Funded by the National Science Foundation, Grant #2127548

CI Compass:
Engaging with NSF Research Infrastructures and the Cyberinfrastructure Community

Ewa Deelman
University of Southern California
CI Compass, PI

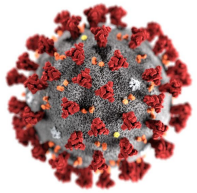
Funded by the National Science Foundation, Grant #2127548

Cyberinfrastructure for Major Facilities Workshops

2024: Collaboration in Action

Goals

- Foster discussions on CI topics relevant to MFs
- Share findings of current CI Compass working groups
- Define new areas of collaboration, potentially form new topical working groups
- Plan for other community meetings RIW'24
- Continue to build a community around CI for MFs and Midscale projects

2017 2019  2022 2024

Funded by the National Science Foundation, Grant #2127548

Mission

CICompass

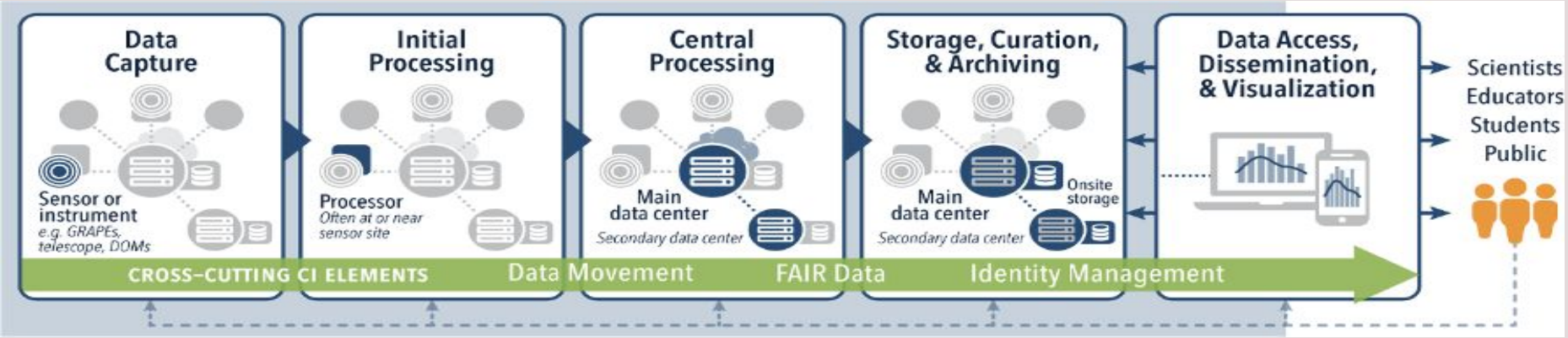
CI Compass provides expertise and active support to cyberinfrastructure practitioners at NSF Major Facilities in order to accelerate the data lifecycle and ensure the integrity and effectiveness of the cyberinfrastructure upon which research and discovery depend.

