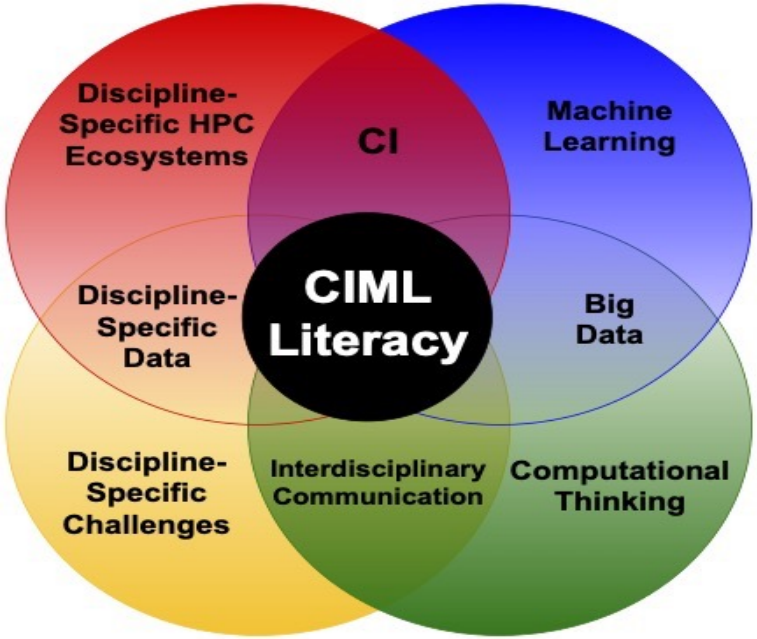
CIML Day2 Welcome, Logistics

HPC & Parallel Concepts

June 27, 2023

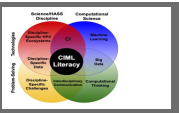


Mary Thomas
(SDSC)

NSF Award 1928224

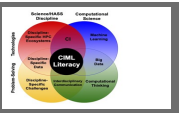
Logistics

- Wednesday, June 20th was “Prep Day”
  - We focussed on making sure you can use your training account connect to Expanse, run jobs, launch notebooks, access the Expanse Portal, etc.
- Web Site:
  - https://www.sdsc.edu/event_items/202306_cimlsi.html
  - https://na.eventscloud.com/website/50410/
- GitHub: https://github.com/ciml-org/ciml-summer-institute-2022



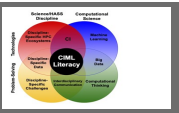
What is CIML?

- NSF CyberTraining Grant: *Developing a Best Practices Training Program in Cyberinfrastructure-Enabled Machine Learning Research (CIML)*
- Objectives: ***Scalable Machine Learning***
  - To *facilitate* researchers and educators who are using machine learning (ML) and big data analytics methods for their domain specific applications or instructional material
  - To develop a *community* of machine learning and data analytics CI Users (CIU) and Contributors (CIC) who actively contribute to the training material repository and incorporate the materials into their projects and courses.
  - *Synthesize* the training material into a domain independent CIML workflow system that can be used for creating applications that run on the NSF HPC ecosystem.
  - To create generalized machine learning training and project materials that run on large-scale NSF funded cyberinfrastructure resources such as XSEDE



Day2: HPC and CI Architecture

- Day 1 covered
  - Connecting to Expanse
    - command line
    - portal
  - Interactive Computing
    - Interactive nodes
    - Notebooks
  - Modules
  - Account Management
- Day 2:
  - HPC concepts and architectures
  - Compiling and Linking Code
  - Running Jobs
  - Hands-on Examples

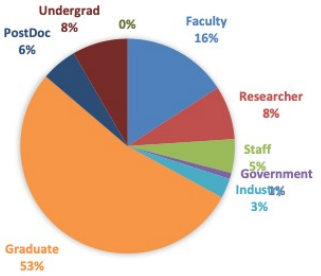


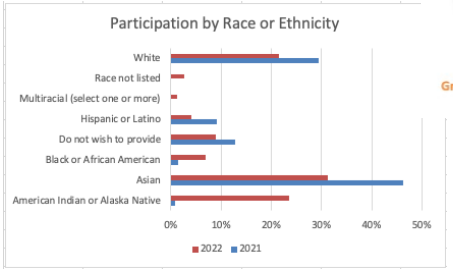
3rd annual CIML institute

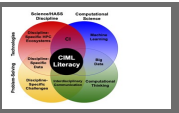
- 2021, 2022 virtual
- 2023: in-person
- Fewer applications, but the same number of participants accepte: ~46 + staff
- Training impacting underrepresented STEM communities
  - Improved participation by gender and Institution: CIML'23: 67% MSI Appl's

