2023 RESEARCH INFRASTRUCTURE WORKSHOP

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

Ewa Deelman
University of Southern California
CI Compass, PI

PROJECT MANAGEMENT. FACILITIES AND OPERATIONS. AWARD MANAGEMENT. EHS. EDUCATION AND PUBLIC OUTREACH. CYBER (CI & CS).
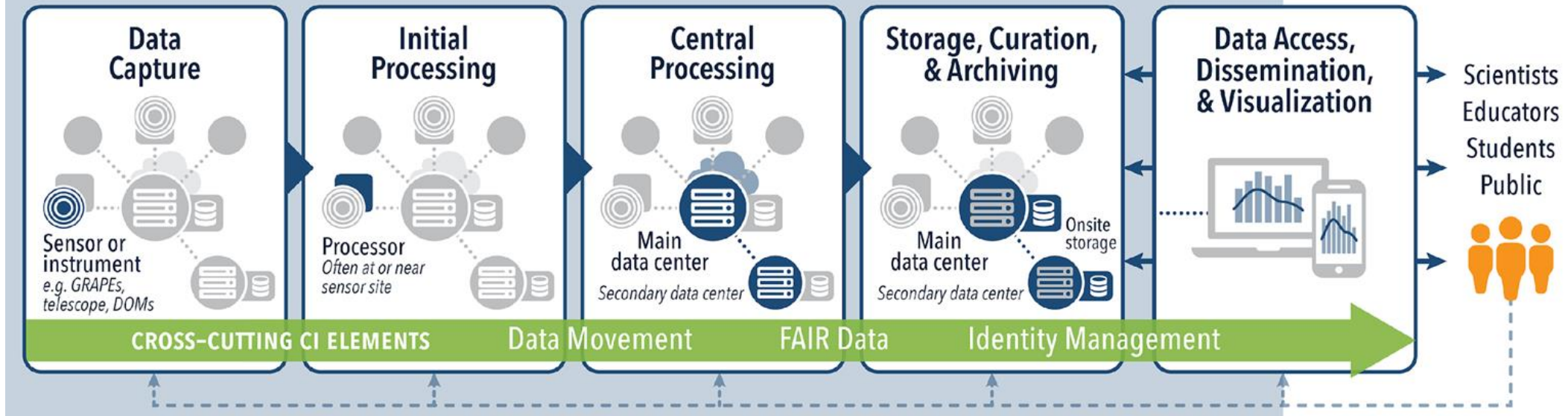
Mission

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.

ci-compass.org

CI Compass Services focus on Major Facilities' Data Lifecycle

Evaluate CI plans, Help architect new solutions, Develop proofs of concept, Assess applicability/performance of existing solutions, Help leverage existing technologies

CI Compass Team: Who we are

Deep expertise in several CI areas critical to the MFs

- Data management, data processing, visualization, archiving, semantic technologies
- Automation, resource management, workflows, sensors
- Networking, clouds, systems and infrastructure
- Large-scale CI deployment and operations, IdM
- Social science, understand the organization structures and culture of MFs

Experience in the management of CI projects

- Conceptualization, design phase, broad adoption
- Project Management and Evaluation
- Organizational science
- Communications & Outreach

Highly collaborative, strong history of working together and with the CS and CI Communities

