3 (2006), 95–108.
- [3] Susan Baldwin. 2012. Compute Canada: Advancing Computational Research, In High Performance Computing Symposium 2011. *Journal of Physics: Conference Series* 341, 012001. Issue 1. <https://doi.org/10.1088/1742-6596/341/1/012001>
- [4] Smriti Bhatt and Izzat Alsmadi. 2021. *Science DMZ and NSF CC\* SWEETER Workshop Report*. Technical Report. SWEETER, Texas A&M University - San Antonio.
- [5] J. Brase, N. Campbell, B. Helland, T. Hoang, M. Parashar, M. Rosenfield, J. Sexton, and J. Towns. 2022. The COVID-19 High-Performance Computing Consortium. *Computing in Science & Engineering* 24, 1 (2022), 78–85. <https://doi.org/10.1109/MCSE.2022.3145608>
- [6] Northeast Cyberteam. 2024. <https://necyberteam.org/>. Last accessed: April, 2024.
- [7] Expanding Computing Education Pathways (ECEP) Alliance. 2024. <https://ecepalliance.org>. Last accessed: April, 2024.
- [8] John Goodhue, Julie Ma, Adrian Del Maestro, Sia Najafi, Bruce Segee, Scott Valcourt, and Ralph Zottola. 2020. Northeast Cyberteam: workforce development for research computing at small and mid-sized institutions. In *Practice and Experience in Advanced Research Computing* (Portland, OR, USA) (PEARC '20). Association for Computing Machinery, New York, NY, USA, 402–406. <https://doi.org/10.1145/3311790.3396662>
- [9] John Goodhue, Julie Ma, Adrian Del Masetro, Sia Najafi, Bruce Segee, Scott Valcourt, and Ralph Zottola. 2020. Northeast Cyberteam - building an environment for sharing best practices and solutions for research computing. In *2020 IEEE High Performance Extreme Computing Conference (HPEC)*. 1–5. <https://doi.org/10.1109/HPEC43674.2020.9286254>
- [10] Great Plains Network. 2024. <https://greatplains.net>. Last accessed: April, 2024.
- [11] William Gropp and Marc Snir. 2009. On the Need for a Consortium of Capability Centers. *The International Journal of High Performance Computing Applications* 23, 4 (2009), 413–420. <https://doi.org/10.1177/1094342009347706>
- [12] Mark Guzdial, Barbara J. Ericson, Tom McKlin, and Shelly Engelman. 2012. A statewide survey on computing education pathways and influences: factors in broadening participation in computing. In *Proceedings of the Ninth Annual International Conference on International Computing Education Research* (Auckland, New Zealand) (ICER '12). Association for Computing Machinery, New York, NY, USA, 143–150. <https://doi.org/10.1145/2361276.2361304>
- [13] James J. Hack and Michael E. Papka. 2020. The U.S. High-Performance Computing Consortium in the Fight Against COVID-19. *Computing in Science & Engineering* 22, 6 (2020), 75–80. <https://doi.org/10.1109/MCSE.2020.3019744>
- [14] Indiana University Center for Postsecondary Research. 2021. Carnegie Classifications > Classification Methodology > Basic Classification. <https://carnegieclassifications.acenet.edu/carnegie-classification/classification-methodology/basic-classification/>. Last accessed: April, 2024.
- [15] Indiana University Center for Postsecondary Research. 2021. The Carnegie Classification of Institutions of Higher Education, 2021 edition. <https://carnegieclassifications.acenet.edu/>.
- [16] Tobin Magle, Brett Milash, Andrew J. Monaghan, Thomas Hauser, Thomas E. Cheatham, Patrick J. Burns, Anita M. Orendt, Shelley Knuth, and Andrew M. Johnson. 2019. RMACC Cyberteam Surveys. <https://doi.org/10.17605/OSF.IO/FDK43>
- [17] Andrew J. Monaghan, Brett Milash, Tobin Magle, Thomas Hauser, Thomas E. Cheatham, Patrick J. Burns, Shelley Knuth, Anita M. Orendt, Suzi White, Dawn Paschal, H. J. Siegel, Timothy Kaiser, Andrew M. Johnson, and Becky Yeager. 2019. Improving regional cyberinfrastructure services through collaboration: Cyberteam for the Rocky Mountain Advanced Computing Consortium. In *Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (Learning)* (Chicago, IL, USA) (PEARC '19). Association for Computing Machinery, New York, NY, USA, Article 41, 4 pages. <https://doi.org/10.1145/3332186.3332232>
- [18] Rocky Mountain Advanced Computing Consortium (RMACC). 2024. <https://rmacc.org/>. Last accessed: June, 2024.
- [19] Andrew Sherman, John Goodhue, Julie Ma, Kaylea Nelson, Eric Brown, Christopher Carothers, Galen Collier, Adrian Del Maestro, Andrea Elledge, Wayne Figurelle, John Huffman, Gaurav Khanna, Neil McGlohon, Sia Najafi, Jeff Nucciarone, Anita Schwartz, Bruce Segee, Scott Valcourt, and Ralph Zottola. 2022. Leveraging Northeast Cyberteam successes to build the CAREERS Cyberteam program: initial lessons learned. *Journal Comp. Sci. Ed.* 13 (April 2022), 38–43. Issue 1. <https://doi.org/10.22369/issn.2153-4136/13/1/7>
- [20] SWEETER. 2024. <https://hprc.tamu.edu/sweeter/>. Last accessed: April, 2024.