OAC-1664061
OAC-1664018
OAC-1664119
2017-2021

ACI-1148453
ACI-1148090
2012-2017

What it takes to achieve water data sharing: lessons learned from CUAHSI HydroShare

Access these slides in HydroShare by searching for "niehs2020"

David Tarboton, Jeffery S Horsburgh, Daniel P Ames, Jonathan L Goodall, Alva Couch, Pabitra Dash, Hong Yi, Christina Bandaragoda, Anthony Castronova, Bart Nijssen, Richard Hooper, Shaowen Wang, Mohamed Morsy, Scott Black, Chris Calloway, Jerad Bales, Martin Seul

HydroShare is operated by CUAHSI with ongoing development through a collaborative project among Utah State University, Brigham Young University, CyberGIS Center University of Illinois, Tufts, University of Virginia, and RENCI University of North Carolina.

HYDROSHARE

http://www.hydroshare.org

CUAHSI
universities allied for water research

Grand Challenges in Hydrology and Water Resources

Advancing understanding and generating knowledge in water research depends on collaboration and data sharing

- Open data
- Integration of information from multiple sources
- Easy to use generally accessible shareable computing
- Working as a team and community

HYDROSHARE

Grand Challenges in Hydrology and Water Resources

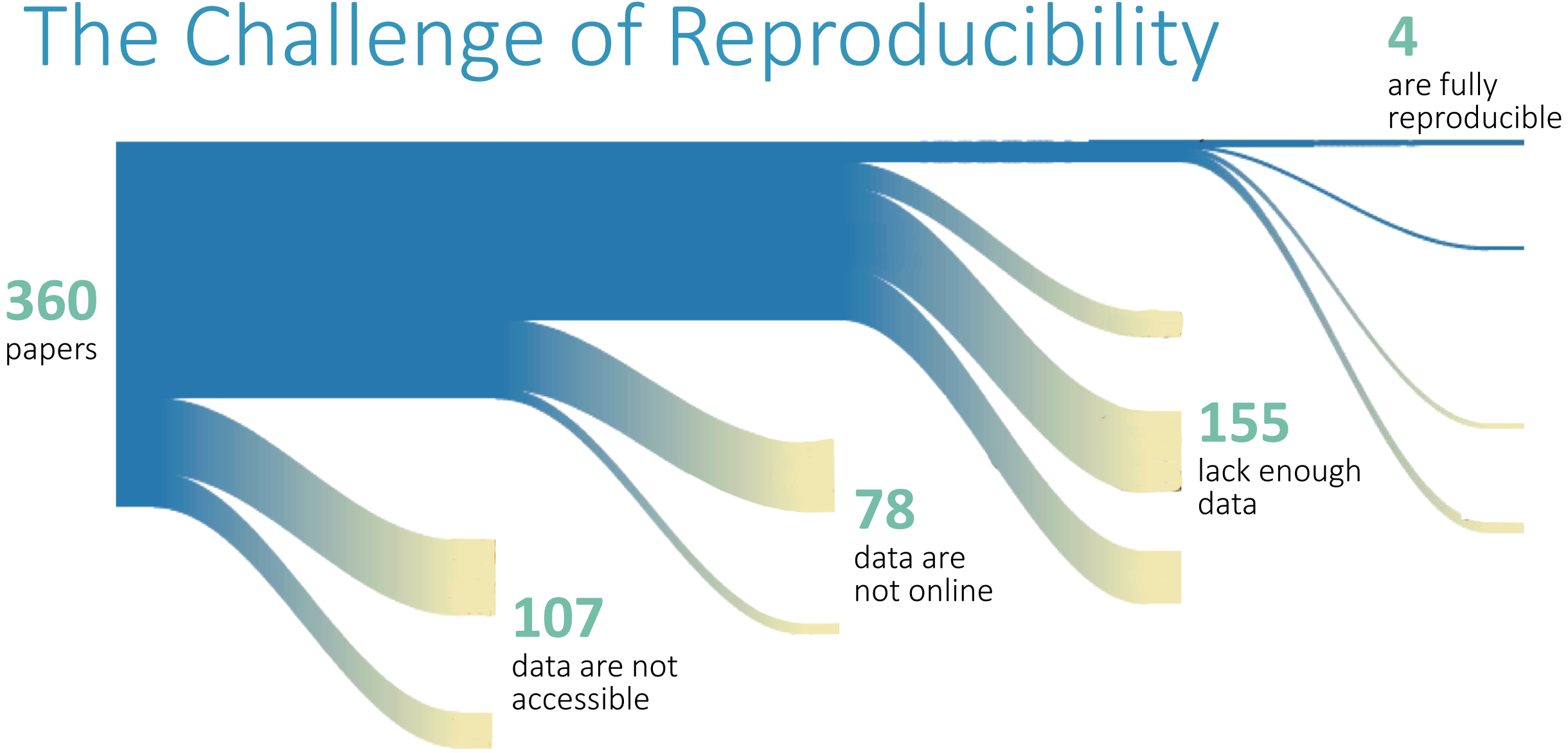
Advancing understanding and generating knowledge in water research depends on collaboration and data sharing

Water Research is a team sport

- C n data
  mation from multiple
  accessible shareable
  computing
- Working as a team and community

HYDROSHARE

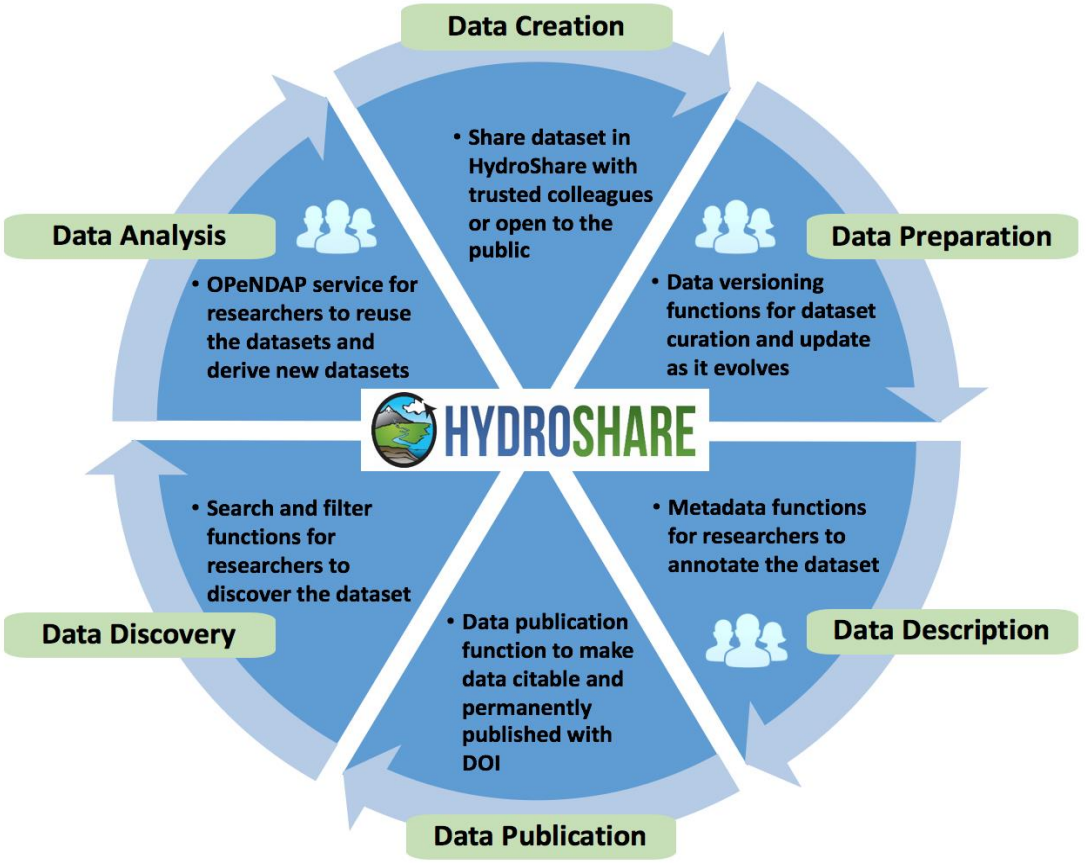
The Challenge of Reproducibility



Stagge, J. H., D. E. Rosenberg, A. M. Abdallah, H. Akbar, N. A. Attallah and R. James, (2019), "Assessing data availability and research reproducibility in hydrology and water resources," Scientific Data, 6: 190030, https://doi.org/10.1038/sdata.2019.30.

The take home message

- Fulfilling the open data mandate and achieving FAIR data and reproducible trustworthy scientific research requires establishing and advancing information, data and model sharing systems that are easy to use and simplify and facilitate data management through the full data lifecycle
- CUAHSI HydroShare is our effort to achieve this for the Hydrology and Water Resources Research community
- Consider using HydroShare if your data are Water related
- Consider ideas from and approaches developed by HydroShare in other information and data sharing systems and repositories



Gan et al. 2020, https://doi.org/10.1016/j.envsoft.2020.104706

HYDROSHARE

HydroShare is a platform for sharing Hydrologic Resources and Collaborating

- File Storage

DropBox-ish Functionality

- Meta Data Descriptions
- Data Access API
- Web Apps
- Social Functions
- DOI Data Publication

Value Added Functionality

dropbox.com

From Dan Ames

HydroShare is a platform for sharing Hydrologic Resources and Collaborating

- File Storage
- Meta Data Descriptions
- Data Access API
- Web Apps
- Social Functions
- DOI Data Publication

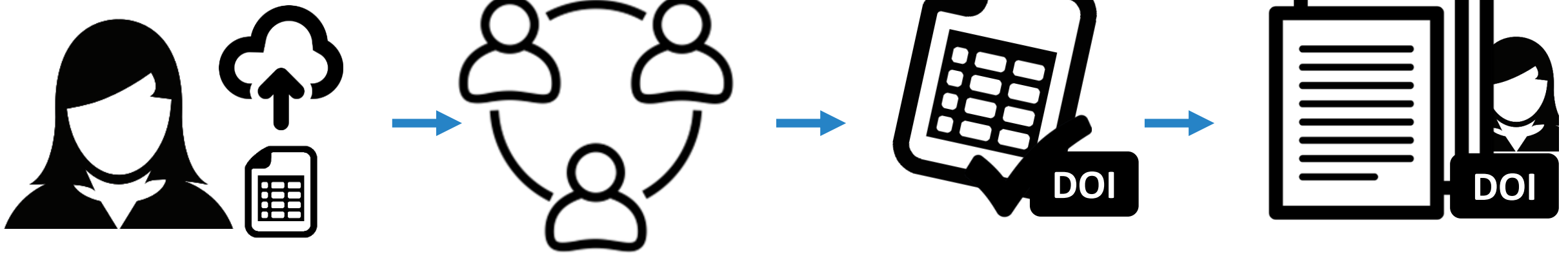
DropBox-ish Functionality

Value Added Functionality

The goal of HydroShare is to advance hydrologic science by enabling the scientific community to more easily and freely share products resulting from their research - not just the scientific publication summarizing a study, but also the data and models used to create the scientific publication.