Funded by the National Science Foundation, Grant #2127548

Cyberinfrastructure for Major Facilities Workshops

2024: Collaboration in Action

- **Survey on topics Fall 2023**



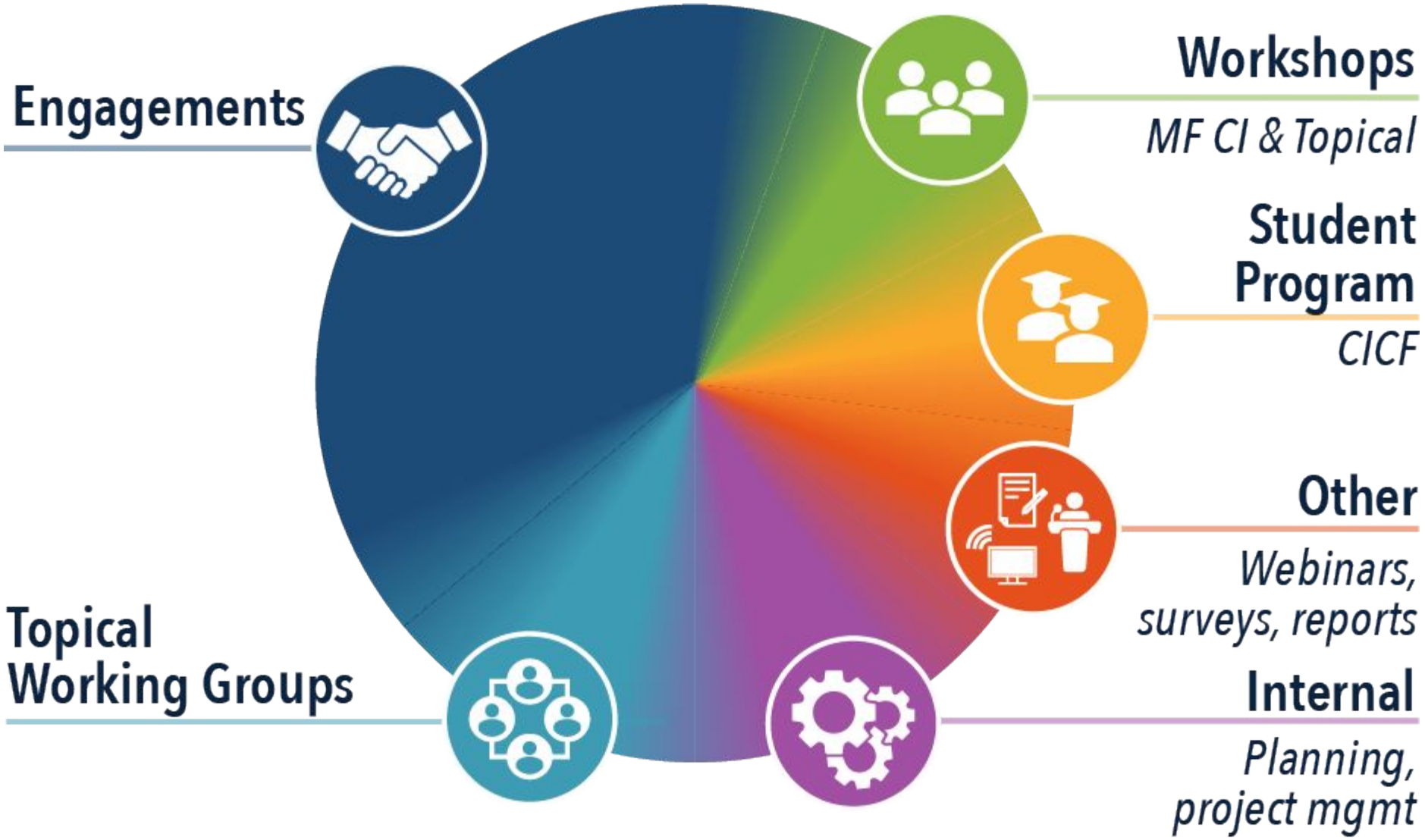
Flow of data through the data lifecycle:

- Data transport, processing, and visualization
- Persistent identifiers and FAIR data
- Designing/Redesigning CI
- Leveraging national CI

Panels
Breakout sessions
Table discussions
Community Lightning talks
Invited speakers

