

Cyberinfrastructure for Major Facilities Workshop

Getting Together, Working Together

In person in Redondo Beach, CA and Online

Ewa Deelman

University of Southern California
CI Compass, PI

March 1, 2022



Automation, Resource Management, Workflows, Project Management

USC

- Ewa Deelman (PI)
- Mats Rynge
- Karan Vahi
- Loïc Pottier
- Rajiv Mayani
- Ryan Tanaka
- Ciji Davis

RENCI

- Anirban Mandal (co-PI)
- Ilya Baldin
- Laura Christopherson
- Erik Scott



Resource Management, Networking, Clouds, Social Science, Evaluation



Data Archiving

Indiana University

Angela Murillo (co-PI)

Texas Tech University

Kerk Kee

Alex Olshansky



Communication & organization science



Workforce development, Sensors, operations, Semantic technologies, Communications and Outreach

University of Notre Dame

- Jarek Nabrzyski (Co-PI)
- Mary Gohsman
- Joanne Fahey
- Christina Clark
- Charles Vardeman
- Don Brower

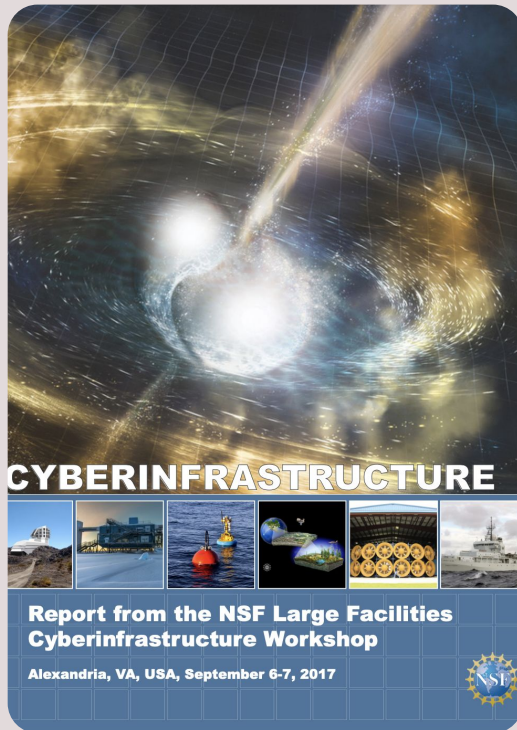


Data management, visualization, clouds, CI deployment

University of Utah

- Valerio Pascucci (Co-PI)
- Steve Petruzza
- Rob Ricci
- Giorgio Scorzelli

Cyberinfrastructure for Major Facilities Workshops

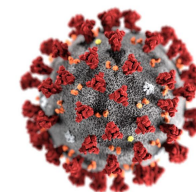


2022: Getting together, working together

- Making data FAIR
- Cloud migration
- Future of CI for MFs
- CI workforce:
 - Developing and retaining talent
 - Developing resilience

2017

2019



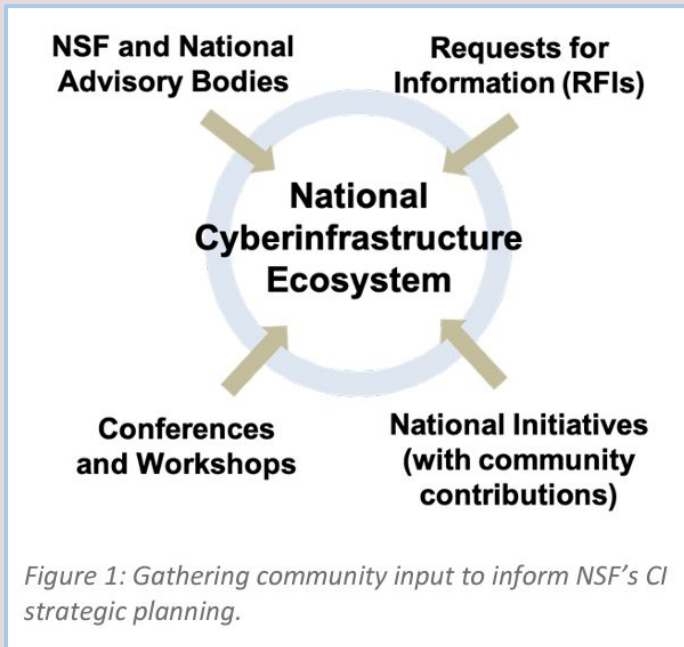
2022

2017 NSF Workshop on CI for Large Facilities

“

- **Establish a center of excellence** (following a model similar to the NSF-funded Center for Trustworthy Scientific Cyberinfrastructure, CTSC) as a resource providing expertise in CI technologies and effective practices related to large-scale facilities as they conceptualize, start up, and operate.