From Dan Ames

HYDROSHARE enables transparency



User creates dataset and uploads to Hydroshare

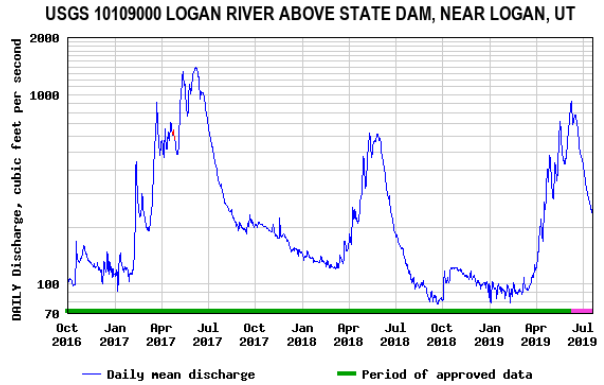
Research team iterates collaboratively

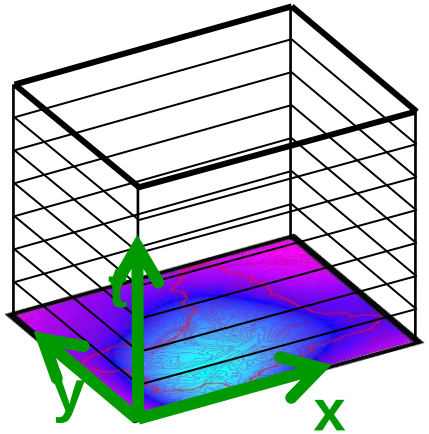
Final version is published and assigned DOI

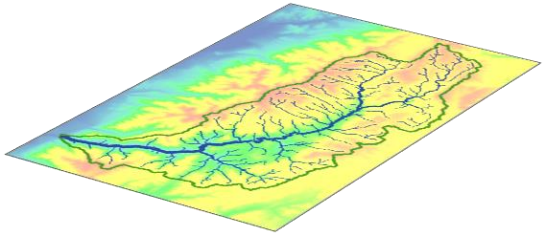
User cites dataset in paper

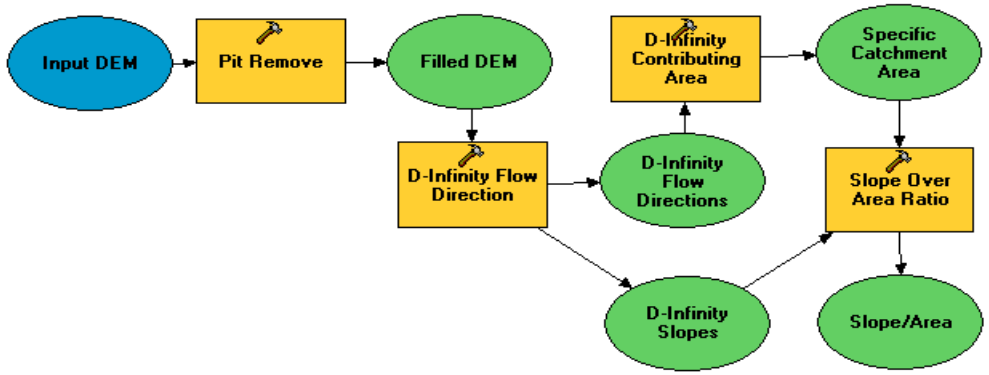
HydroShare founding Principles

- Data and models are first-class products of research and should be shared
- Data and models become **social objects**
- Data and models are stored in and become shareable “Resources” in HydroShare
- Research data management should start at the beginning of a project
- Resources should be created and systematically managed through the full research life cycle and include analysis and modeling workflows









HYDROSHARE

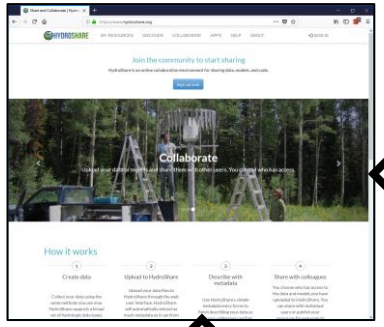
HydroShare Components

Resource exploration

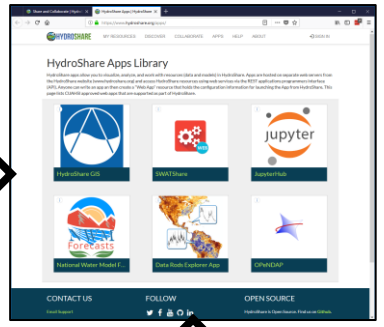
- Discover content
- Organize and annotate your Resources
- Group Collaboration
- Manage access

Django Website

HydroShare Apps



API



OAuth

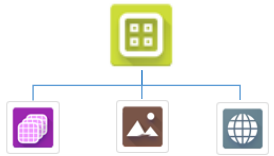
API

API

Resource Storage
iRODS “Network File System”

Distributed file storage

Resource data model concepts



Actions on Resources

- Web software to operate on content you have access to (Apps)
- Computation
- Extensibility through user applications

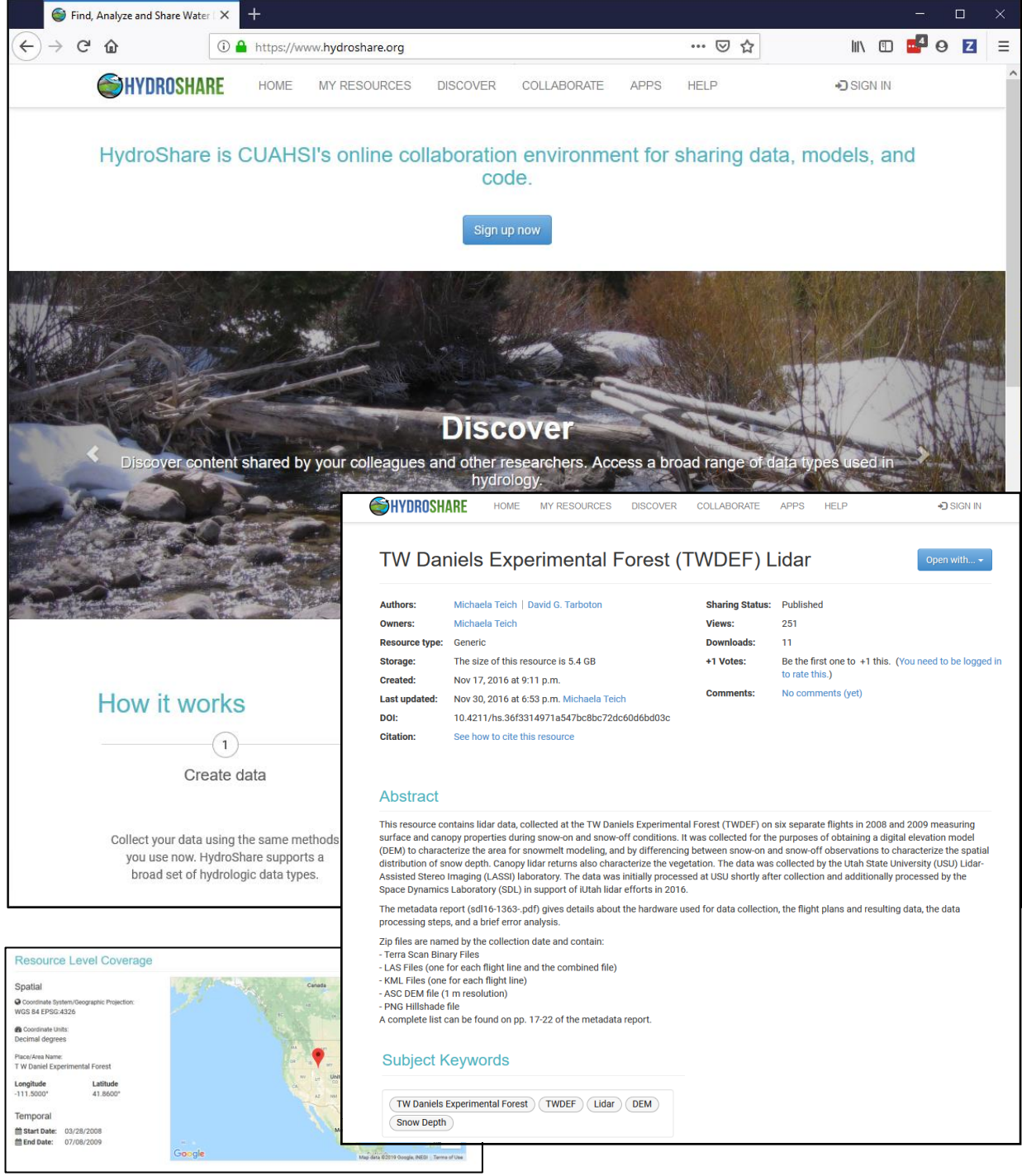
Anyone can set up a server/app platform (software service) to operate on HydroShare resources through iRODS and API

- JupyterHub and CyberGIS Jupyter for Water
- MATLAB Online
- Unidata – THREDDS
- SWATShare (Hubzero)
- …

Moving towards fully web based hydrologic innovation environment

Data and Model Repository
(Django Website)

- Manage data (and models and workflows) throughout research life cycle
- Share data, models, and other products
- Create and edit metadata, some of it automatically
- Permanent publication of data and models with citable digital object identifiers (DOIs)
- Fulfill Findable, Accessible, Interoperable, Reusable (FAIR) open data mandate



HYDROSHARE

Example

Water Resources Research

RESEARCH ARTICLE
10.1029/2019WR024837

Terrain Analysis Enhancements to the Height Above Nearest Drainage Flood Inundation Mapping Method

Irene Garousi-Nejad[1], David G. Tarboton[1], Mahyar Aboutalebi[1], and Alfonso F. Torres-Rua[1]

[1]Department of Civil and Environmental Engineering, Utah Water Research Laboratory, Utah State University, Logan, UT, USA

Key Points:

- Comparison of flood inundation mapped using Height Above Nearest Drainage (HAND) to inundation observed by Planet high-resolution imagery
- Improvements in HAND flood inundation mapping by conditioning the underlying digital elevation model using high-resolution hydrography
- Potential to use satellite observed inundation to infer distributed hydraulic roughness parameters for HAND-based hydraulic routing

Abstract Flood inundation remains challenging to map, model, and forecast because it requires detailed representations of hydrologic and hydraulic processes. Recently, Continental-Scale Flood Inundation Mapping (CFIM), an empirical approach with fewer data demands, has been suggested. This approach uses National Water Model forecast discharge with Height Above Nearest Drainage (HAND) calculated from a digital elevation model to approximate reach-averaged hydraulic properties, estimate a synthetic rating curve,

In an effort to make this study reproducible, the data and computational scripts used to produce the study results have been saved in HydroShare (Garousi-Nejad et al., 2019). The code for the flow direction conditioning tool is part of TauDEM and is available from the TauDEM GitHub repository (http://github.com/dtarb/taudem).