Funded by the National Science Foundation, Grant #2127548

CI Compass activities



Funded by the National Science Foundation, Grant #2127548

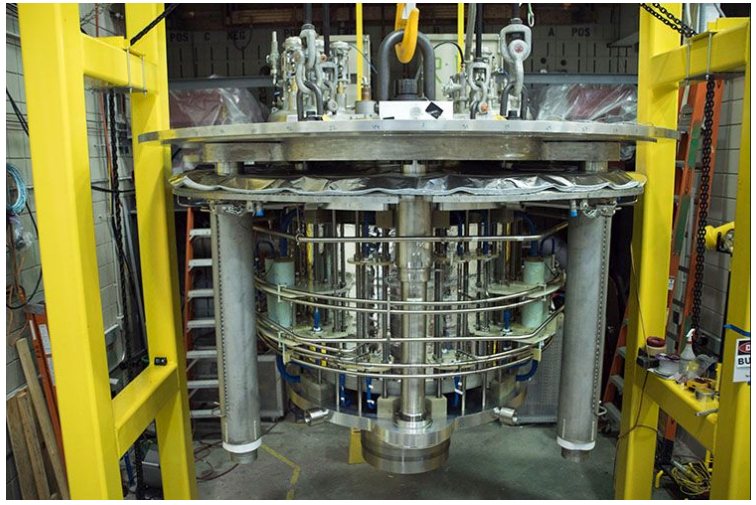
Examples of Engagement topics

- CI design, ab novo and refresh
- Use of cloud computing in the data lifecycle
- Network design review
- Architecting identity management solutions
- Metadata capture and FAIR data
- Visualization, portal-based data display

Help connect with community efforts:

- Research Data Alliance (RDA)
- Earth Science Information Partners (ESIP)
- ORCID

Laser Interferometer Gravitational-Wave Observatory



MAGLab 36-tesla Series Magnet

Funded by the National Science Foundation, Grant #2127548

Topical Working Groups

FAIR Topical Working Group

Interest group for people interested in FAIR and data management at Major Facilities

- Facilitates conversations on FAIR between MFs
- Discussion on best practices and emerging issues
- Open membership with representation from MFs, and Mid-scale facilities, and others
- Engagement with NSF FAIROS RCN groups
- Building collaborations and doing outreach where it makes sense

Other working groups:

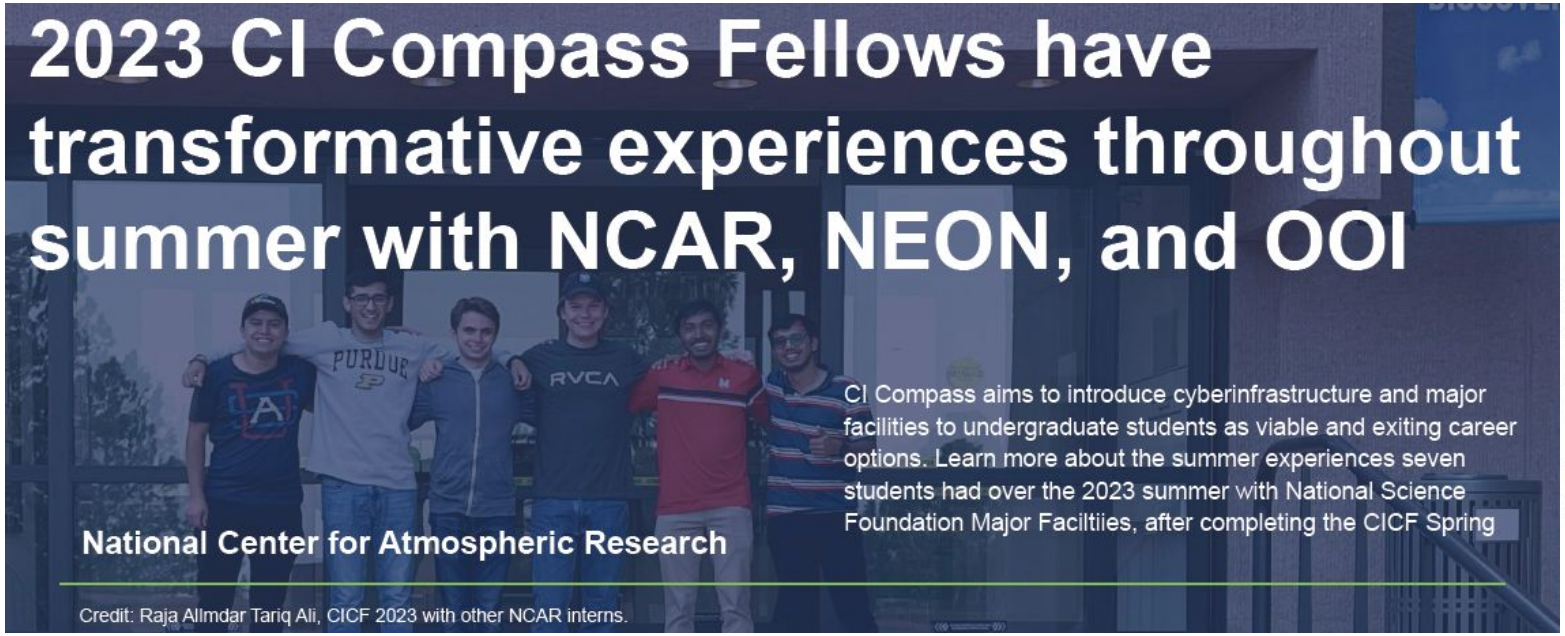
- Identity Management: IdM Cookbook (in Collaboration with Trusted CI)
- Cloud Infrastructure: Working on a report about adopting clouds for MFs
- DLC Guidebook (planned)