“

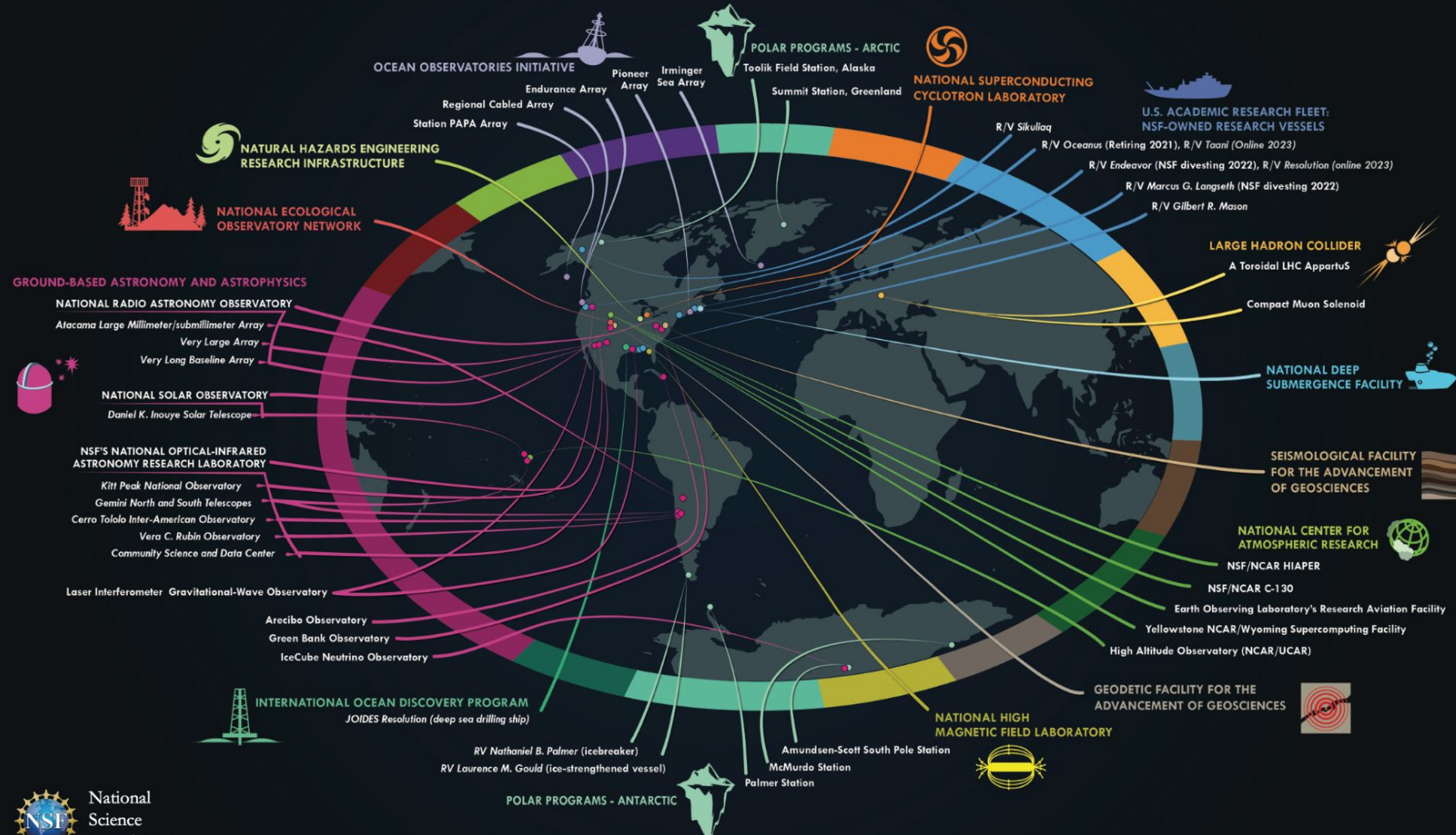


“NSF’s Blueprint for a National Cyberinfrastructure Ecosystem” April 2019



Funded by NSF under grant #1742969

MAJOR MULTIUSER FACILITIES ENABLING BASIC RESEARCH



NSF Major Facilities rely on complex **cyberinfrastructure (CI)** to transform raw data into more interoperable and integration-ready data products

49 MF + mid-scale RI-2 participants

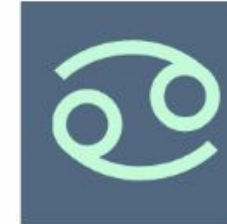
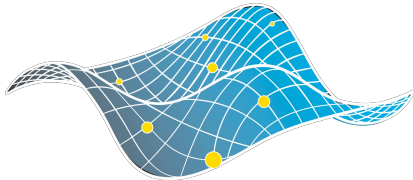
14 NSF PDs

12 Students

NSF CI Ecosystem From the CI Calling Cards



32 Participants



ResearchSOC



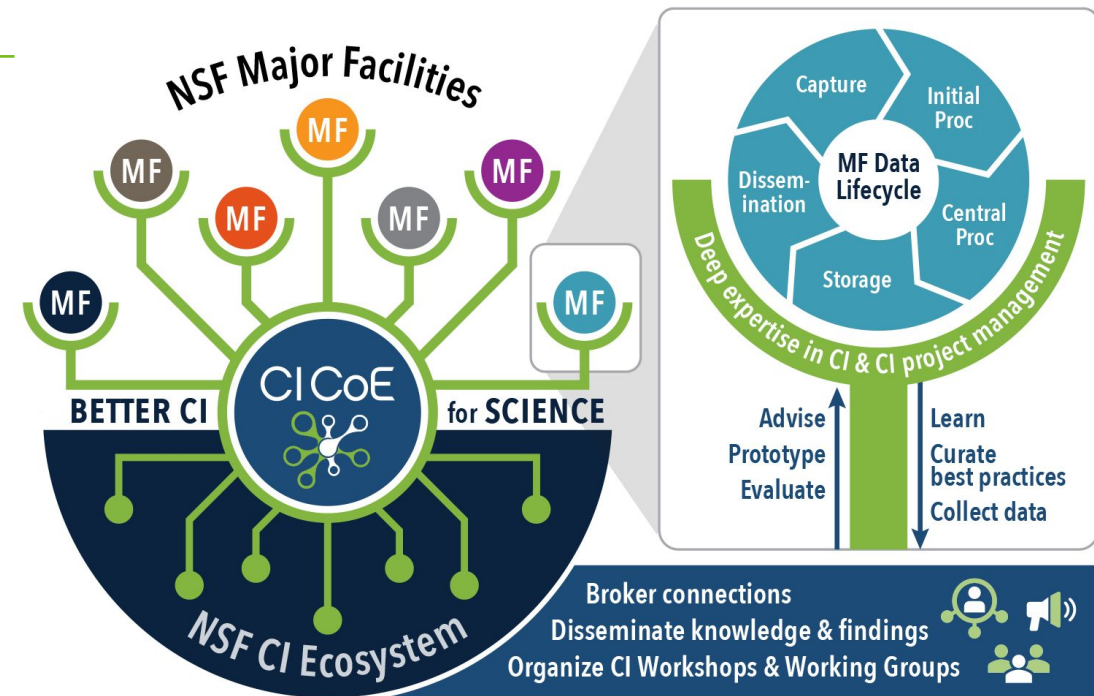
Mission



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.

Overall CI Compass Strategy

1. Recognize the expertise, experience, and mission-focus of Major Facilities
2. Contribute knowledge and expertise to the MF Data Lifecycle (DLC) CI and enhance the overall NSF CI ecosystem
3. Build expertise, not software - Leverage existing knowledge, tools, community efforts
4. Build partnerships to leverage community expertise
5. Broker connections and share knowledge, lessons learned, best practices with MFs, Partners, CI community



Partners: Trusted CI, OSG/PaTh, SGCI, EPOC, RCD Nexus, Fabric, Chameleon

Advisory Committee: *S. Anderson (LIGO), A. Bolton (NOIRLab), B. Hurwitz (UofA), M. Livny, (OSG/PaTh) V. Welch (Trusted CI) M. Zentner (SGCI)*