Garousi-Nejad, I., D. Tarboton, M. Aboutalebi, A. F. Torres-Rua (2019). Data for terrain analysis enhancements to the height above nearest drainage flood inundation mapping method, *HydroShare*, https://doi.org/10.4211/hs.7235a0d6a1834078b2028085b7d8018

Data For Terrain Analysis Enhancements to the Height Above Nearest Drainage Flood Inundation Mapping Method

Open with...

Authors: Irene Garousi-Nejad | David Tarboton | Mahyar Aboutalebi | Alfonso Faustino Torres-Rua

Owners: David Tarboton | Irene Garousi-Nejad

Resource type: Composite Resource

Storage: The size of this resource is 4.0 GB

Created: Aug 19, 2019 at 7:35 p.m.

Last updated: Sep 03, 2019 at 5:50 a.m. Irene Garousi-Nejad

DOI: 10.4211/hs.7235a0d6a18343078b2028085b7d8018

Citation: See how to cite this resource

Content types: Geographic Feature Content | Geographic Raster Content

Sharing Status: Published

Views: 846

Downloads: 618

+1 Votes: 1 other +1 this (You need to be logged in to rate this.)

Comments: No comments (yet)

Abstract

This resource contains the data and scripts used for: Garousi-Nejad, I., D. G. Tarboton, M. Aboutalebi and A. F. Torres-Rua, (2019), "Terrain Analysis Enhancements to the Height Above Nearest Drainage Flood Inundation Mapping Method," Water Resources Research, http://doi.org/10.1029/2019WR024837.

References

Related Resources

The content of this resource serves as the data for:	Garousi-Nejad, I., D. G. Tarboton, M. Aboutalebi and A. F. Torres-Rua, (2019), "Terrain Analysis Enhancements to the Height Above Nearest Drainage Flood Inundation Mapping Method," Water Resources Research, http://doi.org/10.1029/2019WR024837.

Credits

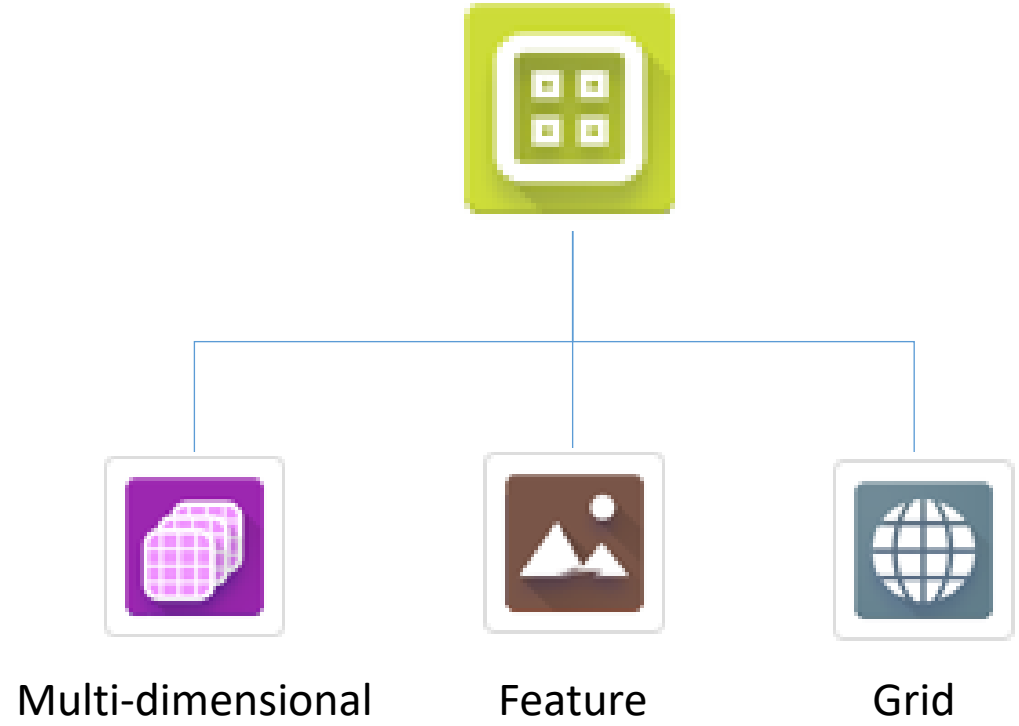
Funding Agencies

This resource was created using funding from the following sources:

Agency Name	Award Title	Award Number
National Science Foundation	Scalable Capabilities for Spatial Data Synthesis	1443080
Utah Water Research Laboratory	Graduate Student Research Assistantship for I Garousi-Nejad	

HydroShare OAI-ORE standard based Resource Data Model

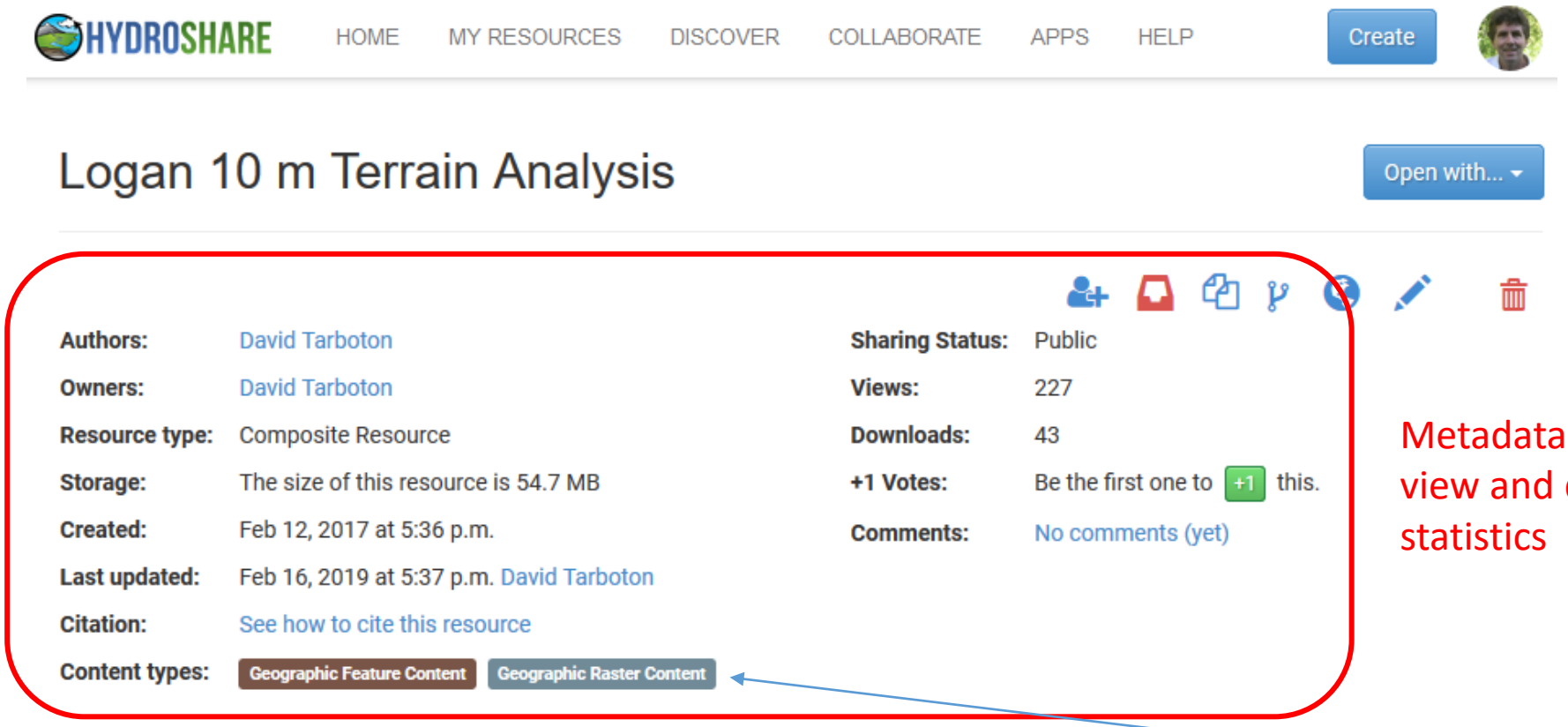
- A resource can hold multiple content aggregations
  - Managed as one discoverable resource
  - One unique identifier
  - One set of resource level metadata
  - One set of access controls (Owners, Editors etc.)
  - May be private, shared with others (users and groups), public, or permanently published to encourage early creation and addition of information through full data lifecycle
- A content aggregation
  - Can hold one or multiple files that comprise a single logical object
  - Each being a different type of data
  - One set of aggregation level metadata
  - Content aggregations are automatically created and metadata generated when a recognized file type is uploaded



Schema.org + Dublin Core machine readable metadata to make data in HydroShare FAIR

Horsburgh, J. S., M. M. Morsy, A. M. Castronova, J. L. Goodall, T. Gan, H. Yi, M. J. Stealey and D. G. Tarboton, (2016), "HydroShare: Sharing Diverse Environmental Data Types and Models as Social Objects with Application to the Hydrology Domain," JAWRA Journal of the American Water Resources Association, 52(4): 873-889, http://dx.doi.org/10.1111/1752-1688.12363.