- Many diverse community connections in astronomy, earth science, physics

Dedicated to the advancement of CI for science, engineering, and education

CI Compass and Trusted CI

- Two of the premier CoEs funded by NSF/OAC to help the NSF science community

- Co-founded the Identity Management Working Group/

- Co-authored IdM cookbook

- Share CoE best practices and lessons learned

- Have standing and open communication and collaboration channels

By Amber Rasche - Senior Communications Specialist, Internet2

CI Compass and its Identity Management (IdM) Working Group recently published *The Federated Identity Management Cookbook*, a new resource that provides time-tested recipes for building IdM capabilities for research cyberinfrastructure. It also serves as a primer on IdM concepts, tools, and best practices for the broader research and education community – with the InCommon Federation as a key ingredient.
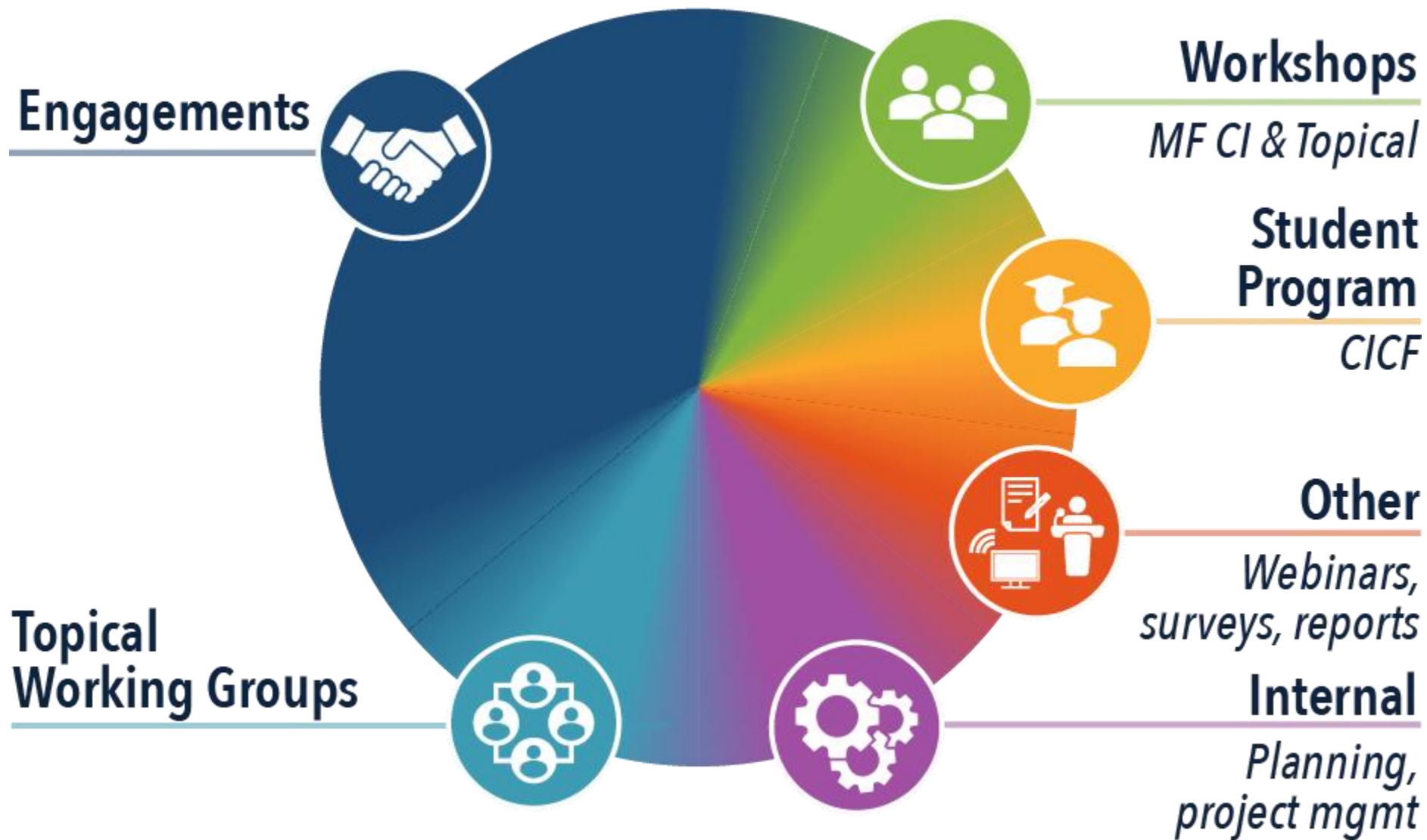
Eric Scott and Josh Drake

Not sure which center to approach with a question or challenge?

Approach either and we'll collaboratively figure out how to best help you.

CI Compass activities

Engagements

Workshops
MF CI & Topical

Student Program
CICF

Other
Webinars, surveys, reports

Internal
Planning, project mgmt

Topical Working Groups

Regional Class Research Vessel (RCRV) Engagement:
Shipboard CI/network plan review

Planned RCRV vessels

January - March 2022

experienced on the vessel. Systems of this sort range from high-resolution video conferencing setups to full-scale virtual reality. As the degree of immersiveness increases, so does the demand on the CI capacity. This portion of the CI is still under design.