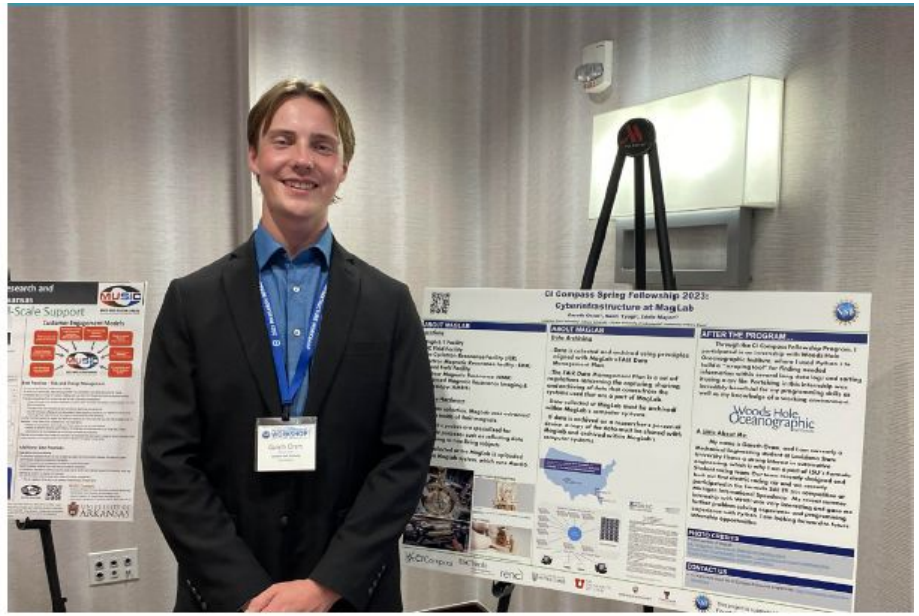
CI Compass
CI4MF 2024 - Collaboration in Action

Funded by the National Science Foundation, Grant #2127548

Undergraduate Student Fellows Program

- Spring Technical Training Program: Virtual
- Summer Internships at MFs
- Opportunity to attend a conference of their choice





Gareth Oram, CICF 2023, had a summer internship with OOI. Credit: CI Compass.

Funded by the National Science Foundation, Grant #2127548

Community Materials and Events

CICompass TECH NOTES

ci-compass.org

Making the Major Facilities Data Lifecycle FAIR

Charles Vardeman
Date Published: January 25, 2022

What is FAIR data?