Resource Landing Page



Metadata header with view and download statistics

Content types

Resources and content aggregations

Resources

Primary

- RESOURCE
- COLLECTION
- WEB APP CONNECTOR

Legacy

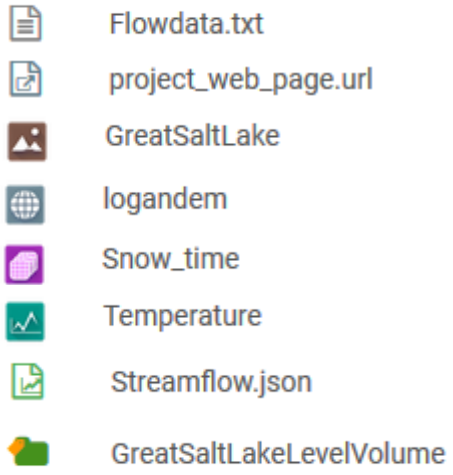
- MODEL PROGRAM
- MODEL INSTANCE
- SWAT MODEL INSTANCE
- MODFLOW MODEL INSTANCE

To be implemented as type aggregations within a composite resource

Composite resource contents

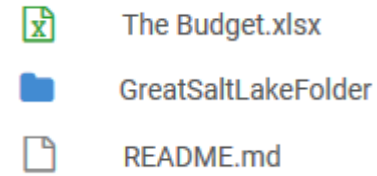
Content type aggregations

- Single File (with external reference as URL as a special case)
- Geographic Feature
- Geographic Raster
- Multidimensional
- Time Series (ODM2 format)
- Referenced Time Series (JSON))
- File set (Folder with files)



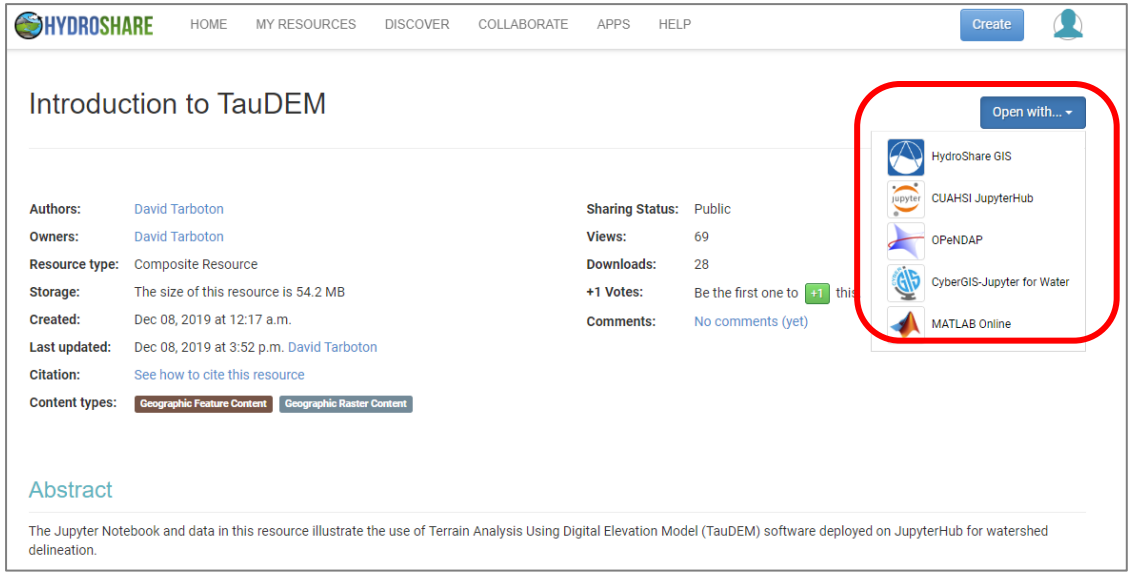
Files

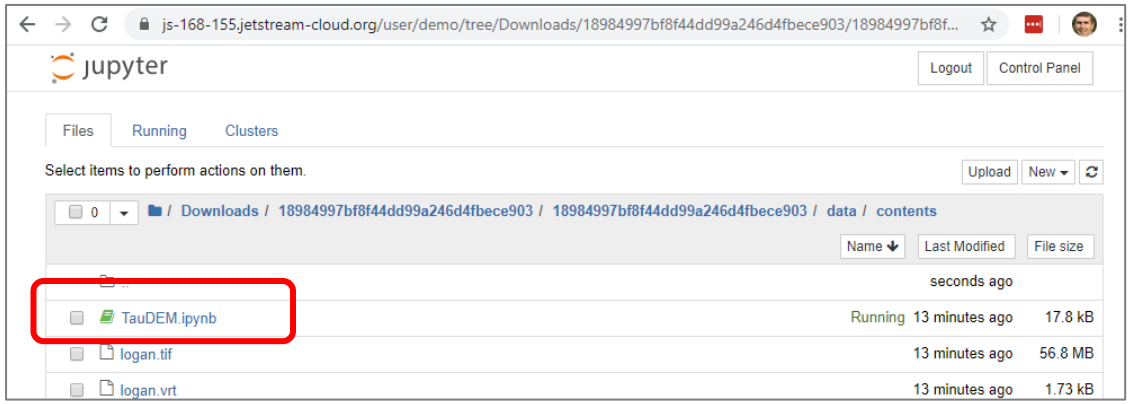
- Single File
- Folder
- Readme (markdown or txt) that renders on landing page

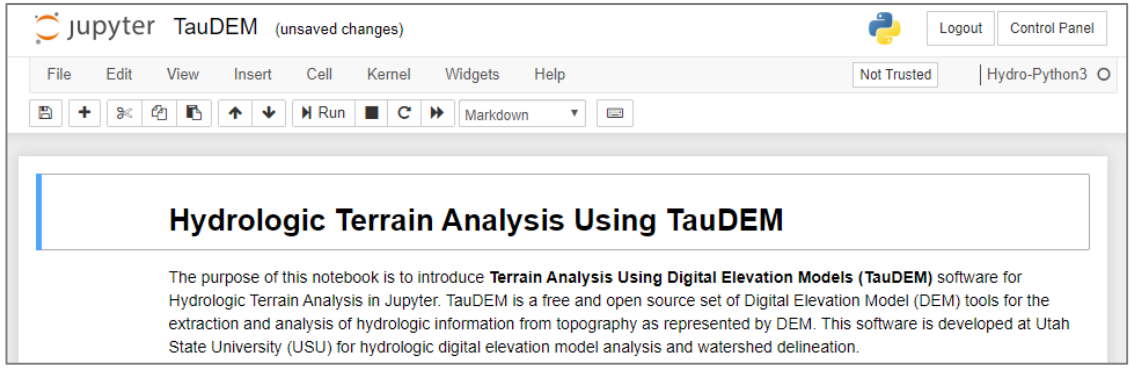


Web based Gateway to computing (Apps and JupyterHub)

- Provide immediate value
  - What can I do now that I may not be able to easily do on my PC
- Model input data preparation
- Model execution
- Visualization and analysis (best of practice tools)
- Reduced needs for software installation and configuration (platform independence)
- Teaching
- Write and execute code in a Jupyter Notebook, acting on content of HydroShare resources and saving results back to HydroShare Repository
  - Collaboration
  - Access to enhanced computation (HPC, Big data)
- Enhanced trust in research through transparency, replicability and reproducibility

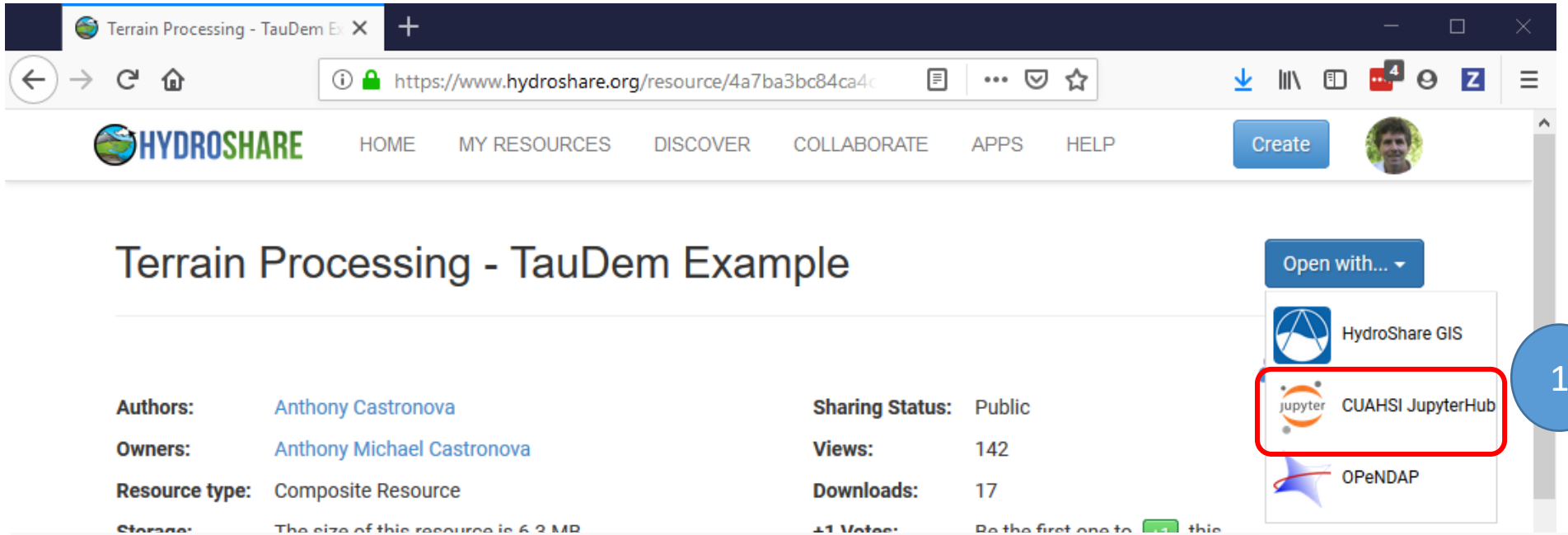






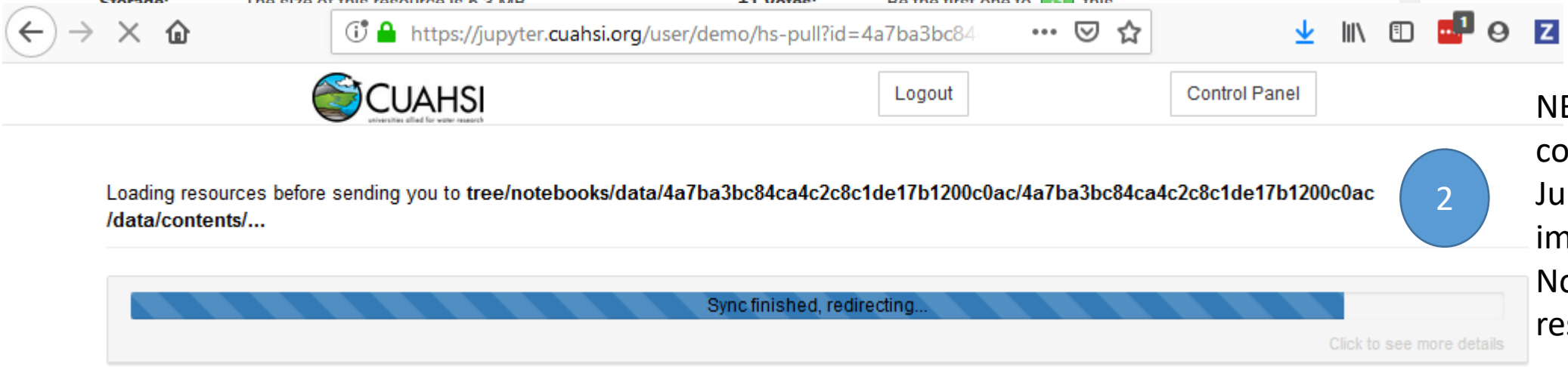
HYDROSHARE

JupyterHub



1

Web app connector launches external App



2

NBFetch used to retrieve contents of resource into JupyterHub to be able to immediately open Notebook and act on resource content files

Implemented through collaboration with Martin Hunt, Science Gateways Community Institute Consultant

