



# CU Boulder Research Computing

Shelley L. Knuth

Assistant Vice Chancellor for Research  
Computing

# Be Boulder.



University of Colorado **Boulder**

# What's our mission?

To help our researchers be the most competitive they can be in their respective disciplines by providing computational research resources beyond the desktop.

# Five Year Strategic Plan

- Infrastructure
  - User Support
  - Data management and storage
  - Secure research data
  - Education
  - Leadership
- October 2023-September 2024



- [https://www.colorado.edu/rc/sites/default/files/attached-files/curc\\_vision\\_2021.pdf](https://www.colorado.edu/rc/sites/default/files/attached-files/curc_vision_2021.pdf)

# Highlighted Accomplishments

- October 2024, 29 people in RC proper
  - Additional four working formally/closely with our team
  - Two vacant positions
- Grew our system
- Received one new grant; two supplements
- Participation in NAIRR
- Supporting Confidential data
- RMACC student cohort
- >\$136M in research grants



# Infrastructure

- On-prem Computing
  - Alpine
  - Blanca
- Data Storage
  - PetaLibrary



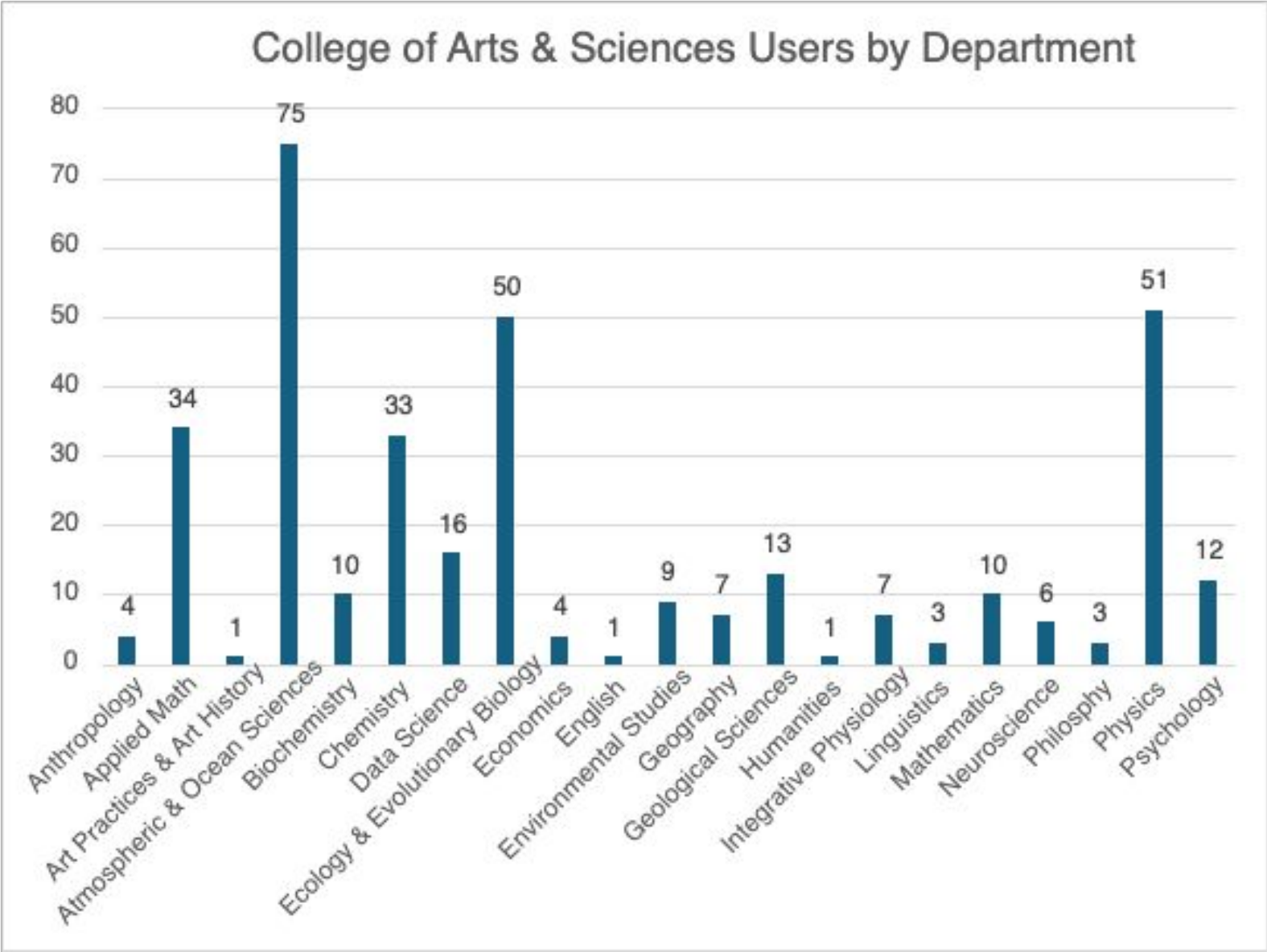
- Secure Cloud
  - The Preserve
  - AWS LZA
- Cloud Foundations Service