Review of Network Architecture

The basic architecture of the on-board network is a switched hub-and-spoke model. The central hub

"One of the primary concerns identified by the review was that the planned 1GbE switch ports in the ship's computer lab should be supplemented with 10 GbE and higher to support deployment of visiting equipment with high-speed network interfaces. "

Field report from ARF, RV Sikuliaq

From Chris Romsos, RCRV, OSU : "Thank you for identifying this as something to address now before delivery of the vessel. We planned for future upgrades like this and have sufficient fiber between the network core and the computer lab to support the upgrade…

Sikuliaq recently upgraded their edge switching throughout the vessel…．A nice piece of corroborating evidence there with Sikuliaq! "

Engagement: MagLab   2023-

- **Goal**: Provide the MagLab assistance in developing FAIR data practices for the DC Field facility and its community of users.

- **Focus**: FAIR instrument descriptions leveraging Research Data Alliance (RDA) Persistent Instrument Identifiers (PIDINST) working group recommendations. Implementation will allow scientists to specify resolvable instrument identifiers in publications.

- **Community**: Engaging Earth Science Information Partners (ESIP) science on schema.org cluster to extended Dataset and Repository with Instrument recommendation. Connecting with ORCID and NSF FAIR Instrument RCN.

- **Deliverable**: Example FAIR instrument documents in JSON-LD along with documentation.

17 Tesla Superconducting Magnet for Optics
Credit: National MagLab

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
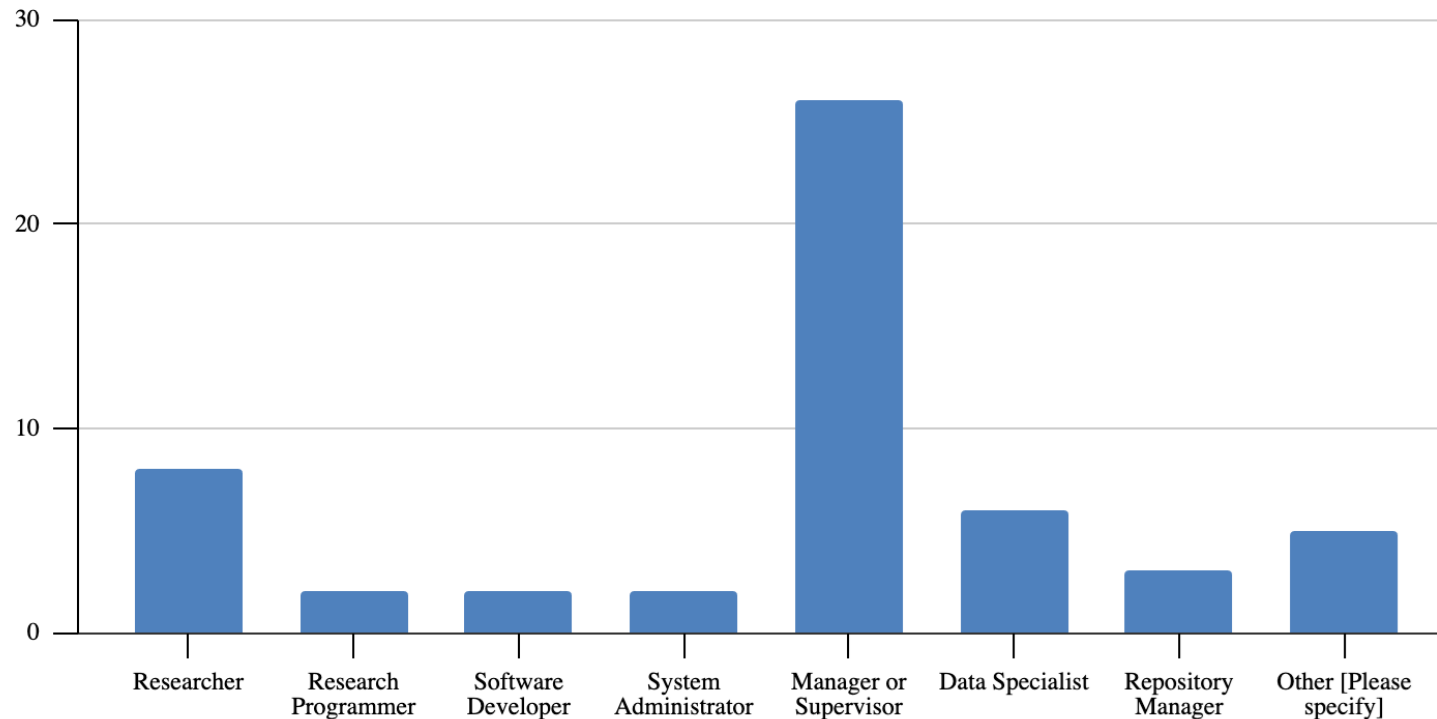- Building collaborations and doing outreach where it makes sense, partnered with **ORCID**