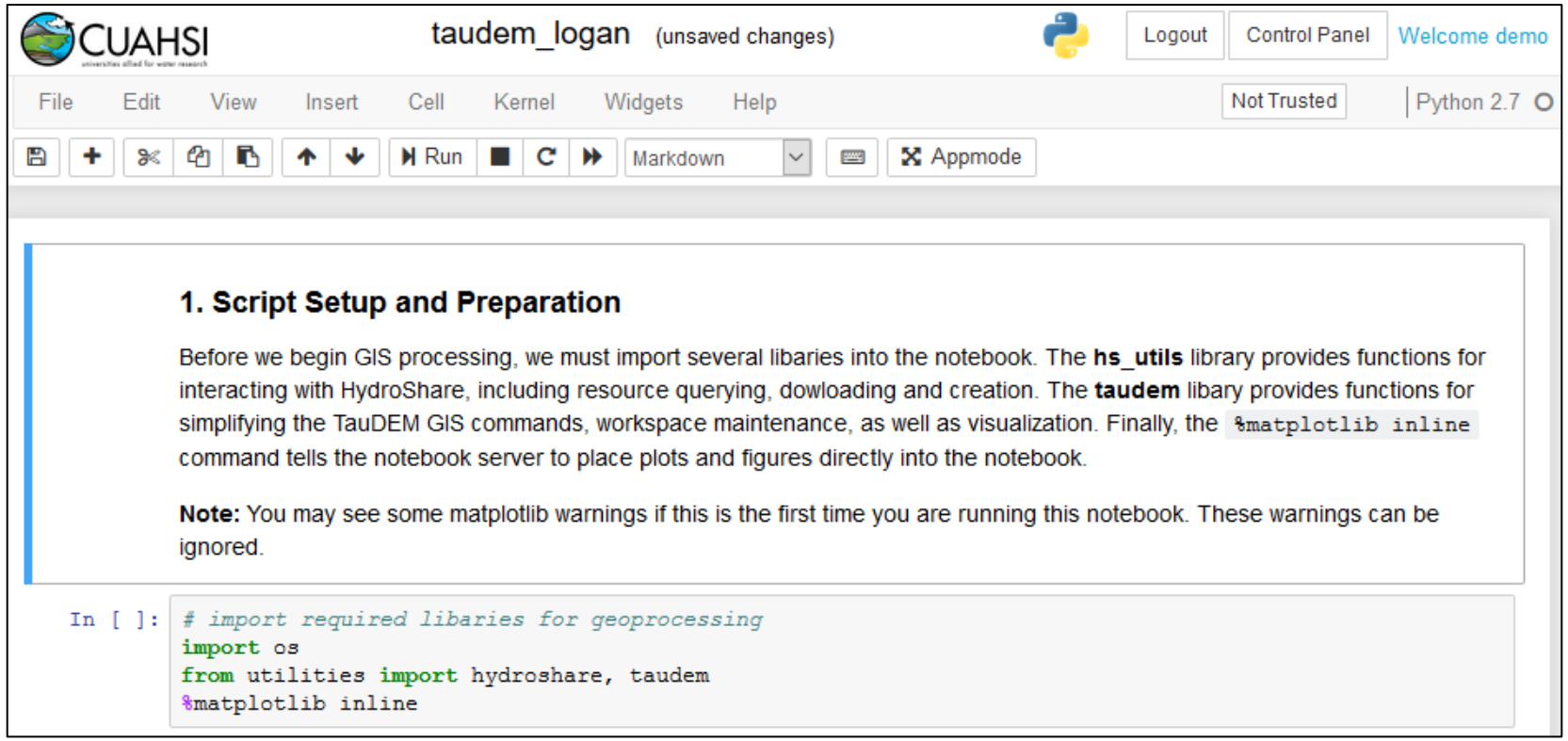
JupyterHub

Open Notebook

3

4

Analysis in JupyterHub

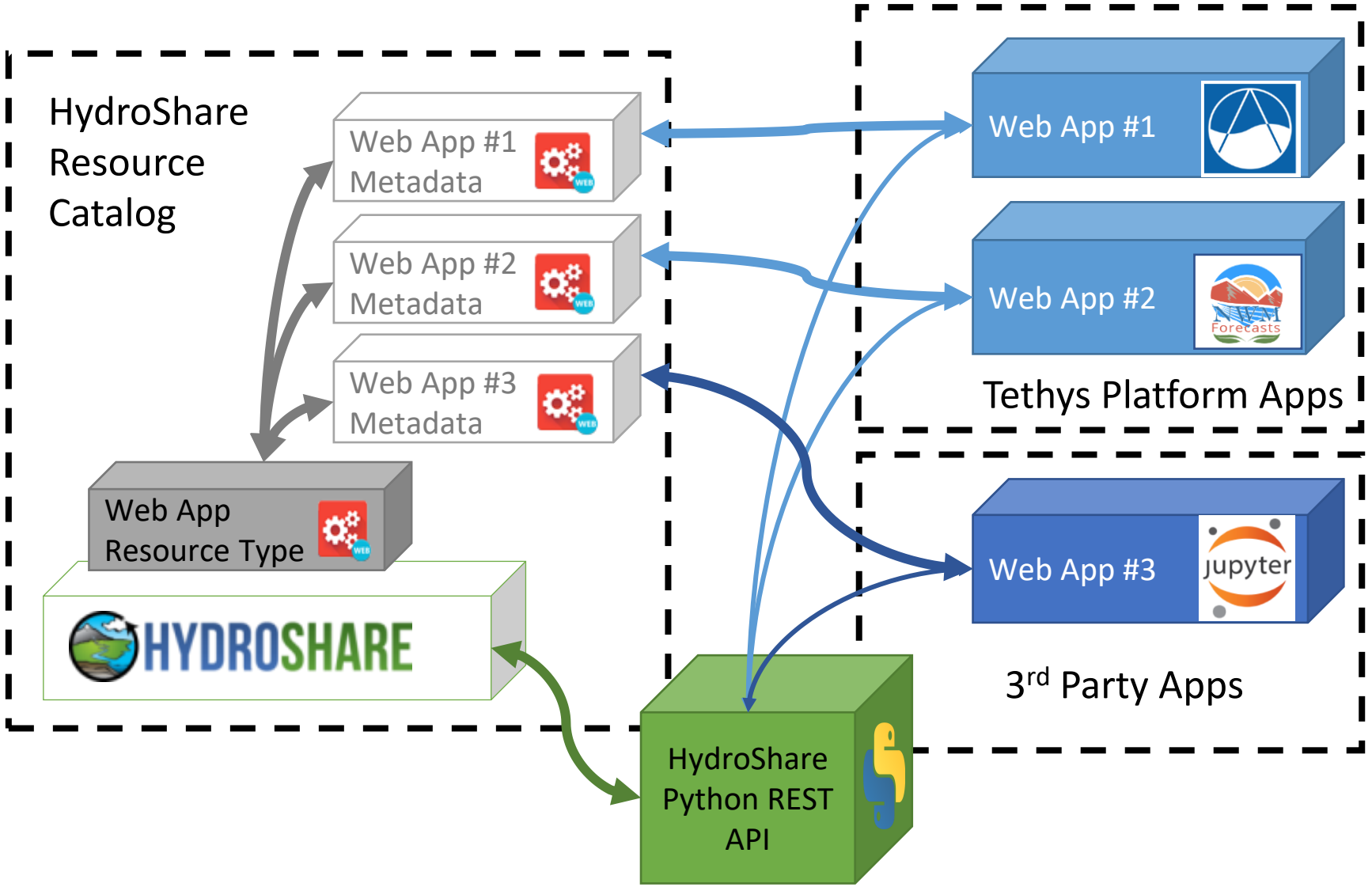


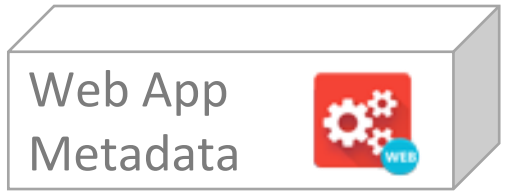
5

Save Results back to HydroShare

Web App Connector

Anybody can create a web app on any web server and configure a web app Connector for it to be launched from HydroShare





Predefined URL Launch Parameters:

Resource ID: ${HS_RES_ID}

Resource Type: ${HS_RES_TYPE}

HydroShare username: ${HS_USR_NAME}

Examples:

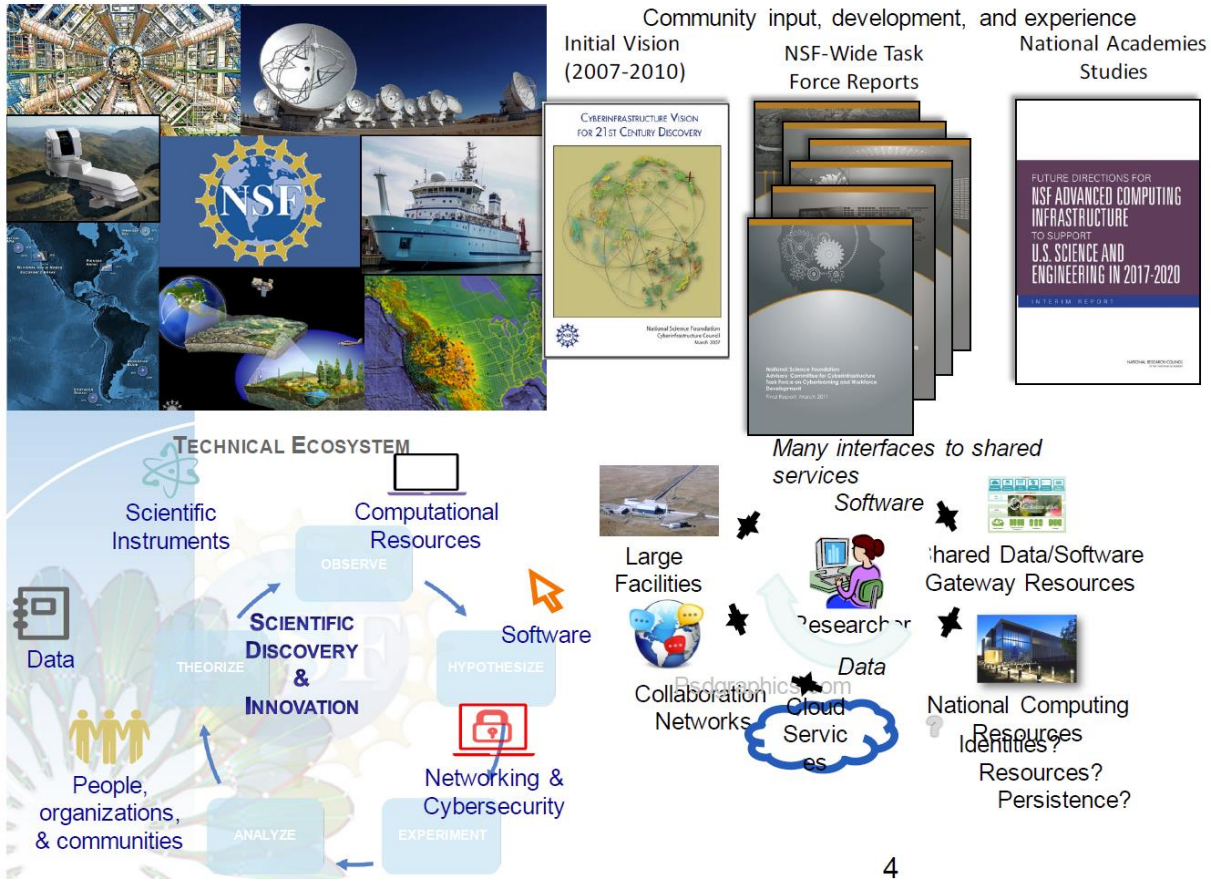
https://apps.hydroshare.org/apps/hydroshare-gis/?res_id=${HS_RES_ID}