# Alpine Supercomputer

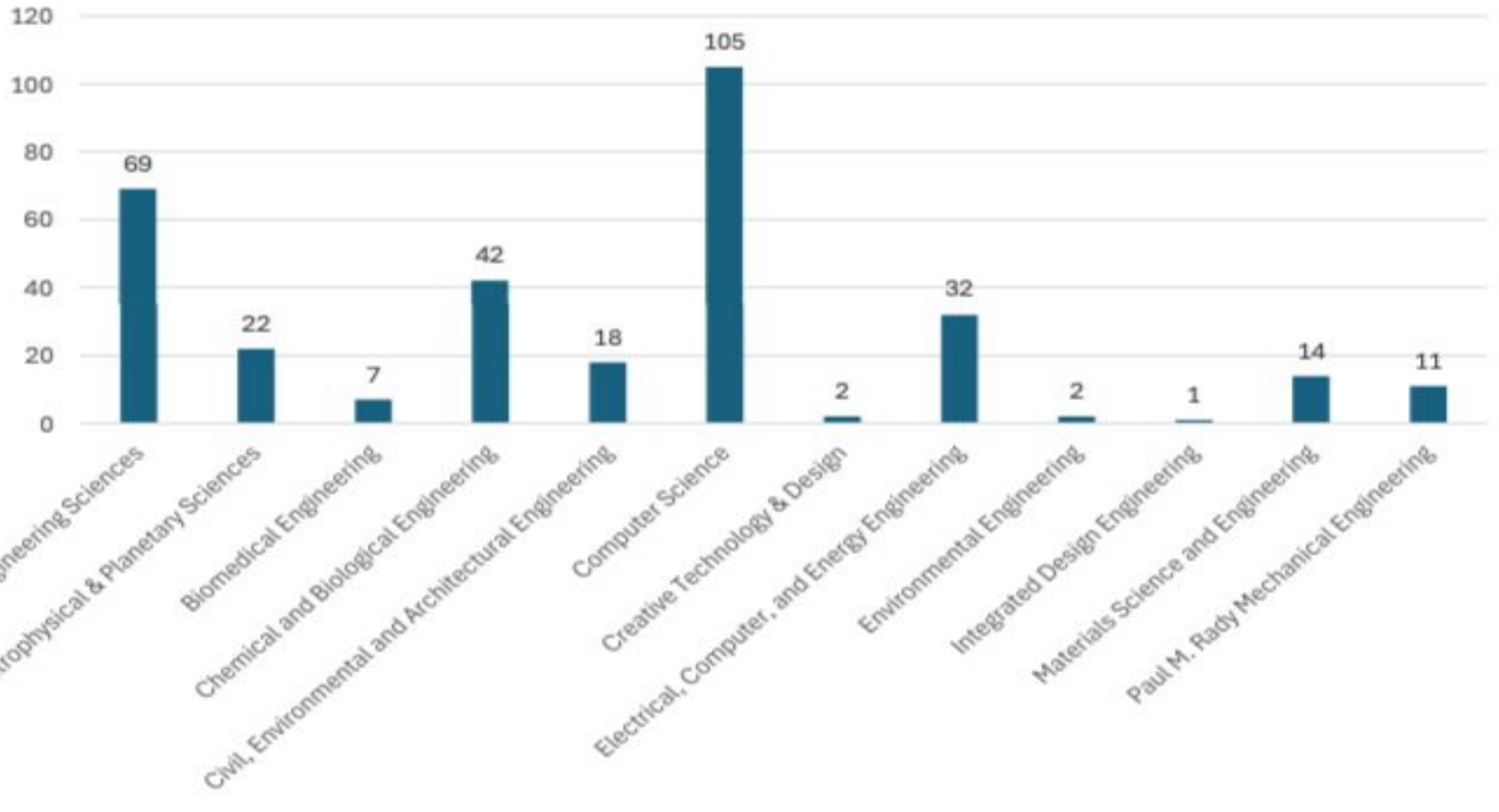
- Alpine
  - CU, CSU, Anschutz, NSF, RMACC
    - CU Denver, University of Northern Colorado, Colorado Mesa University, and New Mexico Tech
  - 389 compute nodes, 22,772 cores, 60 GPUs, 22 high-mem nodes
  - 1.9 PB GPFS scratch filesystem
  - Continued growth this year
  - 1779 users, 1043 from UCB, +28%, +21%
  - 148M core hours, 95M for UCB, +25%, +4%
  - >\$136.8M in research grants, +35%



