Data Archival and Dissemination Tools to Support Your Research, Management, and Education

LIZA BRAZIL
CUAHSI PRODUCT MANAGER
Shout Out: Upcoming Cyberseminars

April 13: Liza Brazil, CUAHSI: *Data Archiving and Dissemination Tools to Support Water Science Research*

April 20: Mark Henderson, CUAHSI: *Cloud-hosting Water Science Data for Collaborative Research*

April 27: Anthony Castronova, CUAHSI: *Cyberinfrastructure to Support Water Science Education and Reproducible Science*
Outline

I. Brief Introduction to CUAHSI
II. Demonstration of CUAHSI resources and tools throughout the presentation
III. Data Management Lifecycle
IV. Data Management Plans
V. Data Management Best Practices
VI. How can CUAHSI help?
VII. Data Archival and Dissemination Using CUAHSI Tools
VIII. How to get involved
IX. Summary
What is CUAHSI?

• CUAHSI is a 501(c)3 Non-Profit Consortium of about 130 U.S. Academic Institutions, Non-Profits, and International Universities; also open to private organizations

  Mission is to shape the future of water science by:
  ◦ Strengthening interdisciplinary collaboration in the water-science community
  ◦ Empowering the community by providing critical infrastructure
  ◦ Promoting education in the water sciences at all levels

• Key Activities
  • Community Services, such as workshops, community meetings, training, etc.
  • Data and Model Services, including HydroShare and time-series services
Community Services

CUAHSI provides access to resources and specialized training to build capacity and extend existing research capabilities.

- Learn a new instrumentation method
- Engage your local community on water issues
- Develop or improve data tools
- Add a new field site to existing research
- Attend the Summer Institute to contribute to the National Water Center goals
- Collaborate with researchers at other institutions

WWW.CUAHSI.ORG
Community Meetings
Save the Date!

2018 CUAHSI Biennial Colloquium

Hydrologic Connections: Climate, Food, Energy, Environment, and Society

July 29 – August 2 at NCTC

Shepherdstown, WV

- Workshops on drones and NASA SWOT mission
- Sessions on water and climate, hydrologic feedbacks, energy-water systems...

www.CUAHSI.org
Data Services

CUAHSI provides access to data and models to facilitate research and education.

- Share, discover, analyze, model, and publish all types of water data in a collaborative environment.
- Free, open-source software
- Fulfill data management plans
- Use for research or in the classroom
- Develop new software to meet community needs
Upcoming CUAHSI Events and Deadlines

◦ April 16, 2018: Instrumentation Discovery Travel Grant (IDTG) Program

◦ May 1, 2018: Submit an abstract to the poster session at CUAHSI's 2018 Biennial Colloquium. Abstracts may be submitted under any of the meeting themes.

◦ June 15, 2018: Register for CUAHSI's 2018 Biennial Colloquium (July 29 - August 1, 2018 at the National Conservation Training Center in Shepherdstown, WV). The theme of this year’s meeting is Hydrologic Connections: Climate, Food, Energy, Environment, and Society.

◦ July 15, 2018: Near Surface Geophysics for Hydrology Workshop (September 20 - 22, 2018 at the University of Wyoming in Laramie, WY)

◦ August 1, 2018: Snow Measurement Field School (January 7 - 10, 2019 in Bozeman, MT)
Questions?
The Data Management Lifecycle

Brainstorming Exercise

Think about or draw a flowchart to describe what you do with your data from the very beginning of your project to the very end.
The Data Management Lifecycle

Brainstorming Exercise

Think about or draw a flowchart to describe what you do with your data from the very beginning of your project to the very end.
Data Management Plans

Usually contain:

- Types of data
- Standards to be applied, for example format and metadata content
- Provisions for archival and preservation
- Access policies
- Plans for eventual transition or termination of the data collection in the long-term future

Many Funders require this, including NSF!

Information taken from DataOne
Data Management Best Practices

- Don’t mess with the raw data
- Use descriptive file names
- Use descriptive headers in files and tables that include the metadata and the data together
- Do not mix data types in table columns

- Use structured data entry forms or templates, validation checks, etc.
- Archive data in non-proprietary data formats
- Preserve data...to the cloud!
- Automate your analyses with reproducible code
- Maintain metadata in an accepted standard
How can CUAHSI help?

◦ Boilerplate text for data management plans if you use our data services
◦ Staff support to review your DMPs
◦ Resources for organizing your metadata into an accepted standard
◦ Tools for archiving and disseminating your data to fulfill your DMP
Questions?
Data Archival and Dissemination Using CUAHSI Resources and Tools
Create a Publishing Account

- [http://hydroserver.cuahsig.org](http://hydroserver.cuahsig.org)
  Create a publishing account using a Google account

- Create a group Google Account for shared access to the database to exist in perpetuity beyond your time at an organization or university
Format Your Data and Metadata

- CUAHSI requires that you format your data and metadata to fit with the Observations Data Model (ODM)
- Widely accepted metadata structure with controlled vocabulary
- Include standard metadata structure in your DMP
- Standard Format Template
Upload Your Metadata and Data Values to Your Database

- Built in checks for formatting errors, updating, and duplicate checking
- Give multiple individuals access to one database to spread the responsibility
- Keep the database private until paper publication, if applicable
Request Publication and Discover on HydroClient

- Public data are updated weekly
- Data are available to the broader community through CUAHSI Water Data Services
- CUAHSI maintains the database on their servers in perpetuity
- CUAHSI provides long term data archiving and data publication through HydroClient.
Proof of Data Publication

CUAHSI provides DOI for your data
DOI directs to CUAHSI Catalog where you can include:
  - Funding information
  - Summary of published data
Point to HydroClient for others to discover, view, and/or download your data alongside data from federal agencies, community groups, university researchers all in the same format!

http://hiscentral.cuahsi.org/pub_network.aspx?n=228
Potential Next Steps

◦ Create publishing account: http://hydroserver.cuahsi.org

◦ Download formatting templates: www.cuahsi.org/data-models/publication

◦ Check out HydroClient to discover data: http://data.cuahsi.org
Questions?
Summary

I. Data management plans should come before data collection
II. Best practices for data management will increase reproducibility
III. CUAHSI provides resources and tools for data archival and dissemination that help with
IV. Attend the upcoming Cyberseminars
Thank you!

Contact Information
Liza Brazil
Product Manager
LBrazil@cuahsi.org
339-933-4130