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): Sharing ideas and developing best practices for CI to support data lifecycle management

FAIR WG Survey Participants

Participants from at least 14 different MFs
Majority were Managers or Supervisors or Researchers

Which NSF Major Facility do you work for? (Count)

Facility	Count
United States Antarctic Program	1
Arecibo Observatory	1
Academic Research Fleet	4
IceCube Neutrino Observatory	3
International Ocean Discovery Program	1
Leadership-Class Computing Facility	1
Large Hadron Collider (including ATLAS/CMS)	1
Laser Interferometer Gravitational-wave Observatory	1
National Ecological Observatory Network	4
SAGE/GAGE	2
National High Magnetic Field Laboratory	1
National Optical-Infrared Astronomy Research Laboratory	2
National Radio Astronomy Observatory	1
Ocean Observatories Initiative	5
Not Listed [Please specify]	22

What is your Primary Job? (Count)
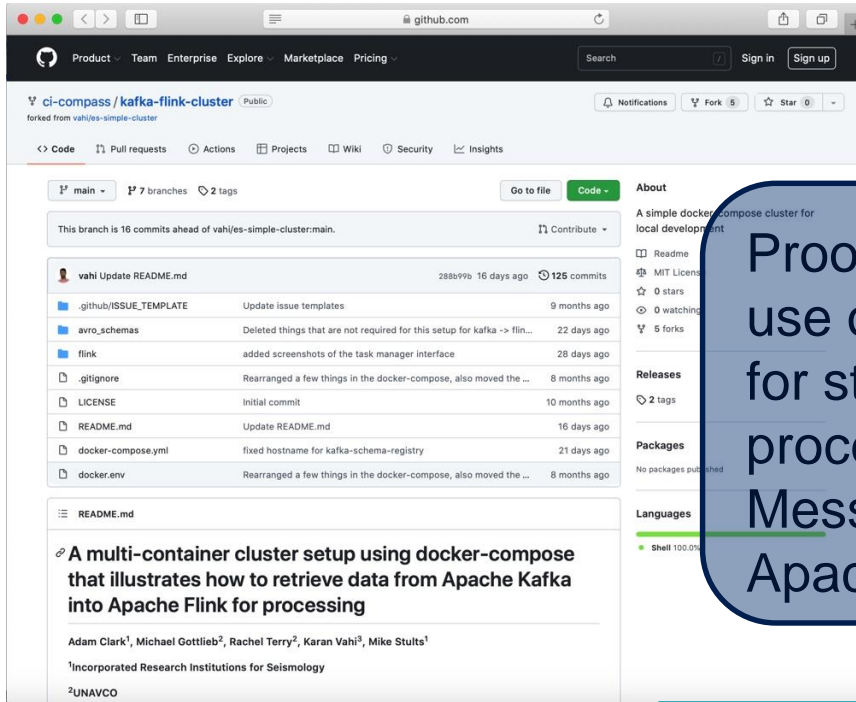
FAIR Topical Working Group Survey