# Who are we supporting?

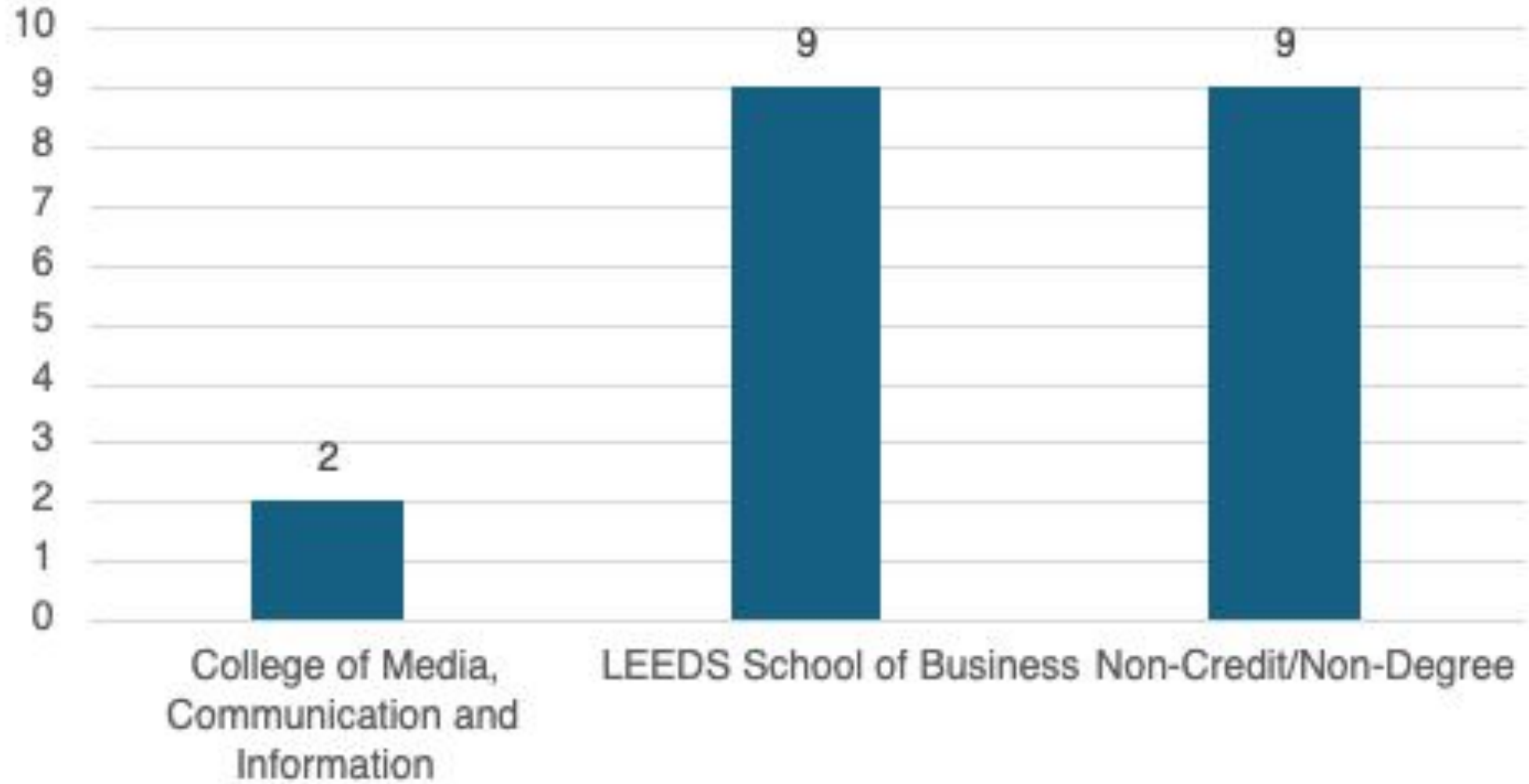


## College of Engineering and Applied Science Users by Department

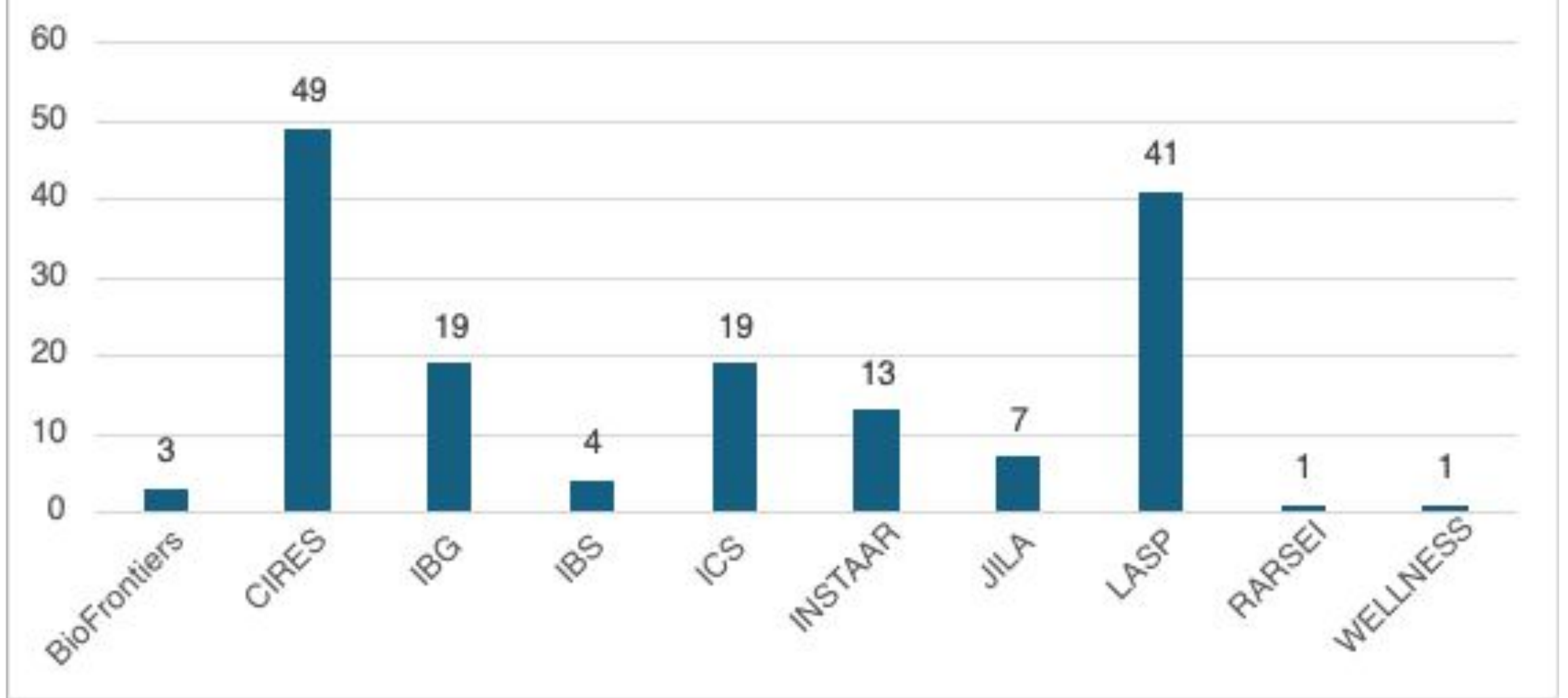




# Other Colleges and Programs



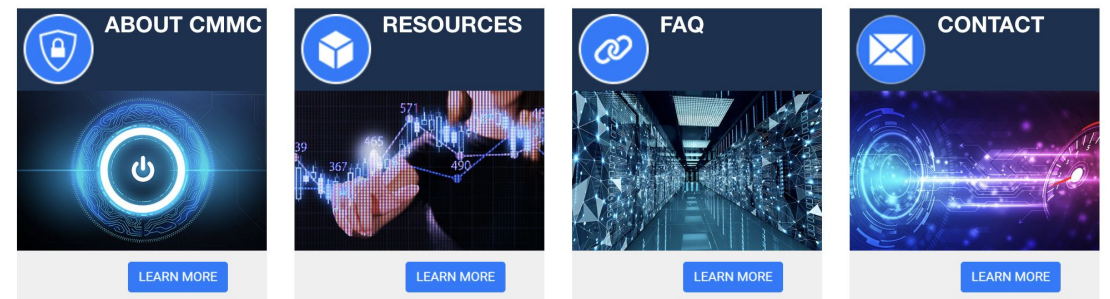
## CU Boulder Institute Users



# Secure Research Data

- The Preserve - Environment for researchers who have projects that require use of highly confidential data
  - Cybersecurity Maturity Model Certification (CMMC) 2.0, Level 2
  - Controlled Unclassified Information (CUI) regulations
- AWS Landing Zone
  - Supports data that requires CU Boulder Confidential status
- Alpine

## CYBERSECURITY MATURITY MODEL CERTIFICATION



<https://dodcio.defense.gov/CMMC/Model/>

# Who are we supporting (The Preserve)?

## Number of active users, including growth by department:

Department	May-23	Sep-24
Aerospace	15	15
LASP	1	2
Mechanical Engineering	0	1
NSI	8	14
<b>Subtotal Research</b>	<b>24</b>	<b>32</b>
OIT	19	18
RIO	8	7
UCCS	1	1
Vendor	1	1
<b>Subtotal Support</b>	<b>29</b>	<b>27</b>
<b>Grand Total</b>	<b>53</b>	<b>59</b>