The notion of the four foundational principles for "data" — Findability, Accessibility, Interoperability, and Reusability or "FAIR" — was proposed by Wilkinson et al. in *"The FAIR*

Knowledge Informed Machine Learning **[2]**, that integrates broader knowledge and context into the machine learning process. Specific attributes for each FAIR principle are contained in **Table 1**

CICompass
Webinar Series

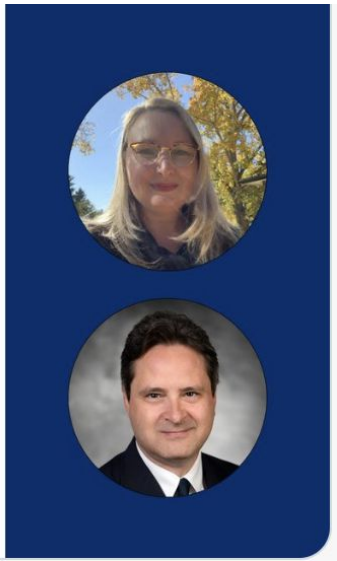
NSF, the Nelson Memo, and Persistent Identifiers

ft. ***Martin Halbert****, NSF Advisor for Public Access, and* ***Shawna Sadler****, Head of Outreach and Partnership at ORCID*

January 31, 2023
2 p.m. EST / 11 a.m. PST, via Zoom

To learn more and register, please visit:
ci-compass.org/news-and-events

ALT



CICompass TECH NOTES

ci-compass.org

Tracking community access to Data Lifecycle data using Knowledge Graphs

Authors: Don Brower and Rodney Ewing

2023 RESEARCH INFRASTRUCTURE
WORKSHOP
JUNE 27 - 30, 2023 • WASHINGTON, DC

Funded by the National Science Foundation, Grant #2127548

RI Workshop Themes

The annual NSF RI Workshop is a collaborative forum for the RI community to share **best practices** and **lessons learned**, **collect community input** on new initiatives, and share project, operations, business tools and techniques.

Contact rioutreach@nsf.gov for more information.

- **Continuations from 2023:**
  - **Mid-scale Research Infrastructure**
  - **Award Management and Guidance**
  - **Cyberinfrastructure/Cybersecurity**
  - **Environmental Health and Safety**
  - **Program and Project Management**
  - **Facilities and Operations Management**
  - **Comms, Education and Public Outreach**
- **New:**
  - **2024 RIG Revisions**

Funded by the National Science Foundation, Grant #2127548

Cyberinfrastructure for Major Facilities Workshops

2024: Collaboration in Action

Outcomes

- Understanding of Research Infrastructure (RIs) current CI needs and challenges
- Sharing of solutions and best practices
- Ideas for potential topical working groups and engagements
- Continue to build a community around CI for RIs
- Topics for RIW'24, please volunteer to help organize with us
- A webinar summarizing the workshop discussions
- A report summarizing the discussions

CICompass
CI4MF 2024: Collaboration in Action
Workshop Notes

How to contribute:

- Participate in the discussions
- Add to the notes
- Fill out Google Form during tabletop discussions (day 2)
- Fill out our post workshop survey
- Reach out to us if you want to host a CI Compass Student Fellow

http://tinyurl.com/ci4mfs2024

Funded by the National Science Foundation, Grant #2127548

Wednesday, January 17 ¶

8:00 AM → 8:15 AM **Workshop Welcome**

8:15 AM → 9:00 AM **Keynote: Leveraging NSF's Cyberinfrastructure to Support Major Facilities Research**

by Katie Antypas (NSF OAC)

9:00 AM → 10:00 AM **Update on CI Compass Activities**

10:00 AM → 10:30 AM **Coffee Break**

10:30 AM → 12:00 PM **Panel: MF Approach to Open Science: FAIR Data, Persistent Identifiers, Etc.**

12:00 PM → 1:45 PM **Lunch Break & Group Picture**

1:45 PM → 2:45 PM **Lightning Talks: 1**

2:45 PM → 3:00 PM **Visualization Demo**

3:00 PM → 3:30 PM **Coffee Break** 30m

3:30 PM → 4:45 PM **Breakout Session: 1**¶

3:30 PM **Making MF Data More Accessible: Data Visualization and Analytics**

3:30 PM **Use of National CI for MFs and Their Users**

4:45 PM → 5:05 PM **Breakout Session Reports Back**

5:05 PM → 7:00 PM **Reception w/ Cash Bar**

7:00 PM → 10:00 PM **Dinner (on your own; groups encouraged)**

CI Compass
CI4MF 2024: Collaboration in Action
Agenda
(login key: ci4mf24_longbeach)