How much of a hindrance have the following been in regards to implementing FAIR at your Facility?
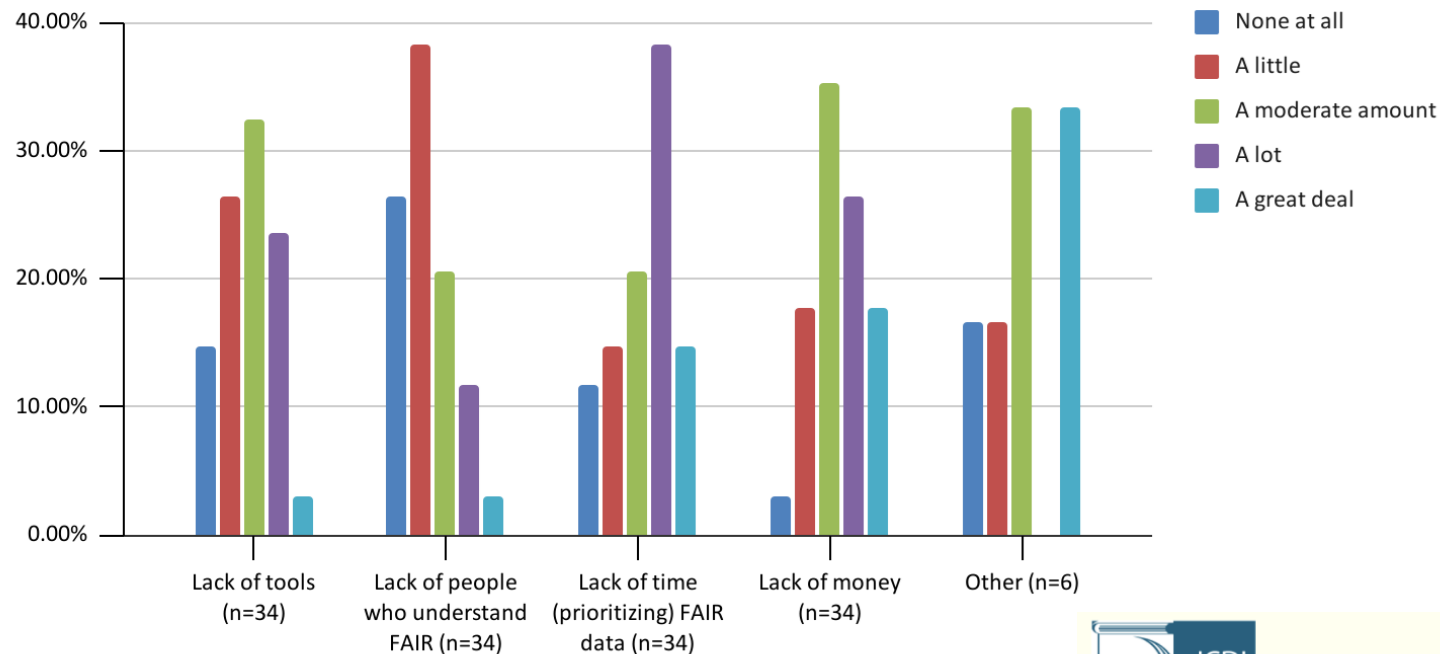
Survey taken March 2023

Biggest hindrance was lack of time

The group will report on the full set of findings

JCDL

Joint Conference on Digital Libraries

Upcoming Conference: JCDL 2023, June 26-30, 2023, Santa Fe, New Mexico, USA.

Proof of concept on use of Apache Flink for stream-based processing of Messages out of Apache Kafka

CI Compass 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**

CI Compass TECH NOTES
ci-compass.org

Building the Next Generation of Cyberinfrastructure Professionals

Author: Angela Murillo

CI Compass TECH NOTES
ci-compass.org

Tracking community access to Data Lifecycle data using Knowledge Graphs

Authors: Don Brower and Rodney Ewing

Webinars

Available online

If you would like to give a seminar or suggest a topic related to DLC, please contact us: contact@ci-compass.org

CI Compass Fellowship Program (CICF)

CICF facilitator Rajiv Mayani works with Fellow Nona Nersisyan at the University of Southern California Information Sciences Institute campus.

Goal: Broaden student participation in CI research, development, deployment, and operations

Virtual Spring Training Program
- *Technical Skills*: Students taught technical skills relevant to CI.
- *Research Skills*: Students research MFs to understand the importance and context of MFs, and the related data and CI.
- Free to undergraduate students, with the possibility of course credit.
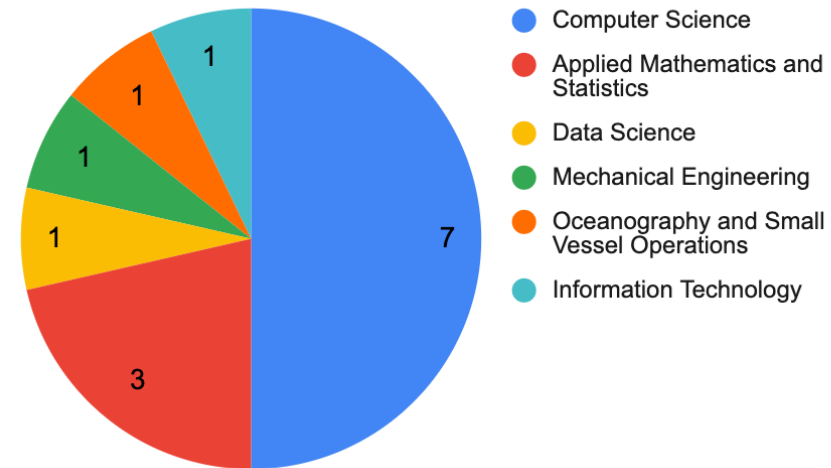
(Optional/Invited) Summer Project Program
- We are collaborating with MFs to provide CI-related summer projects for selected student fellows.
- In-person or virtual, depending on the MF/project.
- Students are paid for their participation.