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

Experience in the management of CI projects

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

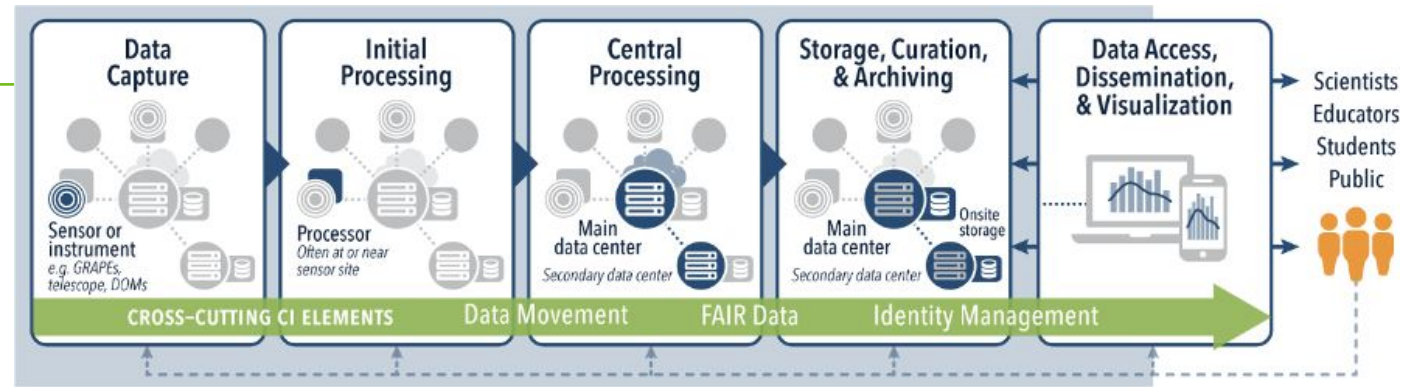
Highly collaborative, strong history of working together

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

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



CI Compass services: What we can help with.



*Evaluate CI Plans, Help Design New Solutions, Develop Proofs of Concept,
Assess Applicability/Performance of Existing Solutions, Help Leverage CI Solutions*

Data Capture: Discuss sensor data annotations, help apply community sensor data models to the MFs data, explore messaging systems, support interoperability across MFs, and connect to industry standards and to similar communities of practice within scientific domains.

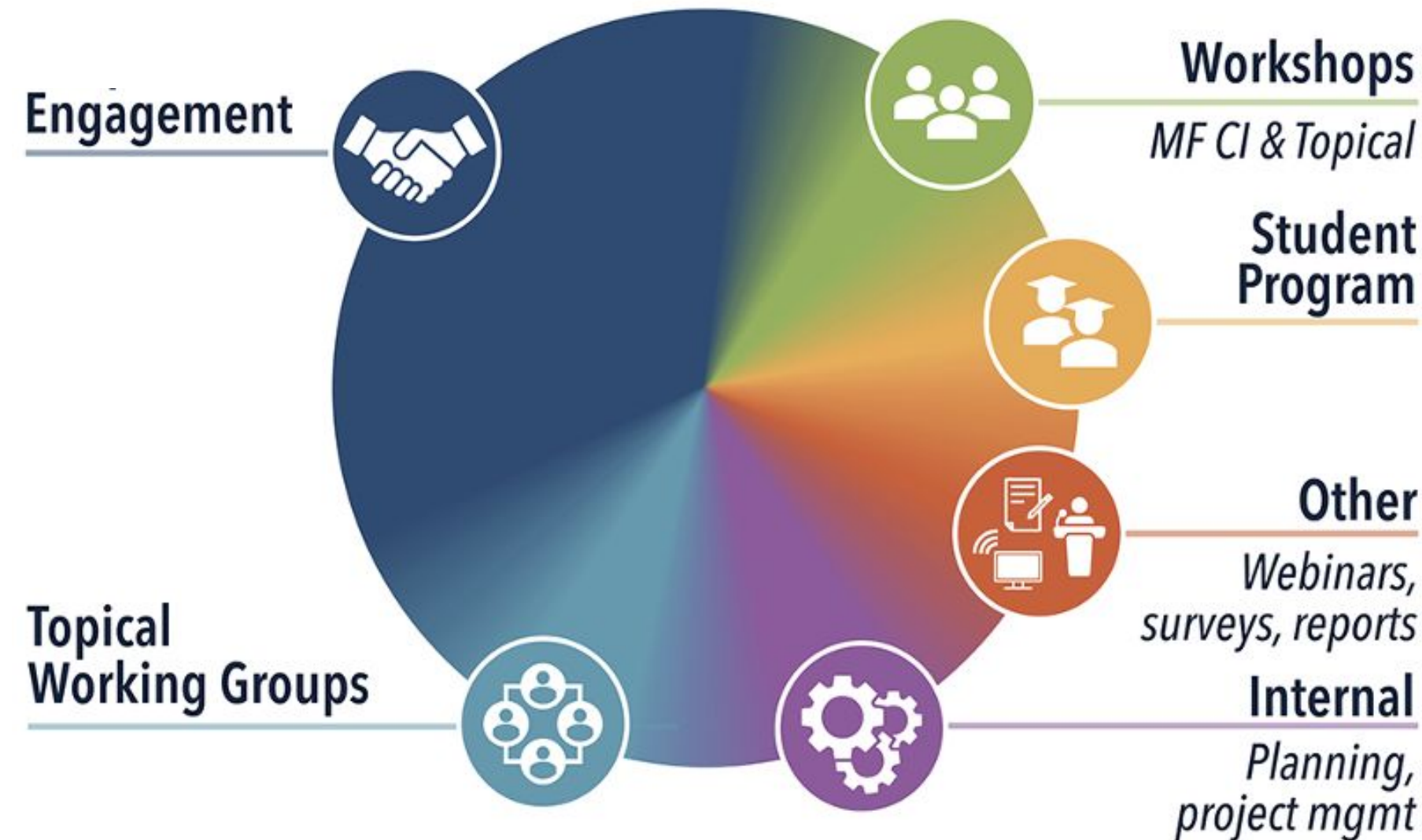
Central Processing: Help leverage existing testbeds to evaluate new software stacks and configurations during CI design/enhancements.

Data access, dissemination & visualization: Assist in designing data visualization solutions (tools, formats, proofs of concept, etc.) for enhancing data discoverability and accessibility.

Identity management: Assist in developing IdM solutions for managing user data access to data, help in using IdM for data usage reporting and tracking.

Offered services' list (non-exhaustive) is published on CI Compass website:
https://ci-compass.org/assets/453214/cicompass_services_table.pdf/

CI Compass activities: What we do.



CI Compass

Examples of Engagements

NCAR Modeling with NEON Data [Nov 2020 - ...]

Project goal: Combine NEON ecosystem data with NCAR modeling capabilities to enable new discoveries; Use cloud technologies to enable data modeling and wide community access.

CI Compass: Provide advice on cloud technologies, including containers; Hands-on-help with container testing; Consult on FAIR aspects of data management; Provide advice on data visualization with proofs of concept;

Current Working Groups (WGs):

1. Data Exchanges
2. Container/Cloud Computing
3. Data Visualization

<https://www.neonscience.org/ncar-neon-community-collaborations>

NCAR neon

NCAR-NEON Data CI Pilot: ABBY: 2019-2021

Abby Road NEON / ABBY

Time-series of Daily average Leaf-level Flux at ABBY

One integrated report compares the daily average leaf-level flux from the ABBY site with the daily average leaf-level flux from the ABBY site.