Funded by the National Science Foundation, Grant #2127548

Thursday, January 18 ¶

7:00 AM → 8:00 AM **Breakfast**

8:00 AM → 8:05 AM **Day 2 Welcome**

8:05 AM → 8:30 AM **Invited Talk: Enabling Open Science and Data Sharing: Trust, Provenance, and Data Integrity** **Speaker**: Michael Corn (NSF/CORF)

8:30 AM → 10:00 AM **Panel: Coordinating and Combining Data Processing, Movements, and Storage**

10:00 AM → 10:30 AM **Coffee Break**

10:30 AM → 10:50 AM **Invited Talk: Designing Cyberinfrastructure for the Antarctic Research Vessel**¶20m

Speaker: Mike Prince (NSF)

10:50 AM → 11:50 AM **Lightning Talks: 2**

11:50 AM → 1:15 PM **Lunch**

1:15 PM → 2:30 PM **Panel: If I knew then what I know now, I would have never designed it this way: Approaches to CI conceptualization and design**

2:30 PM → 2:55 PM **Table-Top Discussions**

2:55 PM → 3:00 PM **Workshop Closing Remarks**

3:00 PM → 4:00 PM **Social Coffee Hour**

Funded by the National Science Foundation, Grant #2127548

Please tell other participants about your biggest recent CI (including workforce development) accomplishment? ---- Summary below provided by ChatGPT

*The collective accomplishments in **Continuous Improvement (CI)** and workforce development shared by participants highlight significant advancements across various scientific and technological domains. Notable achievements include the commissioning and network enhancement of advanced X-ray light source facilities at ASU, the migration of substantial data processing capabilities to cloud-based environments, and the successful processing of complex datasets by leveraging multi-institutional GPU capacities. A standout achievement is the service provision to over 460 users from 100 institutions, managing a remarkable rate of over **five jobs per second**. There's also progress in the design and approval phases of the Antarctic Research Vessel project, integrating cyber infrastructure and security. Participants are actively involved in shaping the future of CI through leadership roles in reports on NSF Major Facilities and cloud use, advocating for the adoption of persistent identifiers for research facilities, and studying the social aspects of CI integration. Contributions to the educational aspect of CI are evident in the CI Compass Fellowship program, which supports the growth of students and professionals in the field. Data management remains a pivotal area, with efforts focused on managing extensive atmospheric datasets, enhancing data discovery, and promoting community-driven open data systems. Lastly, the community has engaged in international collaboration to establish models for global research commons, emphasizing equity in data access and utilization.**

5.18

** 21 responses used*

Funded by the National Science Foundation, Grant #2127548

Please share some thoughts about what are the things that keep you up at night or opportunities you see for your MF/project (in the context of CI) ---- Summary below provided by ChatGPT

1. *Deployment of advanced cyberinfrastructure for the CXFEL project, including network, data storage, and computational upgrades for research facilitation.*

2. *Addressing workforce development by hiring and retraining for evolving CI needs, and promoting open science through enhanced data management.*

3. *Advancing CI through AI integration, cross-disciplinary collaborations, and sustainable, secure, domain-specific infrastructure solutions..**