CI Compass Fellowship Program (CICF)

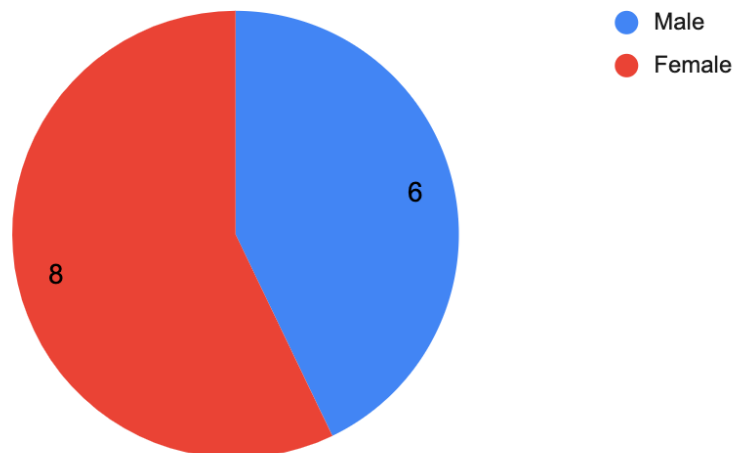| Institution (9 schools)	Count
Indiana University - Purdue University Indianapolis	4
Indiana University - Bloomington	1
University of Notre Dame	3
University of Iowa	1
Louisiana State University	1
Maine Maritime Academy	1
University of Alabama at Birmingham	1
University of California, San Diego	1
Arizona State University	1

Majors

- Computer Science
- Applied Mathematics and Statistics
- Data Science
- Mechanical Engineering
- Oceanography and Small Vessel Operations
- Information Technology

At RIW 2023

Nardine Ibrahim, IU

Gareth Oram, LSU

Gender

- Male
- Female

Year

- 1st (Freshman)
- 2nd (Sophomore)
- 3rd (Junior)
- 4th (Senior)
- 5th+ (Super Senior)

SPRING 2023 PROGRAM (12-weeks, virtual)
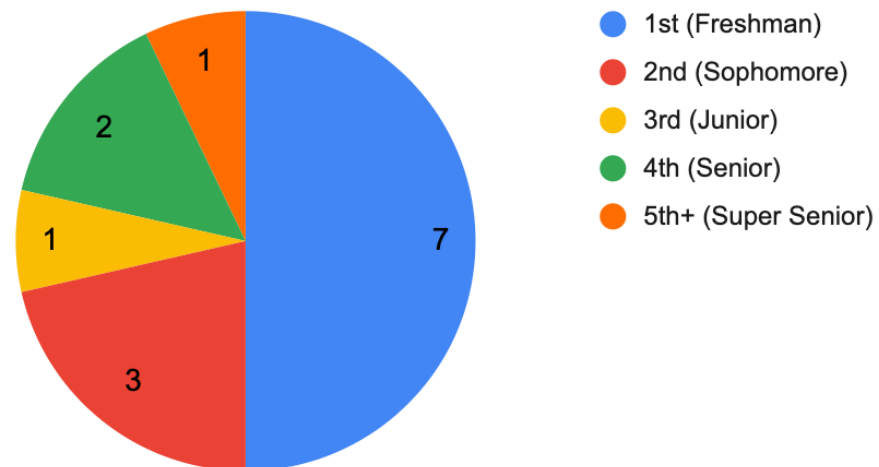
| Week	Technical Training Program	Research Training Program
Week 1	Orientation, Linux/Unix Shell, Terminal	Cyberinfrastructure, Major Facilities, the Data Lifecycle. Robin Nelson, Natural Hazards Engineering Research Infrastructure (NHERI)
Week 2	Introduction to Python Programming	Research Data Management. Shawna Sadler, ORCID
Week 3	Python Programming, Jupyter Notebooks, Python Data Analysis Packages	Research Computing. Dan Stanzione, Texas Advanced Computing Center (TACC)
Week 4	Best Practices in Software Development: Version Control, GitHub, Pytest	Research and Data Ethics
Week 5	Best Practices in Software Development: Container, Docker	Brian Dobbins, National Center for Atmospheric Research (NCAR)
Week 6	Cloud Computing, Part 1	Major Facilities and the Data Lifecycle
Week 7	Cloud Computing, Part 2	Major Facilities and the Data Lifecycle
Week 8	Chameleon Cloud	FAIR Data
Week 9	Spring Break	Spring Break
Week 10	Data Workflows, Pegasus	Professional Skills, Networking
Week 11	David Butcher, the National High Magnetic Field Laboratory (MagLab)	Group Presentations Day 1
Week 12	Machine Learning/AI	Group Presentations Day 2