https://mygeohub.org/.../?res_id=${HS_RES_ID}&usr=${HS_USR_NAME}&src=hs

http://hyrax.hydroshare.org/opendap/${HS_RES_ID}/data/contents/

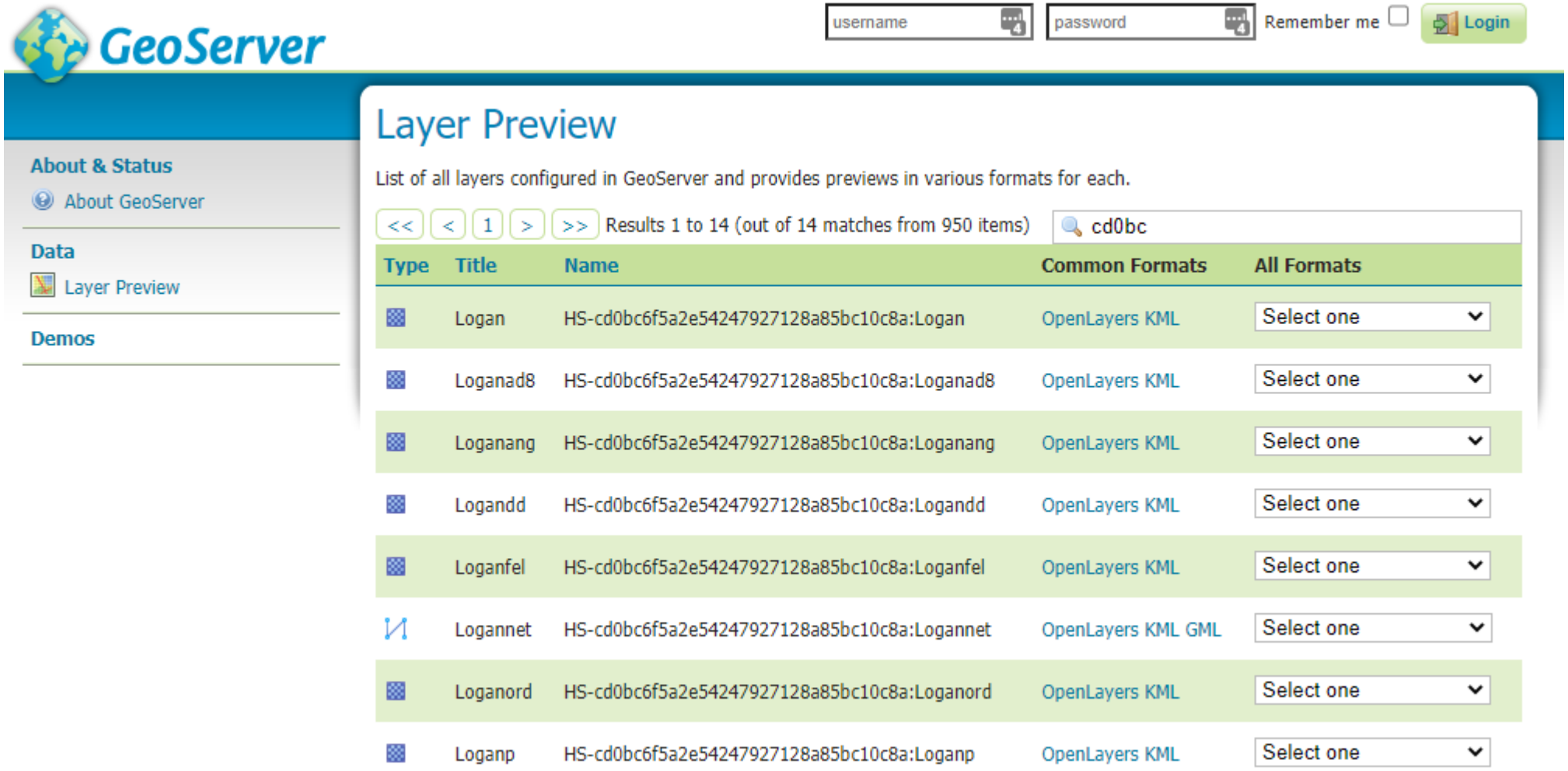
Interoperability and Software Ecosystems

- A foundation of the web
- No one system can do it all
- Applications programming interfaces (APIs)
- Unique Identifiers that enable linked data (web URI's)
- A cyberinfrastructure ecosystem of many interfaces to shared services
- Personal Cyberinfrastructure
  - Individually managed set of CI tools you assemble and learn to use to do your work



NSF vision for a cyberinfrastructure of many interfaces to shared services [Rajiv Ramnath, NSF Division of Advanced Cyberinfrastructure https://doi.org/10.6084/m9.figshare.4676173]

