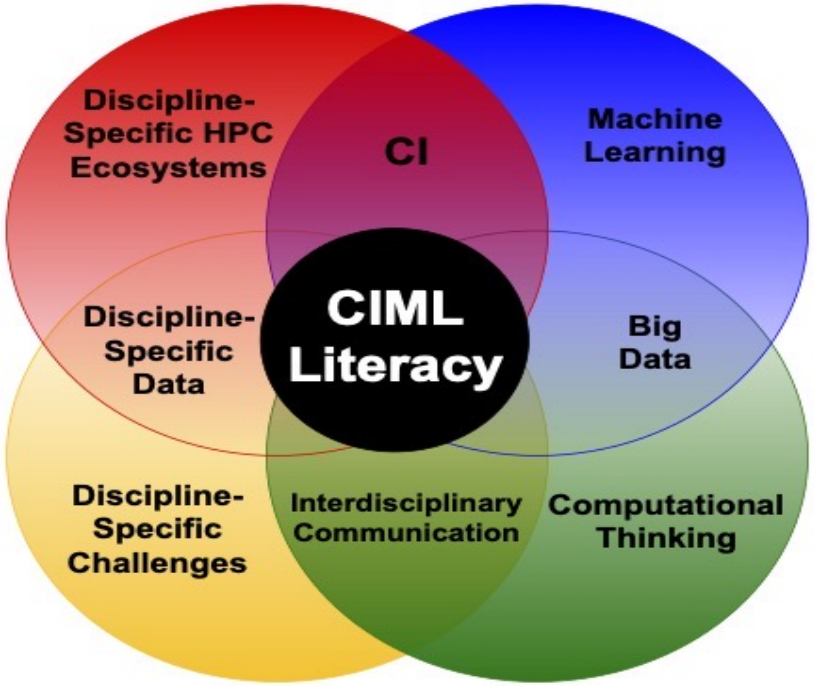
CIML Day2 Welcome and Logistics

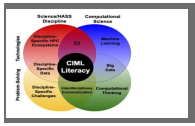
June 27, 2022



Mary Thomas
(SDSC)

Day2: HPC and CI Architecture

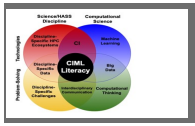
- Day 1 covered
  - Connecting to Expanse
    - command line
    - portal
  - Interactive Computing
    - Interactive nodes
    - Notebooks
  - Modules
  - Account Management
- Day 2:
  - HPC concepts and architectures
  - Compiling and Linking Code
  - Running Jobs
  - Hands-on Examples



Welcome to the 2nd CIML Summer Institute!

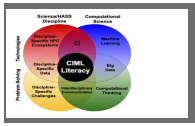
Today is “Day 1”; last week we met for ”PreDay:”

- We’ll focus on making sure you can connect to Expanse, run jobs, launch notebooks so that next week goes smoothly
- We will use Slack for chatting/communicating
- We will use Zoom for
  - all presentations and group discussions
  - Breakout rooms for hands-on sessions
  - *To avoid Zoom fatigue, we’ll have lots of breaks*
- Please be on time so we can stay on schedule
- WebSite: https://na.eventscloud.com/website/22773/home/
- GitHub: https://github.com/ciml-org/ciml-summer-institute-2022



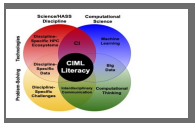
What is CIML?

- 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



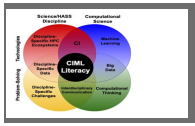
Logistics

- Wednesday, June 22nd was “Prep Day”
- We focussed on making sure you can connect to Expanse, run jobs, launch notebooks, access the Expanse Portal.
- We will use Slack for: chatting/communicating
- We will use Zoom for:
  - All presentations and group discussions
  - Breakout rooms for hands-on sessions
  - To avoid Zoom fatigue, we’ll have lots of breaks
- Web Site: https://na.eventscloud.com/website/35357/home/
- GitHub: https://github.com/ciml-org/ciml-summer-institute-2022



Day 2 Agenda:
HPC/Cyberinfrastructure Concepts (Mon, 06/27/21)

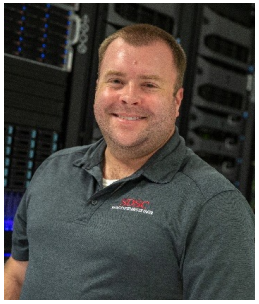
TIME (Pacific time)	TOPIC	PRESENTER
8:00 AM - 8:05 AM	2.1 Quick Welcome	Mary Thomas
8:05 AM - 9:05 AM	2.2 Introduction HPC/Cyberinfrastructure	Robert Sinkovits
9:05 AM - 10:05 AM	2.3 CPU Computing - Hardware, Architecture, Software, Running Jobs	Mary Thomas
10:05 AM - 11:35 AM	2.4 Data Management and File Systems	Marty Kandes
11:35 AM - 12:05 PM	Break/Lunch	
12:05 PM - 1:50 PM	2.5 GPU Computing - Hardware architecture and software infrastructure	Andy Goetz
12:05 PM - 1:50 PM	[Q&A, Wrap-up]	



CIML Instructors

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

Mai Nguyen, Ph.D.
Lead for Data Analytics



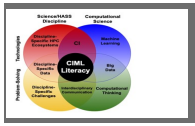
Marty Kandes, Ph.D.
Computational and Data Science Research Specialist

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

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

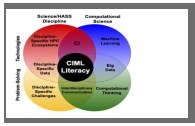
Paul Rodriguez, Ph.D.
Computational Data Scientist

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?

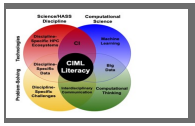


Basic Information

- Expanse User Guide:
  - https://www.sdsc.edu/support/user_guides/expanse.html

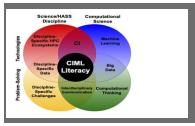
- You need to have an Expanse account in order to access the system. There are a few ways to do this:
  - Submit a proposal through the XSEDE Allocation Request System
  - PI on an active allocation can add you to their allocation (if you are collaborators working on the same project).
  - Request a trial account, instructions @ https://portal.xsede.org/allocations/startup.

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

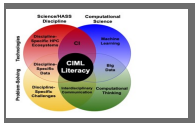


Resources

- 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 hope you all have a great workshop!



Thank You!

If you have problems, please contact help@xsede.org