CICF Student Impact

Testimonials:

- *"As a result of my participation in the program, **I am now more eager than ever to explore opportunities in scientific computing and cyberinfrastructure**."*

- *"**I knew nearly nothing about major facilities coming into this program. I left knowing so much and wanting to continue researching.** I didn't think I was interested in scientific computing until this program."*

Summer 2023:

- 5 students working with NCAR/NEON
- 2 students working with OOI
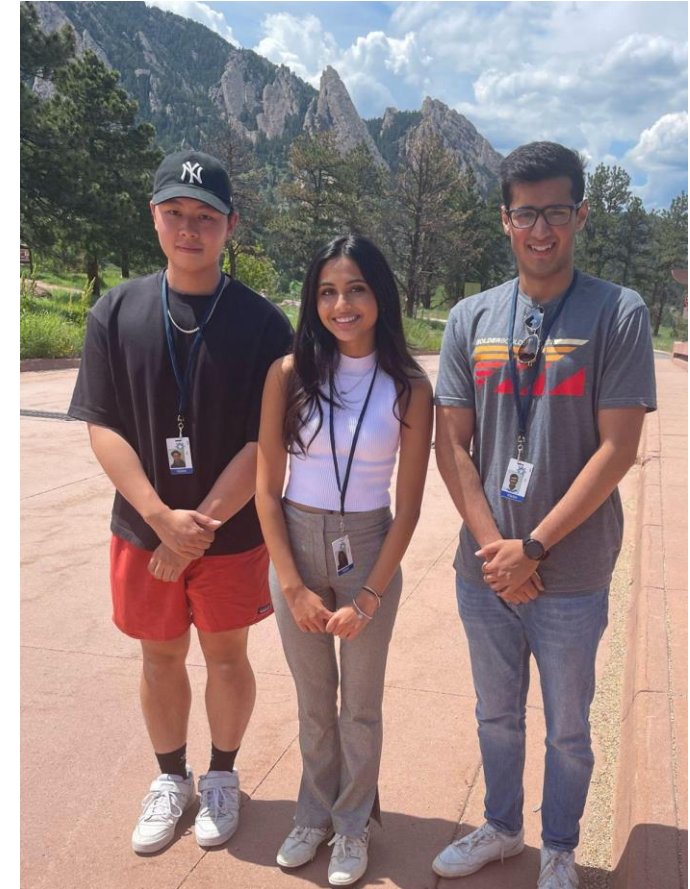
Three of our fellows (Ed Lin, Mahee Shah, and Raja Ali) at NCAR in Boulder, CO this summer

Want to get involved with CICF?

Spread the word to interested undergraduates!

We are currently recruiting Partners:

- **Major Facility Partners**
  - Work with CI Compass to provide summer projects for student fellows
  - Provide feedback to CI Compass on technical skills desired of the next generation of CI Professionals
  - Give a guest lecture related to the Data Lifecycle

- **Faculty Mentors**
  - Help recruit Student Fellows at your Institution
  - Provide course credit or auditing options for students (institutional requirements permitting)
  - Hold periodic meetings with mentees and CI Compass
  - Faculty Mentors will receive a stipend for their time

Connect with us!

Email: cicf@ci-compass.org

Website:

Social Media Outreach

From Ship to Shore
CI COMPASS SUPPORTS RESEARCHERS IN ACCELERATING SCIENTIFIC DISCOVERY

When was the last time you participated in a summer reading program?

- 2023 CI Compass Summer Reading Program, where our members and partners are invited to submit reading materials they have found interesting and valuable.
- These submissions include items on AI, FAIR Data, Machine Learning, Cybersecurity, and more.
- Connection and sharing of materials can keep us all up to speed in the fast-paced CI environment.
- Sharing on our Twitter and LinkedIn, as well as a periodic update of submissions in our Newsfeed.

Follow @CiCompass on Twitter, and CI Compass on LinkedIn

Want to submit items? Visit ci-compass.org/about/contact to find the submission link.

