Workshop Agenda

Date and Time:

August 5, 2025
9:00 am - 5:30 pm

Workshop Title:

Practical strategies for reproducible research and robust data management

Location:

Comer Seminar Room
Lamont-Doherty Earth Observatory

Session / Description	Start Time	Duration
Overview of important concepts for research data management What is reproducibility? How can thoughtful data management make you a better scientist? What is data curation and why is it important? Best practices for FAIR data publishing.	9:00 am	90 Minutes
Morning Break	10:30 am	30 Minutes
Understanding reproducibility and data best practices through a real CZ data workflow Walk through an example of data best practices in action using a CZ data repository. See how different improvements could make the data and code even more robust and learn about the challenges and tradeoffs with implementing these.	11:00 am	60 Minutes
Lunch Break	12:00 pm	120 Minutes
Breakout Sessions → Choose your path	2:00 pm	90 Minutes

Breakout → Data publishing in HydroShare and EarthChem for CZNet researchers:

For those who are associated with the Critical Zone Network (CZNet), this session will introduce the data publishing portals, HydroShare and EarthChem (including sample registration via SESAR), and CZNet specific data publishing best practices.

Breakout → Reproducible workflows in R

Intro to the R package {targets} for reproducibly managing and automating your data workflows.

Breakout → Cloud computing workflows with HydroShare

Learn how to execute code in CUAHSI's free-to-use cloud computing environment with JupyterHub and Jupyter Notebooks.

Afternoon Break	3:30 pm	30 Minutes
Breakout Sessions Continue → This session will offer more interactive opportunities and on-demand support following the breakouts from the previous session. Attendees can try using the tools and resources with their own data or workflows. There will be plenty of time for Q&A during this session.	4:00 pm	90 Minutes

Breakout → Data publishing in HydroShare and EarthChem for CZNet researchers:

Get individual questions answered as you make a plan to publish your CZ data.

Breakout → Reproducible workflows in R:

Try building your own `targets` workflow for data analysis, practice writing your own custom functions, ask questions about robust workflow approaches for your own research, etc.

Breakout → Cloud computing workflows with HydroShare:

Practice running your own code in the cloud using JupyterHub and get questions answered about cloud computing workflows that are launched from HydroShare.