- 25% growth in researcher use from last year
- Supporting up to \$1.1M annually in research grants
- Emails sent – 377
- Emails received - 2089

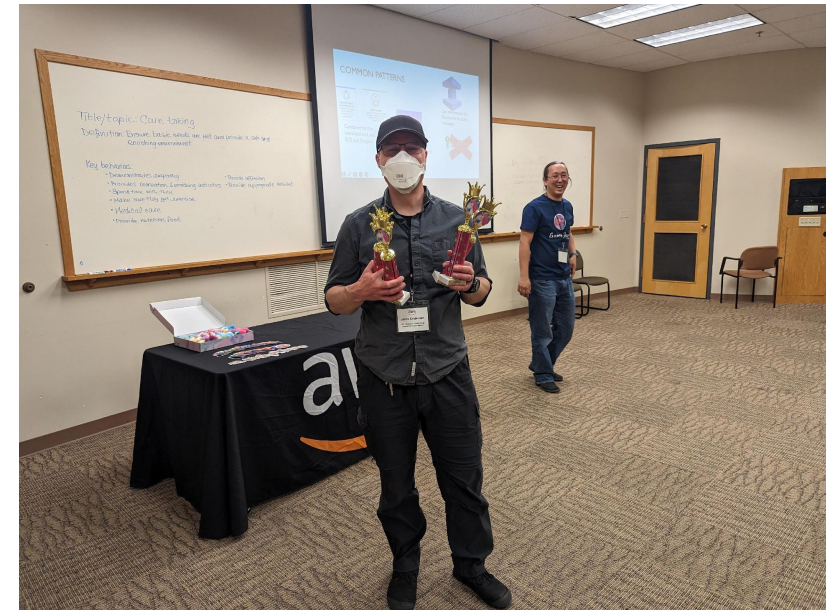
# On-Prem and Off-Prem Cloud

- Supports computing, data, and gateway needs not appropriate for existing HPC and HTC infrastructure
  - Long running jobs – website applications, databases, JupyterHub
- Cloud Foundations Service
  - AWS, Azure, Google Cloud Platform
  - CUmulus Infrastructure is retiring



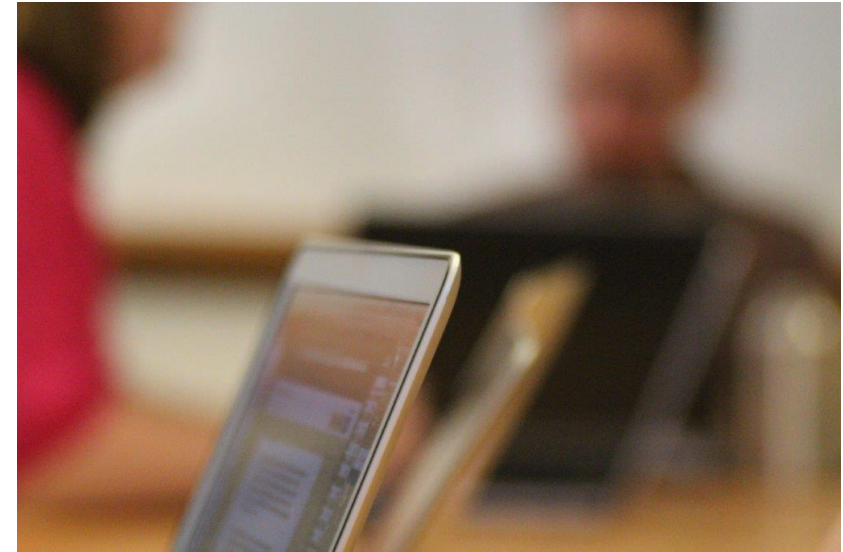
# On-Prem and Off-Prem Cloud

- Who are we supporting?
  - Top users in CIRES, Computer Science, IBS, and Psychology (38 projects)
  - >\$20M in research grants
- Notable accomplishments:
  - AWS LZA
  - Added Google Cloud as a service
  - Seeking status as a foundational service
  - New cloud 101 training