By Christina Clark

CI Compass Summer Reading List

Title: ChemCrow: Augmenting large-language models with chemistry tools

Author(s): Andres M Bran, Sam Cox, Andrew D White, Philippe Schwaller

Why we recommend it?: "I recommended this paper as an example of LLM (GPT) based agents for scientific research to illustrate how scientists might leverage these emerging technologies to advance scientific knowledge."
- Charles Vardeman, CI Compass and the University of Notre Dame

Access the article using this address:
https://arxiv.org/abs/2304.05376

CI Compass Summer Reading List

Title: Cybersecurity in a Large-Scale Research Facility— One Institution's Approach

Why we recommend it?: "A good introduction to MagLab and in-depth discussion of an effective cybersecurity program for an NSF Major Facility."
From Jim Basney, Director of Trusted CI

Author(s): Butcher, David S., Christian J. Brigham, James Berhalter, Abigail L. Centers, William M. Hunkapiller, Timothy P. Murphy, Eric C. Palm, and Julia H. Smith.
Access the article using this address:
https://doi.org/10.3390/jcp3020011
https://www.mdpi.com/2624-800X/3/2/11

March 2022 Cyberinfrastructure for Major Facilities Workshop

Goals: Bring together the MFs, CI community, and NSF to

- Share best practices in use of CI
- Discuss opportunities in leveraging the NSF CI ecosystem (lightning talks)
- Brainstorm solutions to the challenges we face today and in the future

Topics: FAIR data, Clouds, Workforce development, Future CI

Outcomes:

- A webinar summarizing the workshop discussions
- A report summarizing the findings and recommendations, including the analysis of the CI Calling Cards
- **https://tinyurl.com/ci4mf-report-2022**

Planning the 2024 workshop early in January in Southern California, if you would like to attend please email us at **contact@ci-compass.org**

How do we build a community around MFs and CI?

- Conducted 24 in-depth interviews with MF professionals across MFs about community building.
- Provided training opportunity for 3 graduate students at TTU (Hispanic Serving Institution) and 1 graduate student at Indiana University.
- Preliminary findings about community building ideas:
  - Facilitate more working groups on common challenges
  - Create a directory/registry of MF professionals
  - Cultivate interpersonal relationships for building of a trusting and sharing community of practice
- More analysis to come, including qualitative inductive content analysis and textual analysis using topic modelling

Kerk
Kee

Mahedi
Hasan

Sayo
Okunloye

Chaitra
Kulkarni

Angela
Murillo

Jose
Cordova

BUILDING COMMUNITY

CICompass

![2023 Research Infrastructure Workshop logo]

2023 RESEARCH INFRASTRUCTURE
WORKSHOP
JUNE 27 - 30, 2023 • WASHINGTON, D.C.

Come visit us at our table on Tuesday and Wednesday!

Kerk Kee,
TTU
Community Building Research

Christina Clark,
UND
Communications
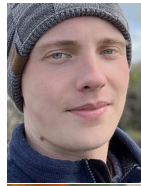
Tuesday, June 27, 2023

1:35 PM - 3:05 PM	Panel: A Discussion on Different Models of Data Governance	Moderator: Jarek Nabrzyski, UND
3:05 PM - 3:25 PM	Break	
3:25 PM - 5 PM	Panel: The Expanding Use of AI in Research Infrastructure Applications	Moderator: Chuck Vardeman, UND

Wednesday, June 28, 2023

12:35 PM - 1:05 PM	Virtual Roundtable: CI Compass	Erik Scott, RENCI
4:15 PM - 5 PM	In-Person Poster Session ● CI Compass: The NSF Cyberinfrastructure (CI) Center of Excellence for Navigating the Major Facilities Data Lifecycle ● CICF Student Fellow Poster ● FAIR Data Implementation at NSF Major Facilities ● Building the Next Generation of Cyberinfrastructure Professionals	Anirban Mandal, RENCI Gareth Oram, LSU Chuck Vardeman, UND Nicole Virdone, USC

Friday, June 30, 2023

12:30 PM - 1:05 PM	Virtual Roundtable: CI Compass	Erik Scott, RENCI

Connect with us!

To learn more about CI Compass services, leadership, news, upcoming events and our resource library,
please visit **ci-compass.org**

Contact the CI Compass Team with questions or requests by emailing

contact@ci-compass.org

Social media

Twitter
Follow **@CICompass**

LinkedIn
Connect with us **linkedin.com/company/ci-compass**

YouTube
Subscribe to our channel
CI Compass