Gender	Count	2022	2021
Man	234	67.8%	68.6%
Woman	94	27.2%	20.1%
Do not wish to provide	14	4.1%	4.3%
Nonbinary	3	0.9%	

Participation by Gender: some improvement

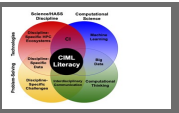






Day 2, June 27 2023, Agenda

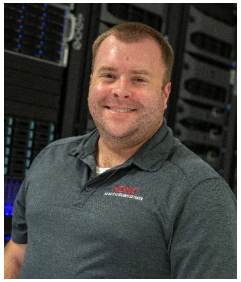
TIME (Pacific time)	*TOPIC*	*PRESENTER*
8:00 AM – 8:30 AM	***Light Breakfast & Check-in***	
8:30 AM – 9:30 AM	2.1 Welcome and Introductions	Mary Thomas
9:30 AM - 9:40 AM	Break	
9:40 AM – 11:00 AM	2.2 Parallel Computing Concepts	Robert Sinkovits
11:00 AM – 11:10 PM	Break	
11:10 AM – 12:30 PM	2.3 Running Batch Jobs on SDSC Systems	Marty Kandes
12:30 PM – 1:30 PM	Lunch	
1:30 PM - 2:50 PM	2.4 Data Management and File Systems	Mahidhar Tatineni
2:50 PM – 3:00 PM	Break	
3:00 PM – 4:30 PM	2.5 GPU Computing - Hardware architecture and software infrastructure	Andreas Goetz
4:30 PM – 5:00 PM	Q&A, Wrap-up	
5:00 PM – 7:00 PM	Evening Reception - 15th Floor, the Village	



CIML Instructors

Andreas Goetz, Ph.D.
Director of Computational Chemistry Laboratory

Mai Nguyen, Ph.D.
Lead for Data Analytics



Marty Kandes, Ph.D.
Computational and Data Science Research Specialist

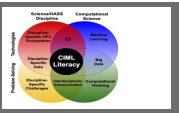
Mary Thomas, Ph.D.
Computational Data Scientists, HPC Trainer

Mahidhar Tatineni, Ph.D
Director, User Support Services

Peter Rose, Ph.D.
Director of Structural Bioinformatics Laboratory

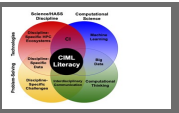
Paul Rodriguez, Ph.D.
Computational Data Scientist

Robert Sinkovits, Ph.D.
Director of Education and Training



Let’s get to know each other

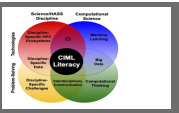
1. Name
2. Institution/Company & Department
3. How do you like to spend your time when not at work?
4. What have you binged watched?



Basic Information

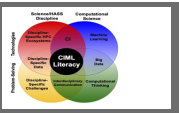
- Expanse User Guide:
  - https://www.sdsc.edu/support/user_guides/expanse.html
- After the institute is over, you need to have an Expanse account in order to access the system. There are a few ways to do this:
  - Submit a proposal through the ACCESS system (https://access-ci.org/)
  - PIs on an active allocation can add you to their allocation (if you are collaborators working on the same project).
  - Request a trial account, contact consult@sdsc.edu

- Online repo and information:
  - https://github.com/sdsc-hpc-training-org/expanse-101
  - https://hpc-training.sdsc.edu/expanse-101/

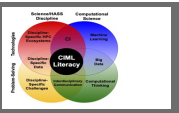


Resources

- Expanse User Guide
  - https://www.sdsc.edu/support/user_guides/expanse.html
- GitHub Repo for this webinar: clone code examples for this tutorial – clone example code:
  - https://github.com/sdsc-hpc-training-org/expanse-101
- SDSC Training Resources
  - https://www.sdsc.edu/education_and_training/training
  - https://github.com/sdsc-hpc-training/webinars
- XSEDE Training Resources



We know you will all have a great workshop!



Thank You!

If you have problems, please contact consult@sdsc.edu