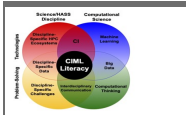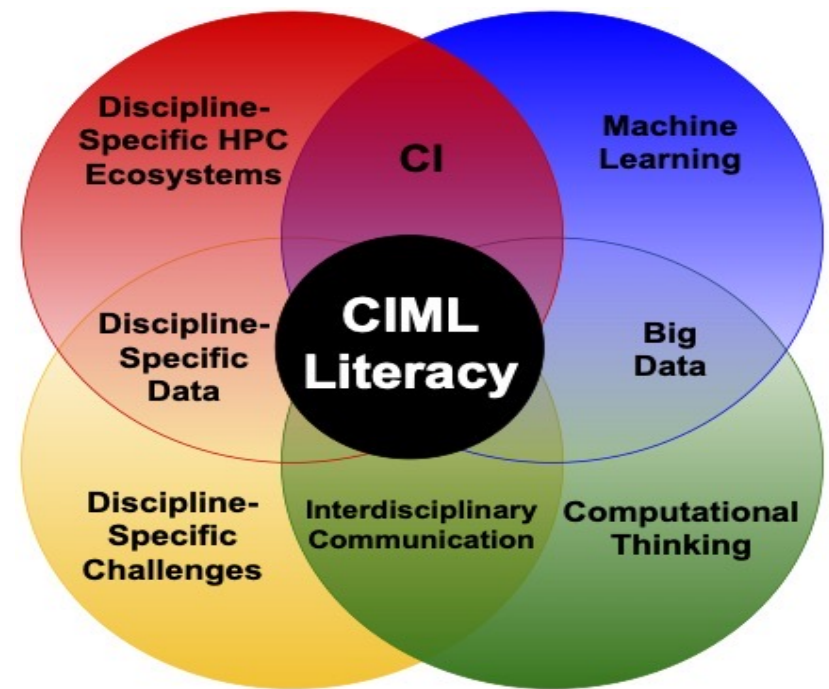# CIML Summer Institute, Day 1
Prep Day: Logistics and Introductions
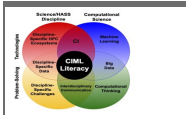
June 22, 2022

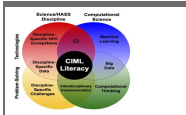Mary Thomas

San Diego Supercomputer Center

Welcome to the 2nd CIML Summer Institute!

- Today is "Prep Day"

- 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

Day1: Prep Day Agenda

9:00 AM - 9:20 AM	**1.1. Welcome & Orientation** Mary Thomas, Computational Data Scientist & Director of the CIML Summer Institute
9:20 AM – 9:40 AM	**1.2 Accounts, Login, Environment, Running Jobs and Logging into Expanse User Portal** Mary Thomas, Computational Data Scientist & Director of the CIML Summer Institute
9:50 AM – 10:10 AM	**1.3 Running Jupyter Notebooks on Expanse** Marty Kandes, Computational and Data Science Research Specialist
10:10 AM – 10:30 AM	**Q&A wrap up**
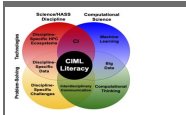
CIML Instructors

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 Scientist, HPC Training Lead

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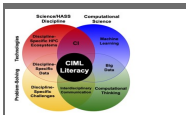*Director of Structural Bioinformatics Laboratory*

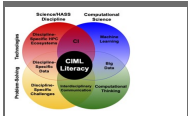**Paul Rodriguez, Ph.D.**
Computational Data Scientist

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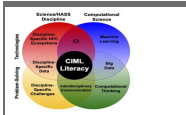*Director of Education and Training*

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?
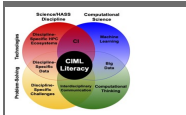
Basic Information

- 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://education.sdsc.edu/training/interactive/
- XSEDE Training Resources
  - https://www.xsede.org/for-users/training
  - https://cvw.cac.cornell.edu/expanse/

Problems? Contact help@xsede.org

Thank You!

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