NCAR-NEON community workshop on 11/09/21

CI Compass

TECHNICAL REPORT

ci-compass.org

Making the Major Facilities Data Lifecycle FAIR

Charles F. Vandeman II
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 Guiding Principles for scientific data management and stewardship" [1] and envisages a set of first principles for research communities with respect to the management and curation of scientific data. These principles were created from the point of view that data should be *structured* in a way that the data itself is "smart data" which can be queried for information relative to the four FAIR principles. That is, given the "4 Vs" of big data of Volume, Variety,

for machine learning, and in particular, *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** and require implementation relative to a specific scientific community through community based recommendations.

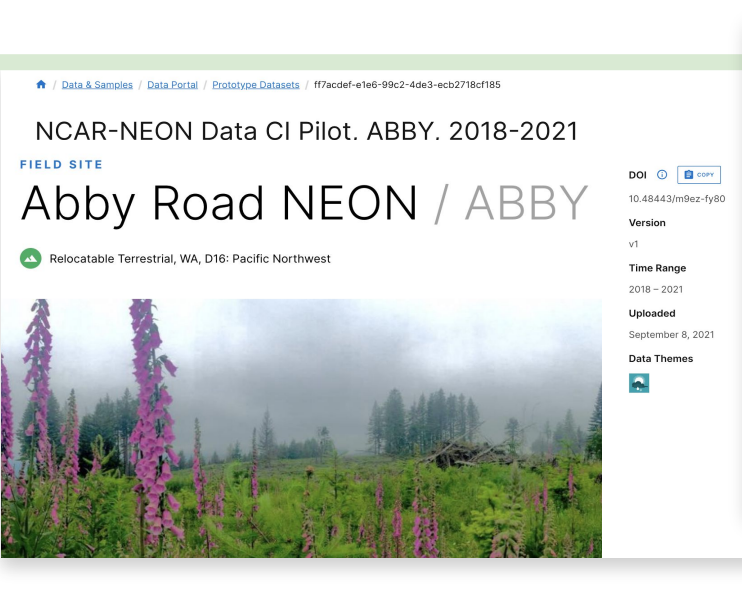
Table 1: The FAIR Guiding Principles

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier

Examples of 1-1 Engagements

2021-2022

Major Facility	NEON / NCAR	ARF/RCRV	ARF	NOIRLab	SAGE / GAGE	Arecibo
Collaborator			Trusted CI	Initiated by Trusted CI	Internet 2	EPOC, TACC Globus, UCF
Topics	Data Exchange, Clouds, Visualization	Shipboard CI/network plan review	Identity Management	Identity Management	Cloud platform design	Data management



NCAR-NEON Data CI Pilot. ABBY. 2018-2021

FIELD SITE

Abby Road NEON / ABBY

Relocatable Terrestrial, WA, D16: Pacific Northwest

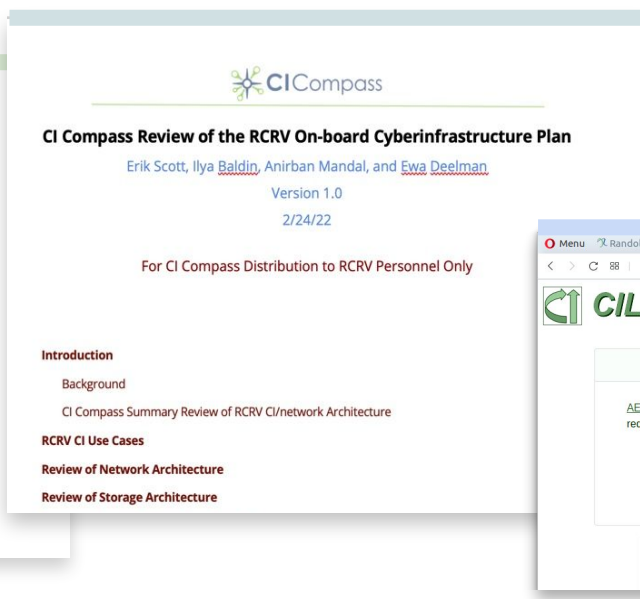
DOI: 10.48443/m9ez-fy80

Version: v1

Time Range: 2018 - 2021

Uploaded: September 8, 2021

Data Themes



CI Compass

CI Compass Review of the RCRV On-board Cyberinfrastructure Plan

Erik Scott, Ilya Baldin, Anirban Mandal, and Ewa Deelman

Version 1.0

2/24/22

For CI Compass Distribution to RCRV Personnel Only

Introduction

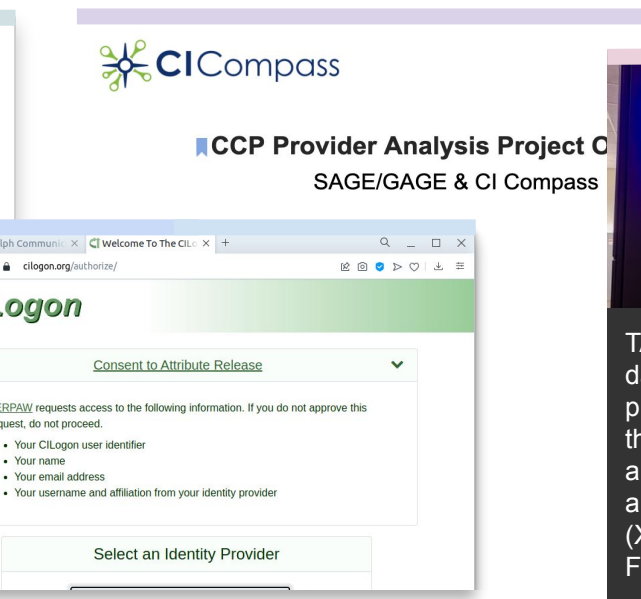
Background

CI Compass Summary Review of RCRV CI/network Architecture

RCRV CI Use Cases

Review of Network Architecture

Review of Storage Architecture



CICompass

CCP Provider Analysis Project O

SAGE/GAGE & CI Compass

Menu Randolph Community College Welcome To The CICompass

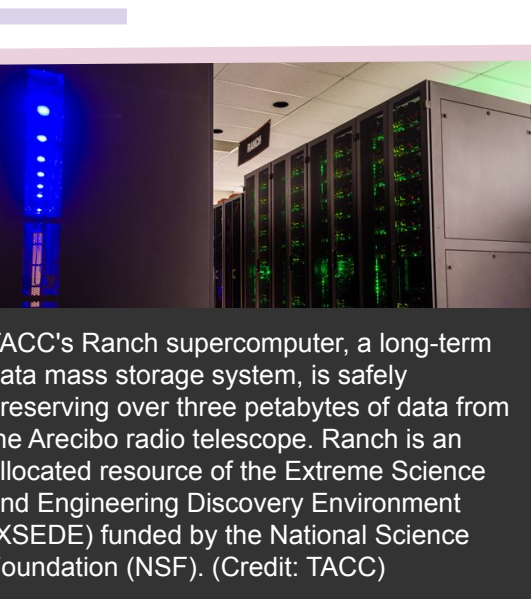
CICompass

Consent to Attribute Release

AERPAW requests access to the following information. If you do not approve this request, do not proceed.