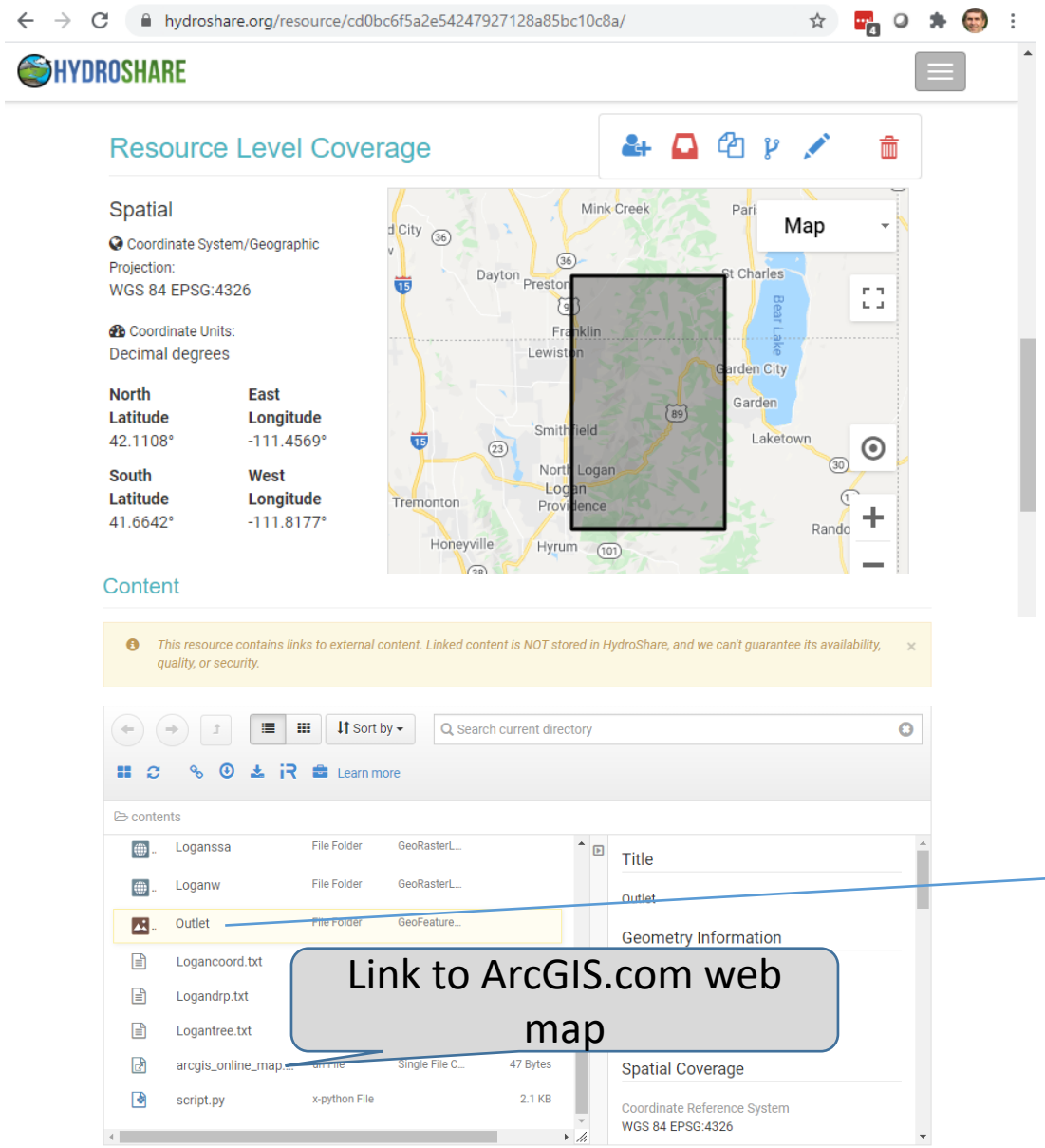
geoserver.hydroshare.org



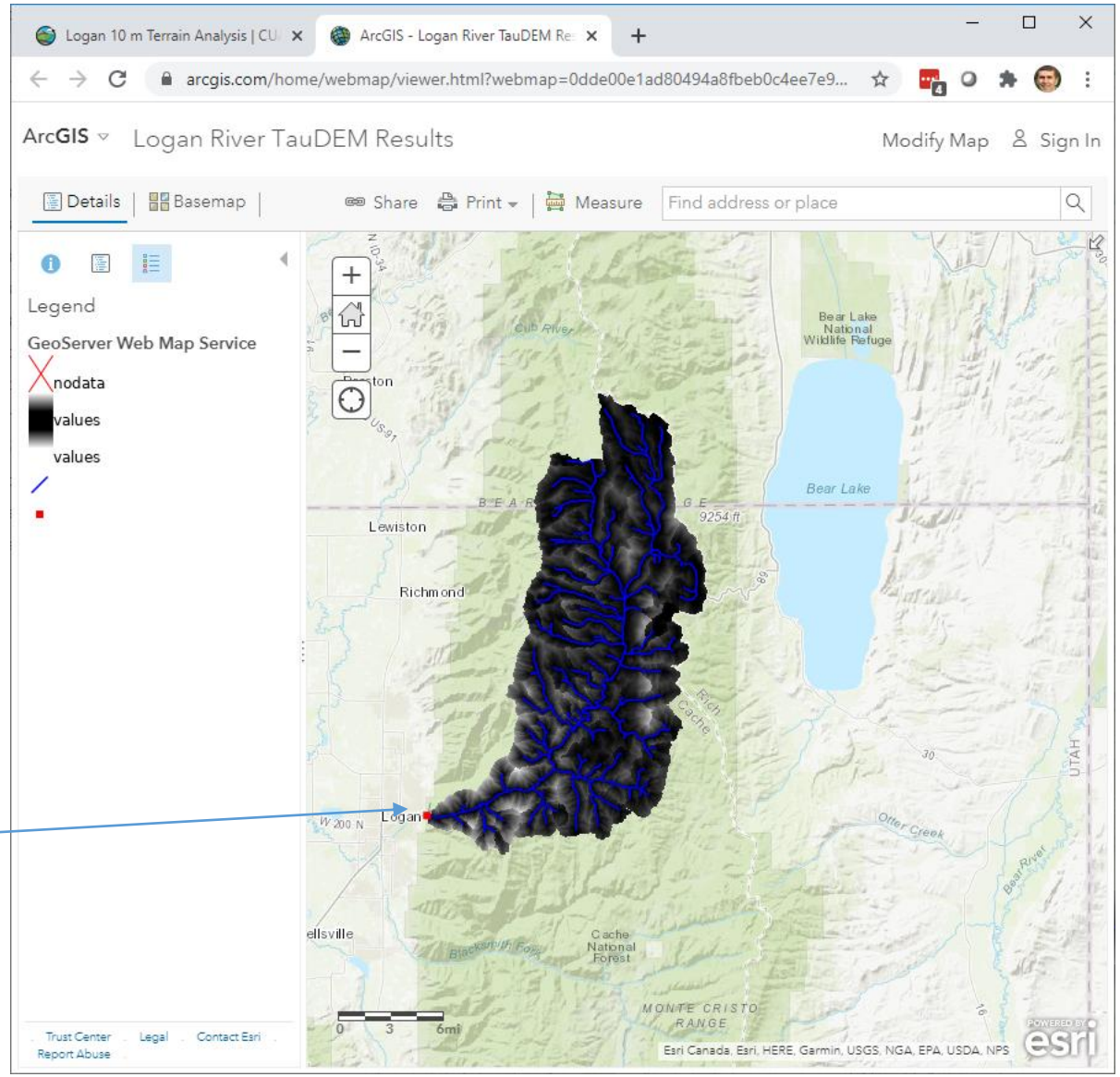
Public Resources with Geographic Content are published on Geoserver using OGC Web Services

https://geoserver.hydroshare.org/geoserver/HS-cd0bc6f5a2e54247927128a85bc10c8a/wms?request=GetCapabilities&service=WMS

HydroShare GeoServer contents in Arcgis.com



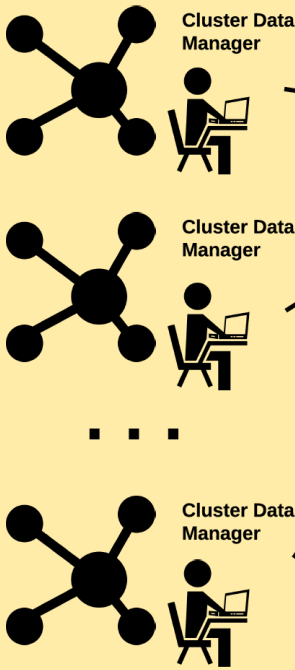
Link to ArcGIS.com web map



NSF Critical Zone (CZ) Collaborative Network Data Management

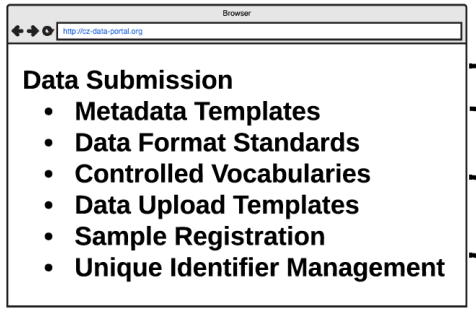
CZ Thematic Cluster Network

- Data collection
- Data aggregation
- Local data management
- Quality assurance/quality control
- Metadata creation



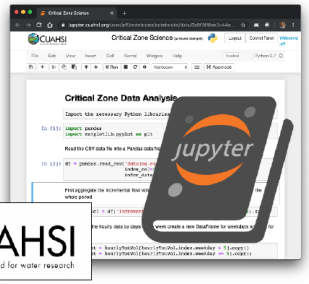
Data Submission Portal

- User support specialist
- Data curation support
- Documentation



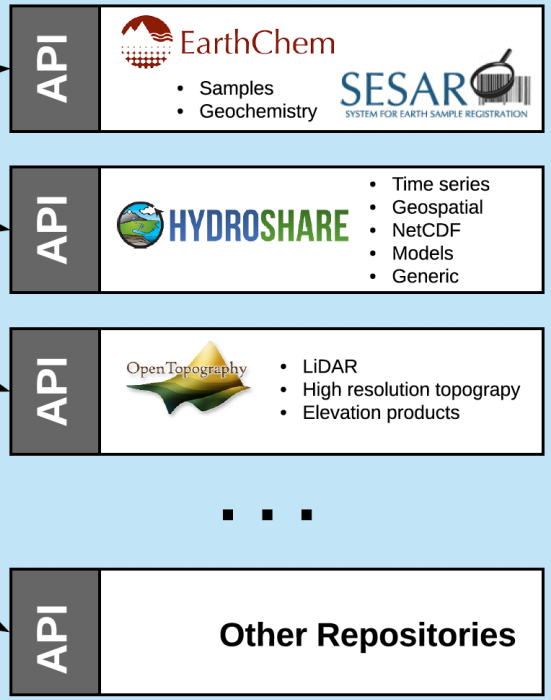
Computational Resources

- Jupyter Notebook support
- Computation and modeling
- Synthesis studies
- Reproducible analyses



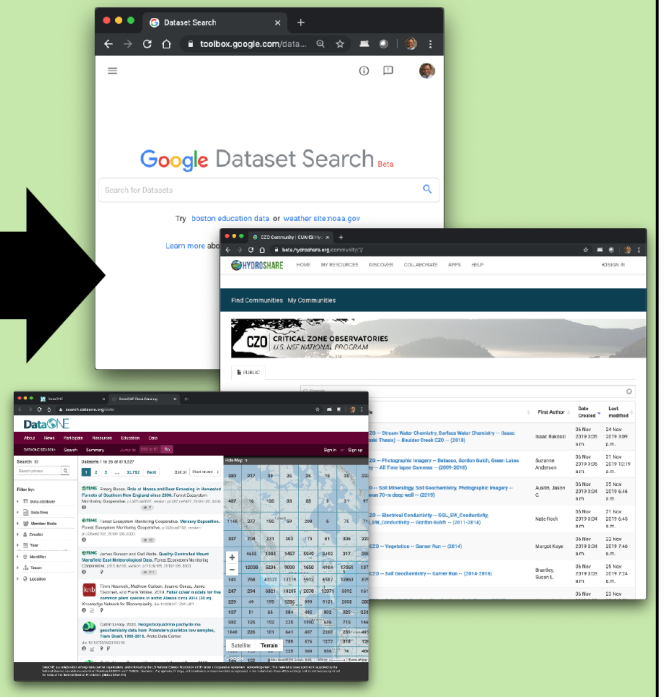
Repositories for Data and Research Products

- Permanent data archival and publication
- Access control for embargoed data
- Open access for public datasets
- Citable data



Catalog Services

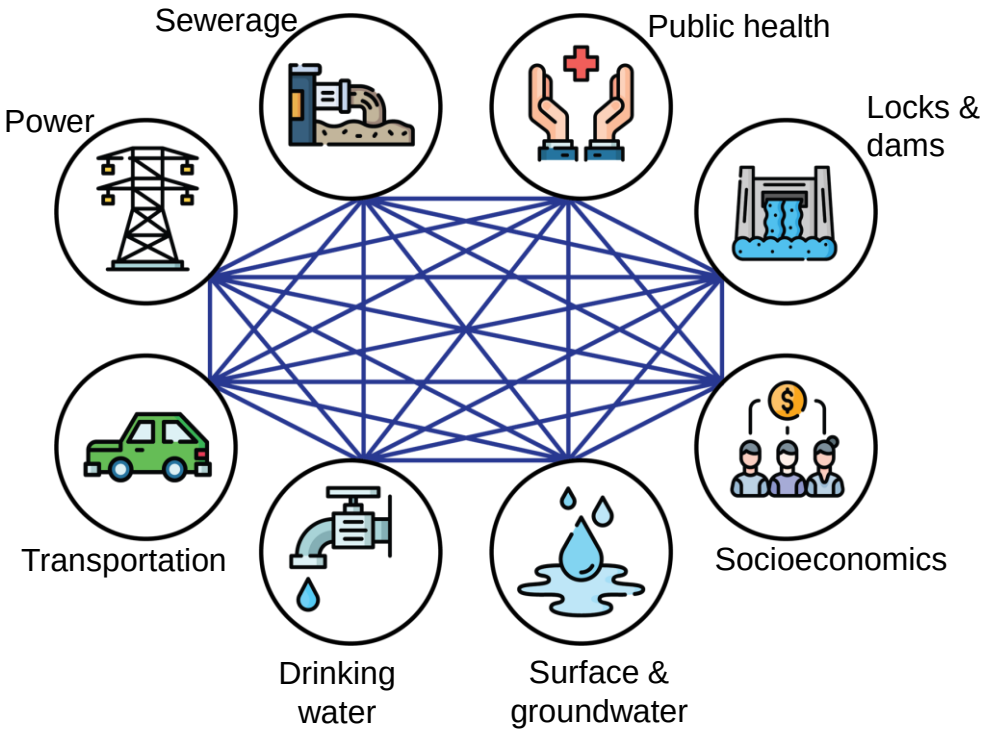
- Cross-repository view of CZ data and research products
- Discovery based on authors, keywords, geograhic area, time, Cluster
- Schema.org metadata implementation



UF-OKN Urban Flooding Open Knowledge Network

- Delivering flood information to anyone, anytime, anywhere
- Holistic approach for better understanding how floods impact the urban multiplex (water, transportation, power, civic infrastructure networks etc.)
- Knowledge graph to deliver results based on natural language queries

What areas of my city are at risk?