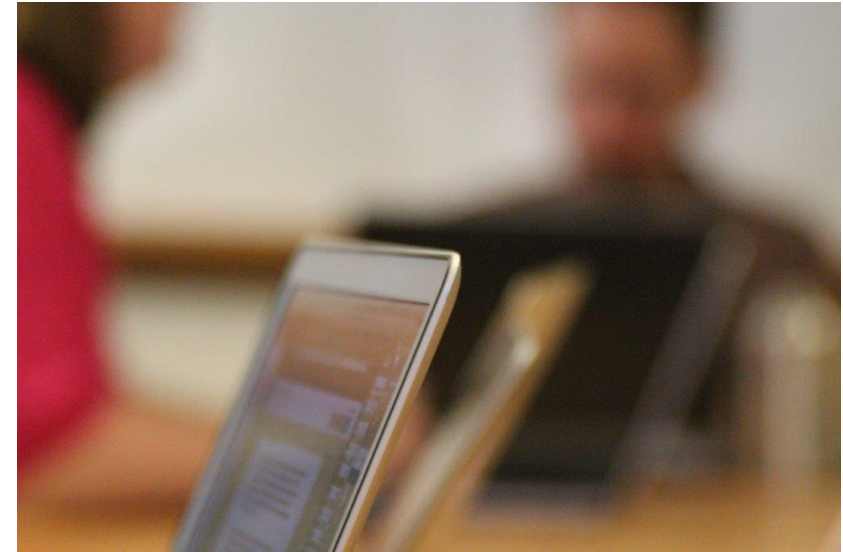
# PetaLibrary and Blanca

- PetaLibrary
  - Large scale data storage
  - Archive and Active
  - Backup options
  - Offsite backup
- Blanca condo cluster
  - Faculty can purchase hardware that we house in our datacenter and manage



# PetaLibrary and Blanca

- Allocated storage:
  - 7.0 PB (8.8 PB with backups)
  - 5.1 PB (6.6 PB with backups) for UCB
- Notable accomplishments
  - 18 new Blanca orders provisioned in June
  - PetaLibrary active hardware refresh
    - Replaced hardware losing support
    - Replaced and rebuilt several nodes
    - New OS
  - Archive plus Disaster Recovery Service (offsite)





# User Support

- Trainings (246 attendees)
  - RC Quick Byte, RC Primer, RC Short Course, RC Tech Lab, RC Workshops, and RC Micro-credentialed course
  - AlphaFold2, R, Chatbots, Generative AI apps, AWS Immersion Days, Foundations of HPC
- Interactions (nearly 500)
  - Consultations
  - Office hours
- Tickets – 263/month
- Growth in cloud skill set on team



# Grants



- NSF CC\* with CSU, OAC #2201538 (ended)
  - \$400,000 over 2 years
  - Brings CSU nodes to Alpine
- NSF ACCESS, OAC #2138286
  - \$11,947,478 over five years
  - Support NSF funded advanced cyberinfrastructure
- NSF CC\* for RMACC, OAC #2322260
  - \$1,033,411 over 2 years
- Several other ongoing grants related to ACCESS, collaborations with other groups

# ACCESS and NAIRR

- National Artificial Intelligence Research Resource Pilot
- Executive order from President Biden
- Connects U.S. researchers and educators to computational, data, and training resources
- Federal agencies are collaborating with government-supported and non-governmental partners to implement the Pilot as a preparatory step toward an eventual full NAIRR implementation
- The Pilot helps bridge the gap for researchers who lack access to AI resources needed to conduct research
- Aim to ensure that AI resources and tools are equitably accessible

# ACCESS and NAIRR

- Researchers and Educators can apply to open calls to get access to some of the premier AI resources in the country
  - Computational infrastructure
  - Tools, models, and software
  - Datasets
  - Training and educational resources
- All skill levels are invited to participate
  - Already using AI
  - Just learning – we're all new!