   *21 responses used

Funded by the National Science Foundation, Grant #2127548

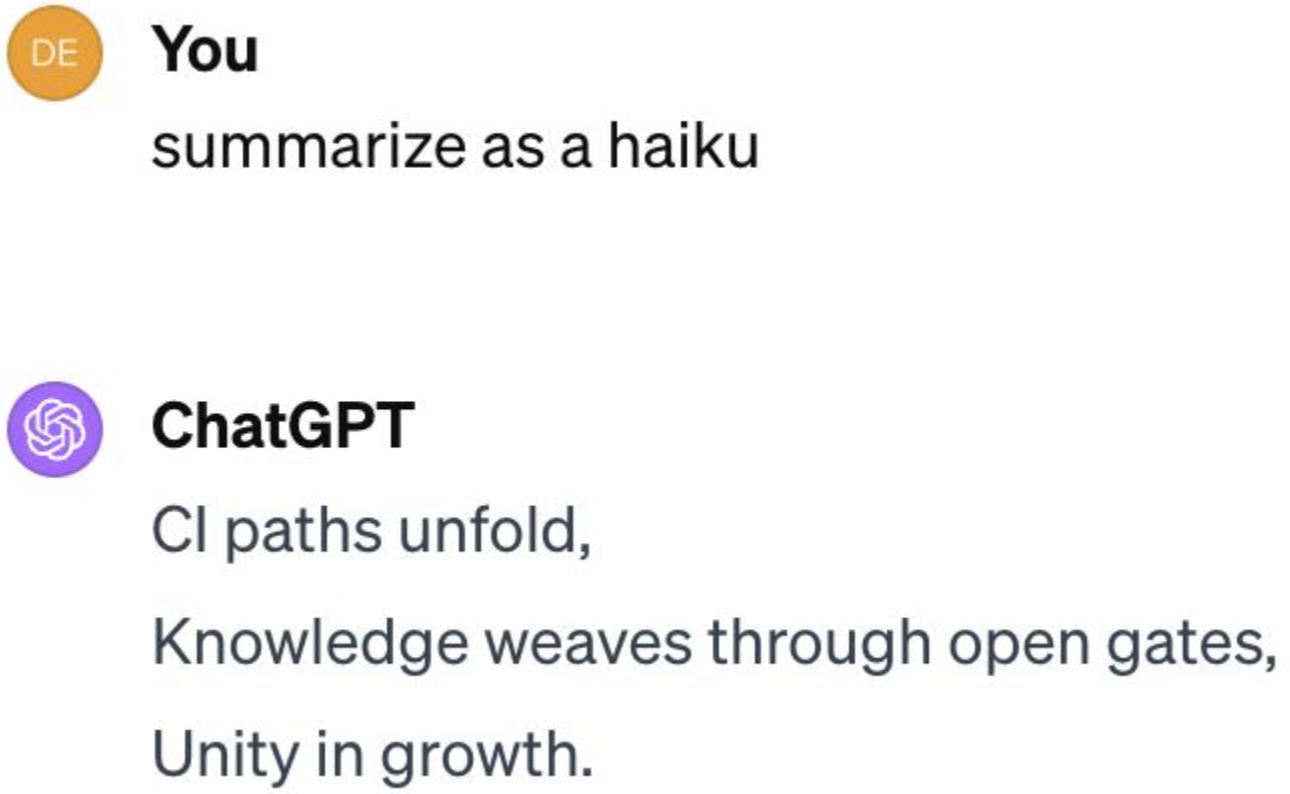
Please share your ideas about topics you would like our CI community to discuss together (technical or social)? ---- Summary below provided by ChatGPT

*Participants are keen on leveraging their interactions with the expert community to enhance their Cyberinfrastructure (CI) projects, like the CXFEL, through the adoption of best practices, especially in critical procurement stages. Extending software technology reach to more Major Facilities (MFs) and addressing the challenges of workforce development and education for new CI users are highlighted as significant objectives. Designing CI that balances open access with robust cybersecurity and reliability is another focus, alongside the strategic use of cloud computing and the advancement of open science principles. There's a strong desire to engage younger audiences in CI and to foster a more socially connected community. Enabling efficient data analysis through tools like a Data Workbench and integrating AI into data center operations is seen as an opportunity for performance enhancement. Participants also aim to facilitate a more united CI community by aligning common issues across various funding agencies and international efforts. Expanding and refining educational programs like the CI Compass Fellowship, promoting open data practices, and addressing scalability, interoperability, and legacy system integration challenges are other key points. Engaging with MFs for feedback and collaboration to support the scientific community's evolving needs and training programs is also a shared ambition..**

- *21 responses used*

Funded by the National Science Foundation, Grant #2127548

Please share your ideas about topics you would like our CI community to discuss together (technical or social) ---- Summary below provided by ChatGPT



*

*21 responses used