- Your CICompass user identifier
- Your name
- Your email address
- Your username and affiliation from your identity provider

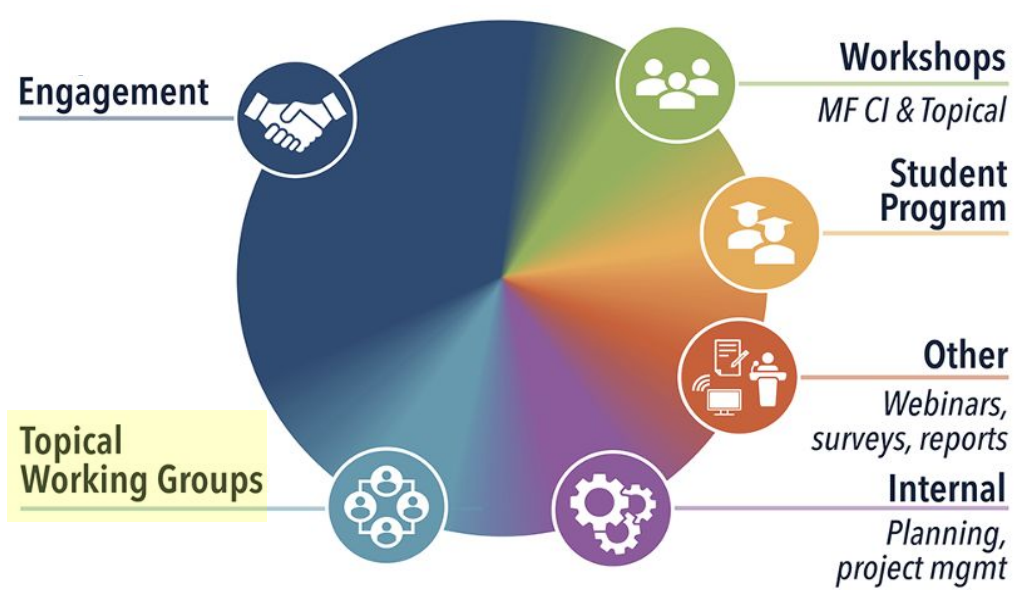
Select an Identity Provider



TACC's Ranch supercomputer, a long-term data mass storage system, is safely preserving over three petabytes of data from the Arecibo radio telescope. Ranch is an allocated resource of the Extreme Science and Engineering Discovery Environment (XSEDE) funded by the National Science Foundation (NSF). (Credit: TACC)

For details of these engagements, please visit our website <https://ci-compass.org/> and view [Anirban's webinar](#) from February 2022

CI Compass activities: Topical Working Groups



Identity Management Topical WG

Disseminate IdM information

- Quarterly meetings with speakers and discussions on topics relevant to MFs: e.g. CILogon
- Issues of identifying data usage and enabling reporting

Cloud Infrastructure Topical WG

- Understand the current practices for clouds used by MFs
- Research alternative solutions and keep up to date with emerging cloud technologies
- Develop a general set of best practices that can inform the MFs

Send email to
engage@ci-compass.org to
participate in any of these or
future Topical Working Groups
or propose new ones!

Organizing a panel tomorrow!

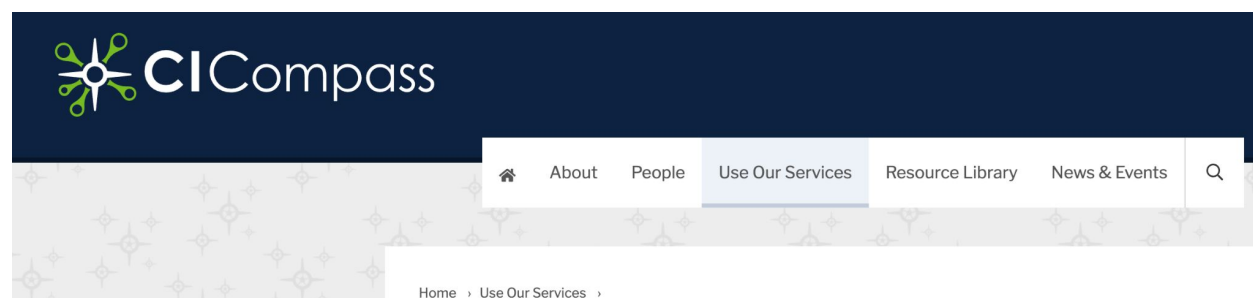
Let's Work Together

To explore potential opportunities for an engagement with CI Compass regarding a CI issue you might have, please email

engage@ci-compass.org

To respond to our **Call for Engagement**, please click the following link:

<https://ci-compass.org/services/apply-for-engagement-with-ci-compass/>



Apply for Engagement with CI Compass

One of CI Compass's core activities is conducting engagements with NSF Major Facilities (MFs) and large-scale projects at all stages of their lifecycle, from proposal to construction, operations, upgrades. CI Compass conducts one-on-one engagements with MFs to help them diagnose problems they are experiencing with their CI, research existing and new solutions, assess the feasibility of possible solutions, provide advice on adopting possible solutions, develop prototype solutions to the problems, and assist with solution implementation.

Engagement teams include staff from the MF and CI Compass. Engagements can range from short discussions to activities lasting 3 to 6 month period. We also support engagements that last for a longer

Please feel free to reach out to us, we are here to work with you!

Student Internship Program (Pilot)