Comet, San Diego  
Supercomputing  
Center, from [sdsc.edu](http://sdsc.edu)

# CU's Role in NAIRR

- Been asked to lead the NAIRR User Experience Working Group
  - ~220 projects
  - What is the user experience like?
  - What recommendations do we have to best meet the needs of users?
  - Trainings and webinars
    - AI showcases
    - Upcoming AI workshop
- **Get involved!!**



Comet, San Diego  
Supercomputing  
Center, from [sdsc.edu](http://sdsc.edu)

# Diversity and Inclusion Efforts

- RMACC student cohort (17 for 2 events)
  - Shadow our sys admins!
- RMACC and ACCESS interviews
  - Interviews, Feedback Forums
  - Workforce development, education
- ACCESS
  - CCEP – 178 applications, 60 institutions, 11 minority serving
- RMACC DEI coalition



# Spotlight: RMACC Student Cohort

- Opportunity for undergraduate and graduate students to learn HPC system administration by learning from and shadowing RC staff
- Target: workforce development and expanded opportunities in HPC
- 17 have participated in some capacity: 11 on-site January, 7 on-site July (+1 remote)
- Students have come from 8 institutions including CU Boulder
- NSF CC\* OAC #2322260

# Spotlight: RMACC Student Cohort



January 2024

- Week on site
- Removal of Summit hardware from CU Boulder's HPCF data center



# Spotlight: RMACC Student Cohort

## RMACC Symposium

- Gathered to meet with RC staff and director as well as with members of RMACC institutions
- New and returning members



# Spotlight: RMACC Student Cohort



August 2024

- Week on site
- In-depth conversations with members of CURC staff regarding job responsibilities in HPC
- User support, data center, system administration
- Rollout - CPU nodes, Alpine
- (1 remote participant)

# Spotlight: RMACC Student Cohort

## Future work

- Announcement and recruiting messages go out in late 2024
- Scratch expansion and nodes will be deployed in an in-person session in spring 2025
- Students from RMACC member institutions welcome to participate
- Compiling educational material into a shareable form for use in increasing opportunities in HPC

# Events

- Stakeholder reviews (ongoing)
- RMACC Symposium (May 20-22, 2024)
  - 2023: 261 participants, 70 students
- Artificial Intelligence, NAIRR workshop
  - Dates TBD, spring
- Check our webpages for ongoing events!
  - <https://www.colorado.edu/rc/events>
  - <https://www.colorado.edu/crdds/events>
  - <https://rmacc.org/trainings-and-events>
  - <https://support.access-ci.org/events>



# Looking Ahead – 2024-2025

- Expansion of Alpine
  - High memory, scratch, CPU, GPUs
  - Anschutz deployment
  - System expansions for AI and Machine Learning
- Updating Blanca's scratch
- Rebuild of datacenters to accommodate future growth and power requirements for Alpine
- Secure Alpine – CU Boulder Confidential status



# Looking Ahead – 2024-2025

- AWS LZA
- Technical debt
- Supporting users with AI projects – determine needs
- User efficient workflows
- Continued growth in PetaLibrary disaster recovery
- Achieve CMMC certification (stretch goal)
- RMACC in person cohort
- Grow with NAIRR
- AI workshop
- Research IT governance group



# Thank you!

- Visit our websites
  - Research Computing [www.colorado.edu/rc](http://www.colorado.edu/rc)
  - CRDDS [www.colorado.edu/crdds](http://www.colorado.edu/crdds)
  - RMACC [www.rmacc.org](http://www.rmacc.org)
- Questions?
  - Email me: [shelley.knuth@Colorado.edu](mailto:shelley.knuth@Colorado.edu)
  - Email RC: [rc-help@Colorado.edu](mailto:rc-help@Colorado.edu)

