Building the Next Generation of Cyberinfrastructure Professionals

ANGELA P. MURILLO[1], DON BROWER[2], SAROWAR HOSSAIN[1], KERK KEE[3], ANIRBAN MANDAL[4]
JAREK NABRZYSKI[2], ERIK SCOTT[4], NICOLE VIRDONE[5], EWA DEELMAN[5]

[1]Indiana University-Indianapolis, Luddy School of Informatics, Computing, and Engineering, [2]University of Notre Dame, Center for Research Computing,
[3]Texas Tech University, [4]University of North Carolina, RENCI, [5]University of Southern California, Information Science Institutes

WHAT IS CI COMPASS?

CI Compass is a National Science Foundation Cyberinfrastructure Center of Excellence

- We provide expertise and active support to cyberinfrastructure (CI) practitioners at NSF Major Facilities (MFs) to accelerate the data lifecycle (DLC) and ensure the integrity and effectiveness of the CI upon which research and discovery depend.



Photo 1: Maximilien Brice, CERN
Photo 2: RCRV Taani - Photo provided by Oregon State University
Photo 3: NAIC - Arecibo Observatory

CI COMPASS STUDENT FELLOWSHIP PROGRAM (CICF)

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

CICF provides undergraduate student fellows the opportunity to:

- Learn about CI development and MFs
- Develop CI-related skill sets important to the work of MFs
- Engage with CI Compass and MF personnel through a virtual training and research program
- Participate in an optional/invited summer program to apply the skills learnt for a particular MF project

SPRING PROGRAM STRUCTURE

The virtual Spring Training Program has two components:

- *Technical Program*
  - Learn technical skills relevant to CI (basic software development, programming for scientists, systems, ML/AI relevant to CI and MFs)
- *Research Program*
  - Learn about the importance and context of MFs, the related data lifecycle, and CI
  - Engage with guest speakers from MFs and the greater CI community

CICF is free to students and the Spring Program can be taken for course credit, depending on their institutions requirements.

To learn more about the CI Compass Fellowship program visit, https://ci-compass.org/student-fellowships/ or email us at cicf@ci-compass.org.

SPRING 2023 PROGRAM

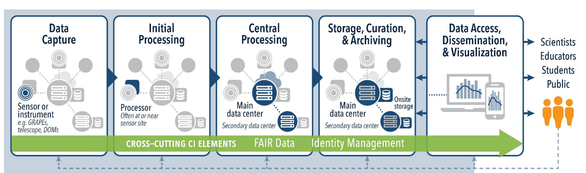
(12-weeks, virtual)

Week	Technical Training Program	Research Training Program
Week 1	Orientation, Linux/Unix Shell, Terminal	Cyberinfrastructure, Major Facilities, the Data Lifecycle. Guest Speaker from the Natural Hazards Engineering Research Infrastructure (NHERI)
Week 2	Introduction to Python Programming	Research Data Management. Guest Speaker from ORCID
Week 3	Python Programming, Jupyter Notebooks, Python Data Analysis Packages	Research Computing. Guest speaker from the 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	Guest Speaker from the 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	Guest Speaker from the National High Magnetic Field Laboratory (MagLab)	Group Presentations Day 1
Week 12	Machine Learning/AI	Group Presentations Day 2

SPRING 2023 PRESENTATIONS

In groups, students research a specific MF to learn about its science mission, CI, and data lifecycle.

- During the Spring 2023 Program, students researched Texas Advanced Computing Center (TACC), National High Magnetic Field Laboratory (MagLab), Ocean Observatories Initiative. (OOI), and Cornell High Energy Synchrotron Source (CHESS).
- Students reviewed websites and published documents, and met with MF staff to learn about the MF they choose.
- At the end of the Spring Program, students presented their work to demonstrate the knowledge gained.



SUMMER PROGRAM

CICF student fellows may apply for a paid, optional, and invited Summer Program to:

- Employ the relevant skills they learned during the Spring Program
- Gain experience applying their technical skill sets on projects relevant to MFs and CI

2022 CICF Pilot Summer Program

- Three students at USC worked with Chameleon Cloud, the Pegasus Workflow Management System, and HTCondor, to test and classify lake zooplankton to learn about the impacts of environmental changes.
- Their work was presented at eScience 2022 (see QR Code).

2023 CICF Summer Program

- Five students at NCAR/NEON and 2 students at OOI.

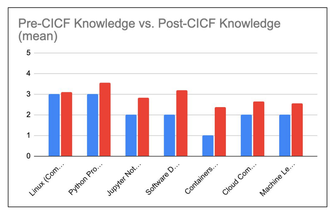
CICF IMPACT

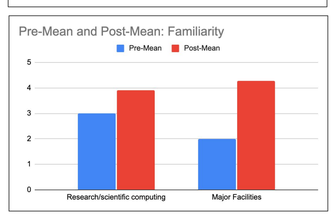
2022 CICF Pilot Program

- Six student fellows participated from two institutions (University of Notre Dame and the University of Southern California).

2023 CICF Program

- Fourteen students from nine institutions, including Indiana University, Louisiana State University, University of Iowa, University of Alabama, and Arizona State.
- Six majors and spanned from 1st years through 5th years.
- Eight female and six male students; three first-generation.





Testimonials from Student Fellows

"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."

WAYS TO GET INVOLVED

- Provide feedback on the program,
- Participate as a guest speaker during the Spring Program,
- Work with the CI Compass team to provide summer opportunities for students.

This project is supported by the National Science Foundation Grant no: 2127548