Undergraduate Students in CS

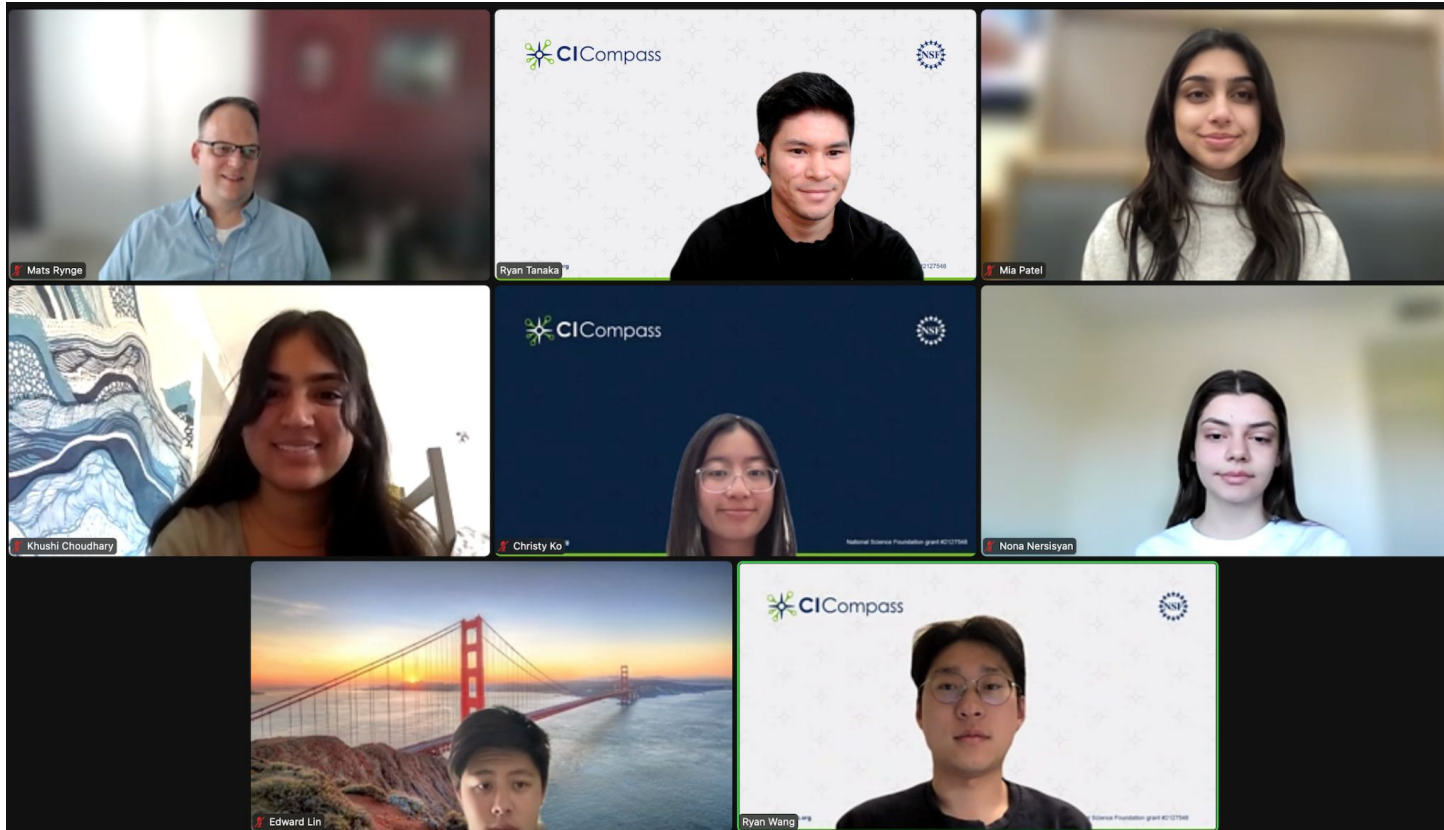
- 5 at USC, 1 at UND
- Spring with CI Compass directed research
- Summer Internships: working on projects, ultimately with MFs

Led by Angela Murillo, Laura Christopherson, Jarek Nabrzyski

Directed by Ryan Tanaka, Loïc Pottier

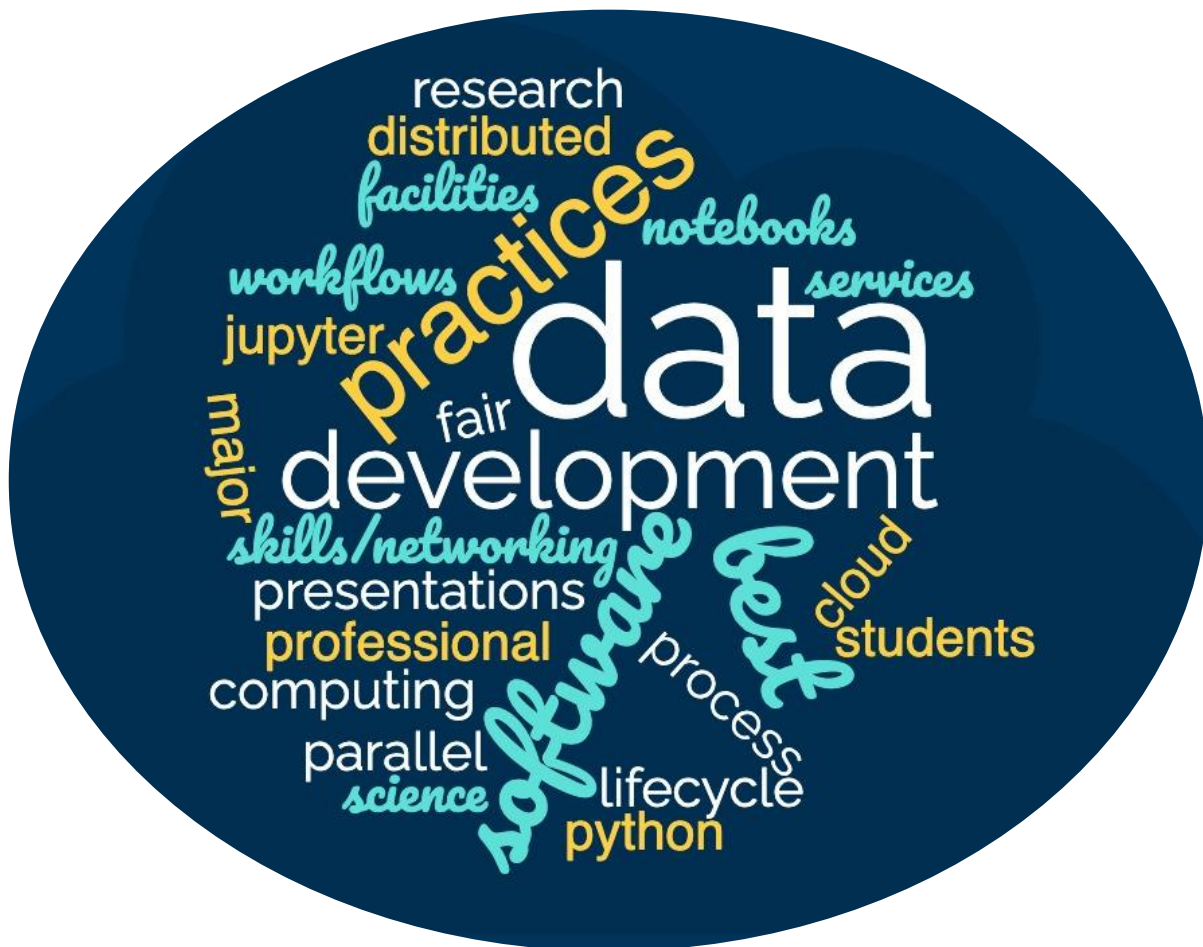
Check out their calling cards!

February 28, 2022, Topic: **Software Best Practices II, Containers!**



Practical

Research



Lesson Topics



Research lectures notes

Help us connect with you and your programs: contact@ci-compass.org


Answers to Q2 on Calling Cards


Time (PST)	Session Title
8:00-8:30	Workshop Welcome and Overview of CI Compass
8:30-9:15	NSF's Perspectives on Cloud Computing
9:15-9:35	Thematic Lightning Talks (Cloud, CI Challenges)
9:35-9:55	Brainstorming: Developing and Retaining a Vibrant Team of Workers (for day 2)
9:55-10:00	Group Zoom picture
10:00-10:30	Break
10:30-11:00	NSF CI Ecosystem Lightning Talks
11:00-12:00	Developing Resilience and Managing Uncertainty during the Pandemic (Breakout)
12:00-1:30	Lunch Break
1:30-3:15	Making the Major Facilities Data Lifecycle FAIR to Provide AI-Ready Data (Group discussion)
	Office Hours (remote only)
3:15-3:45	Kerk Kee, Angela Murillo, Jarek Nabrzyski, Charles Vardeman
4:00-6pm	Reception - in person

Please contribute notes, comments <https://tinyurl.com/ci4mfs2022>


Time (PST)	Session Title
8:00-10:00	Cyberinfrastructure for Major Facilities: Challenges and a path forward (breakout)
10:00-10:30	Break
10:30-12:00	The March Toward the Clouds: MF Perspectives (Topical Cloud WG)
12:00-1:20	Lunch Break
12:45-1:20	Office Hours (online only) Ilya Baldin, Ewa Deelman, Anirban Mandal, Valerio Pascucci
1:20- 2:20	Brainstorming: Developing and Retaining a Vibrant Team of Workers (breakout)
2:20-3:00	NSF CI Ecosystem Lightning Talks
3:00-3:05	Closing
3:05 - 4:00	Social Coffee Hour - in person

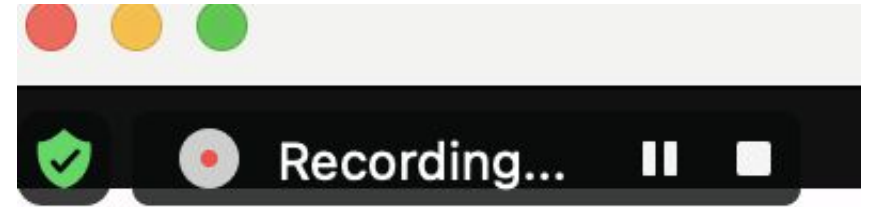
Wrong time zone?


Whova Guides ▾ Organizing your own event?


2022 NSF Cyberinfrastructure for Major Facilities Workshop: Getting Together, Working Together

Redondo Beach, CA (View map) Mar 1 - 2, 2022 | Displaying in the event's time: 11:27 AM (PST)





Recording

- The whole workshop is being recorded
- Only formal presentations will be edited and posted online
- Other recordings will be used to clarify notes

Virtual Breakout Group Assignments

(For all breakouts **except those for the Brainstorming session.**)

Breakout Room 1 Brad Barber Devan Bougie Doug Ertz Jeannette Dopheide Ken Feldman Sharon Broude Geva Margaret Levenstein Kevin Porter Paul Redfern Jennifer Schopf	Breakout Room 2 Maggie Benoit Adam Brazier Jerry Carter Lauren Clay Jeannette Jackson David Kratz Greg Madden Brandi Murphy Joy Pauschke John West	Breakout Room 3 Brian Dobbins James Edson Jeremy Fischer Stephen Jacobs Bogdan Mihaila Giri Prakash Roland Roberts Dan Stanzione Alex Szalay Von Welch	Breakout Room 4 Kathy Benninger Bruce Berriman Alisdair Davey Lee Ellett James Holik Miron Livny Margaret Martonosi Craig Risien David Schultz Adam Shepherd
Breakout Room 5 Christopher Cameron Robert Casey Dru Clark Josh Drake Jeff Glatstein Jeff Hoch Dawn Lenz Richard Oram Vladimir Papitashvili Benedikt Reidel	Breakout Room 6 Kay Avila Shafaq Chaudhry Chris Davis Clint Dawson David Halstead Christine Laney Sarah Ruth Mike Simpson Bob Tawa	Breakout Room 7 Steve Barnet Philip Gates John Haverlack Mark Krenz Allen Pope Kelli Shute Laura Stolp Werner Sun Ross Thomson Michael Zentner	Breakout Room 8 Peter Beckman Gordon Bonan Patrick Brady Mark W. Coles Douglas Fils Terry Fleury Nirav Merchant Eric Palm John Zage

**During
breakouts,
please go to
your
assigned
room**

**We have note
takers, but
please add
your
thoughts!**

**[tinyurl.com/
ci4mfs2022](https://tinyurl.com/ci4mfs2022)**

***No report
back right
after breakout***

The questions will be the same for each breakout room

Workshop materials

Welcome packet, Notes, Slides, Calling cards:
<https://tinyurl.com/ci4mfs2022>

Please contribute to discussions and notes!

To ask questions or make a comment

Please raise your hand in Zoom or use Zoom chat

- The moderator will call on you

Please **do not use** zoom chat for other discussions as this makes it hard to see the questions.

To chat with other participants



2022 NSF Cyberinfrastructure for Major Facilities Workshop: Getting Together, Working Together

Redondo Beach, CA (View map) Mar 1 - 2, 2022

- Home
- Agenda
- Attendees**
- Community
- Messages

Messages

Search thread title...



Whova Team

Hi Ewa, Are you interested in making your ever



Soyeon Park



2022 NSF Cyberinfrastructure for Major Facilities Workshop: Getting Together, Working Together

Redondo Beach, CA (View map) Mar 1 - 2, 2022

- Home
- Agenda
- Attendees
- Community** 1
- Messages
- Win a Prize
- PhotoContest
- Leaderboard
- Resources

Add a topic or social group

Search...

All Topics Followed New Topics

Mark all as read



Meet-ups & Virtual Meets

Suggest a dinner together, a morning ru...
0 activities • 10/21/2021



Ask Organizers Anything

Have any question for the organizer? As...
0 questions • 10/21/2021



Organizer Announcements

Check all announcements sent by the...
0 announcements • 10/21/2021



CI for Major Facilities Workshop General


We are looking forward to your...
2 messages

You can view session in Whova or popped out into Zoom

1

To get Help

2


2022 NSF Cyberinfrastructure for Major Facilities Workshop: Getting Together, Working Together
 Redondo Beach, CA (View map) Mar 1 - 2, 2022

Home
 Agenda
 Attendees
Community
 Messages
 Win a Prize
 PhotoContest
 Leaderboard
 Resources
 Feedback to Whova

Add a topic or social group
 Search...
 All Topics
 Followed
 New Topics
 Meet-ups & Virtual Meets
 Suggest a dinner together, a morning ru...
 0/2/2021
Ask Organizers Anything
 Have any question for the organizer? As...
 questions • 10/21/2021
 Organizer Announcements
 Check all announcements sent by the...
 0 announcements • 10/21/2021
 Break the Ice!
 Introduce yourself to everyone, so other...
 7 introductions • 2/24/2022
 Following

Ask Organizers Anything
 Description: Have any question for the organizer?

3

Email ci4mfs@ci-compass.org

Home
 Agenda
Attendees
 Messages
 Win a Prize
 PhotoContest
 Leaderboard
 Resources

Attendees
 All
 Recommended
 Bookmarked
Categories
 117 attendees total

Ewa Deelman
 Research Professor, Computer Science Depart...

In-person Attendees
 34 people

Remote Attendees
 96 people

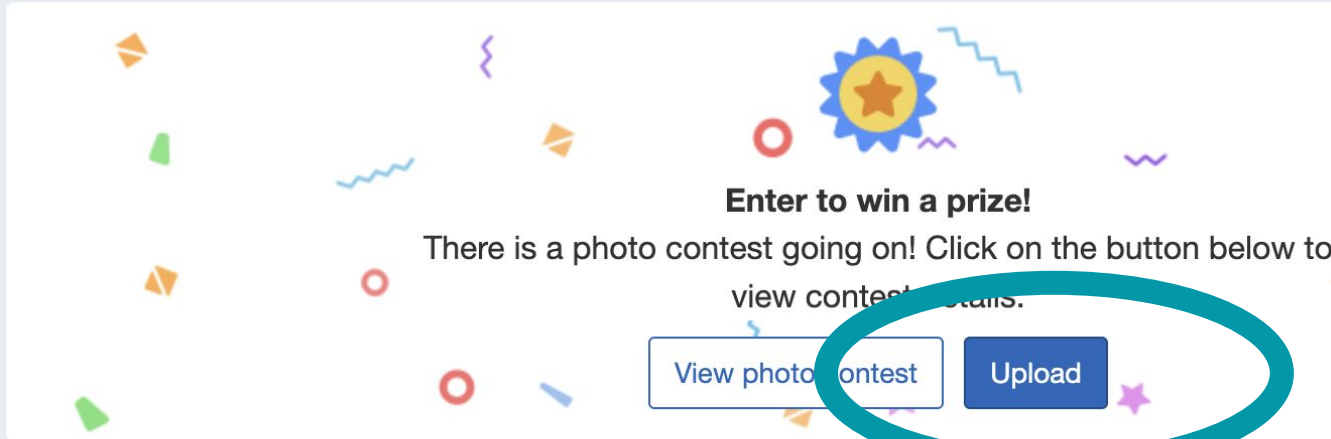
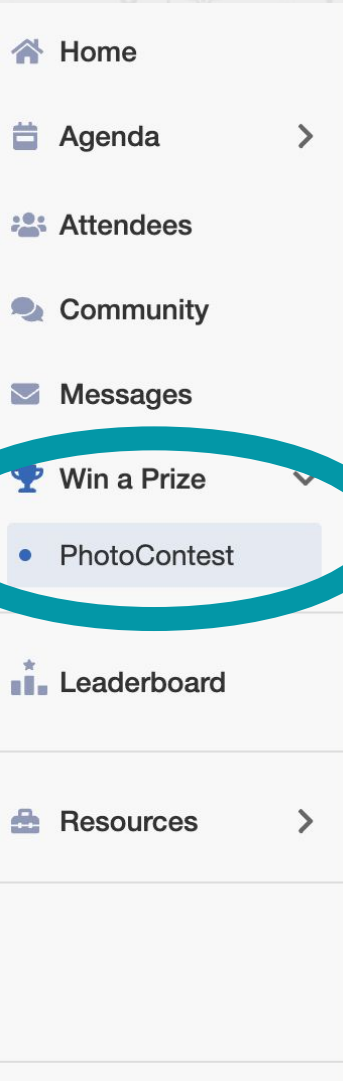
Speakers
 43 people

Organizers
 13 people

CI Compass
 24 people

Virtual Support
 4 people

Best Home Office Companion Photo Contest

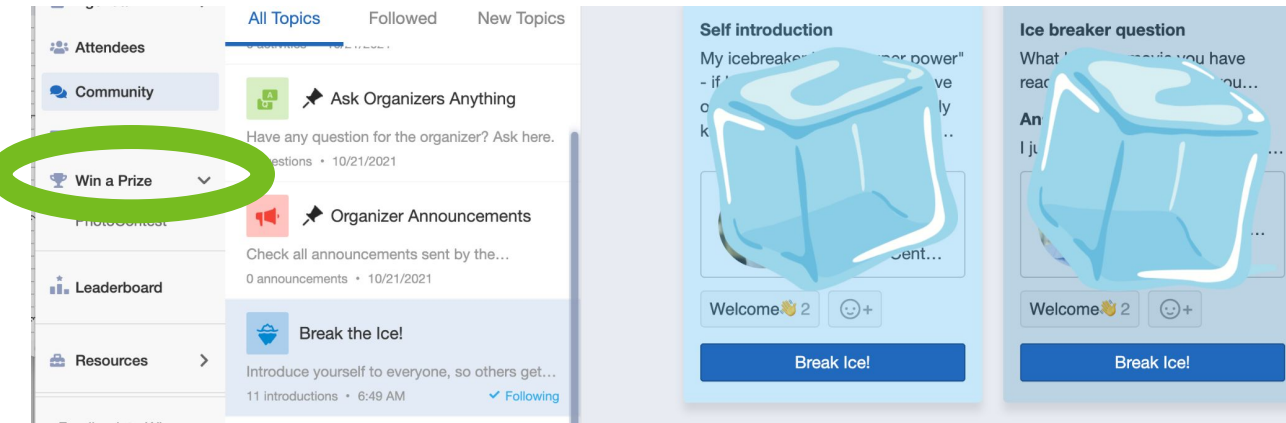


*The 3 photos with the most votes win!
Prizes are a gift certificate,
a box of chocolates, or a CI
Compass T-shirt.*

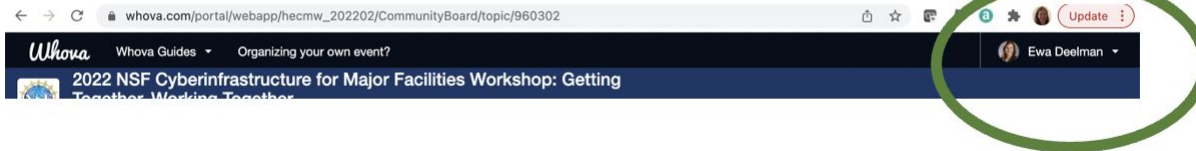


- **Before COB Tue**, Post your scaly/feathery/furry work-from-home companion photo
- **Before lunch break on Wed**: Vote for Photos you like best: click the thumbs up icon
- **Winners announced after lunch**

Whova Ice breakers



Change profile?



Questions?

Self introduction

Hi all. My ice breaker is that I literally just got off a short 10 day cruise on an actual ice breaker. Granted no ice was broken durin...