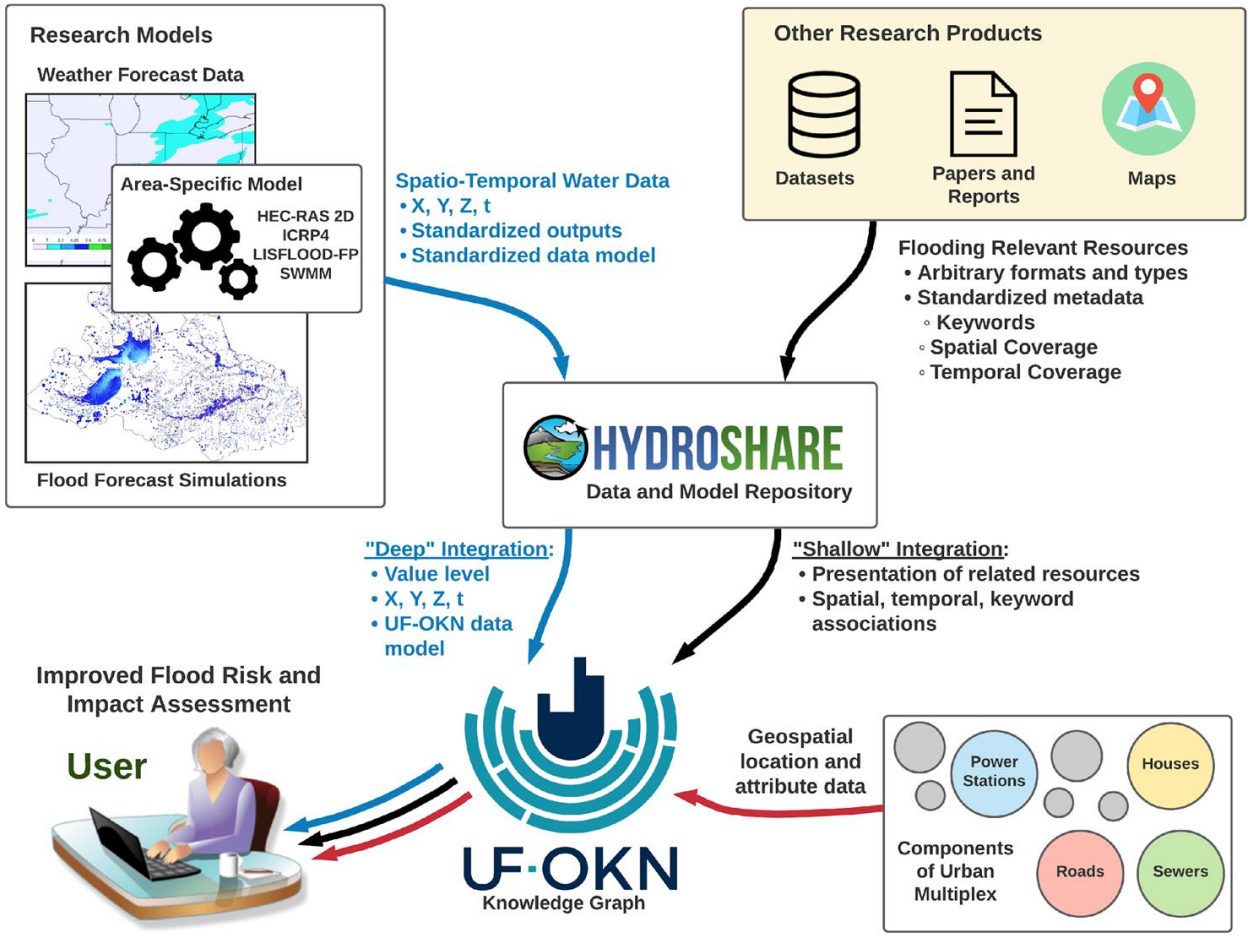


- NSF Convergence Accelerator project

https://ufokn.com/

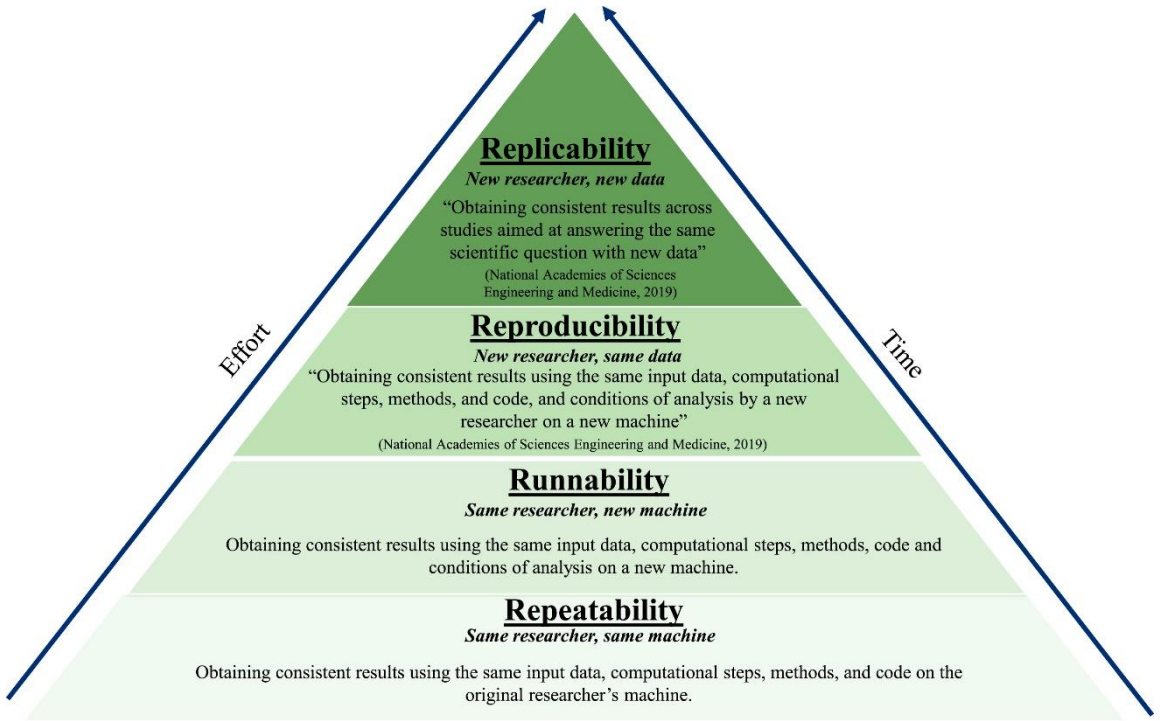
Contact
lilit.yeghiazarian@uc.edu

HydroShare supporting UF-OKN Research Community and Data Integration



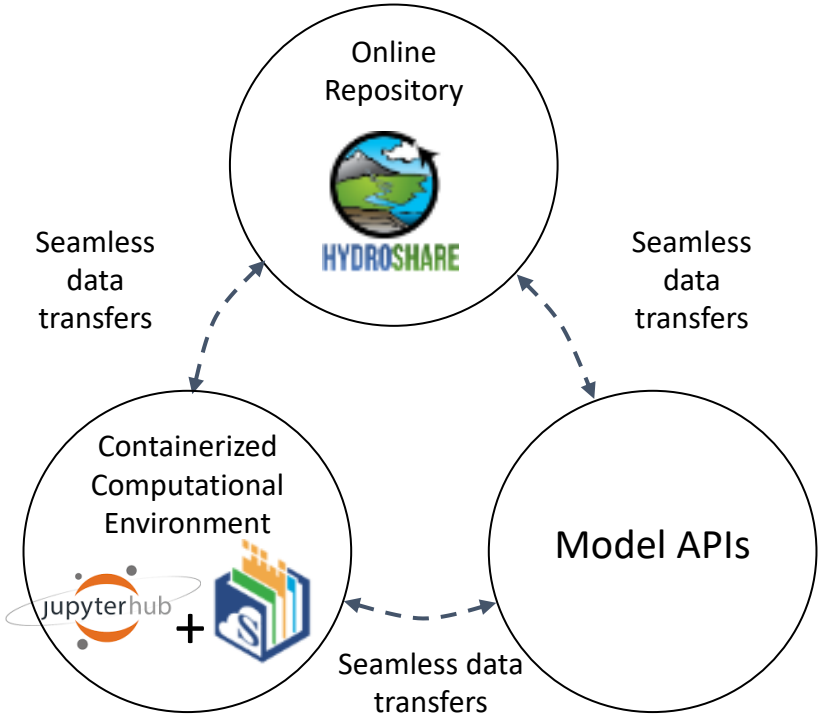
Reproducibility

What it means to reproducible



Essawy, et al., 2020, A taxonomy for reproducible and replicable research in environmental modeling, *Environmental Modelling & Software*, 134:104753, https://doi.org/10.1016/j.envsoft.2020.104753

EarthCube Reproducibility Workbench

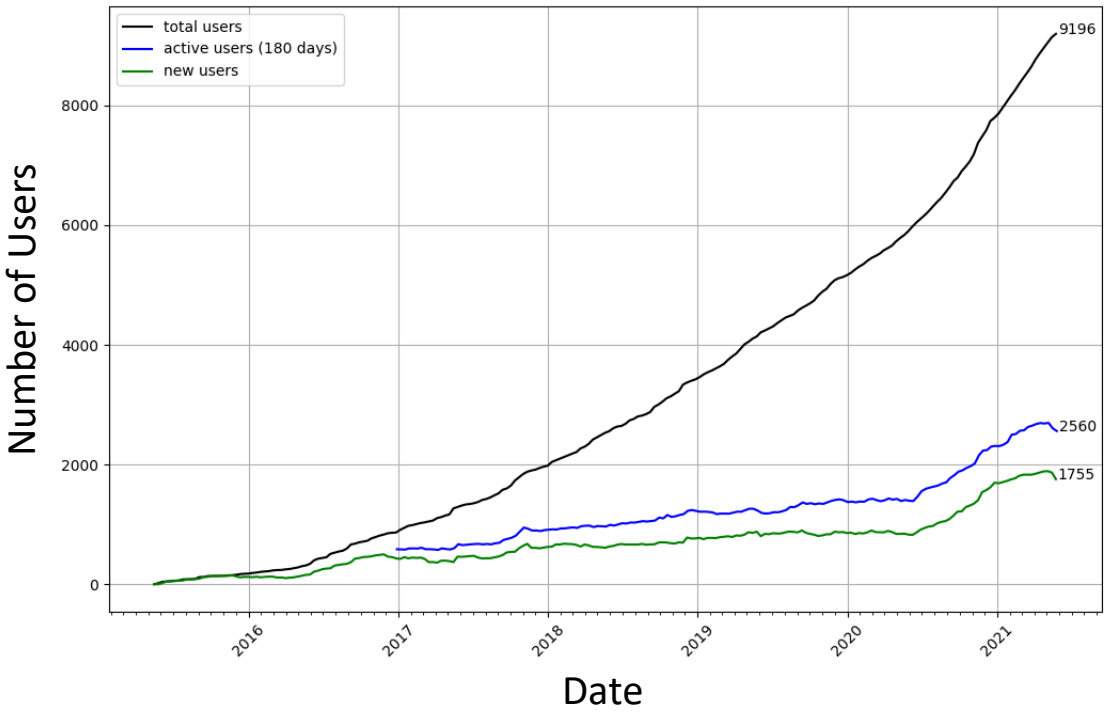


Choi, et al., 2021, Toward open and reproducible environmental modeling by integrating online data repositories, computational environments, and model application programming interfaces, Environmental Modelling & Software, 135:104888, https://doi.org/10.1016/j.envsoft.2020.104888

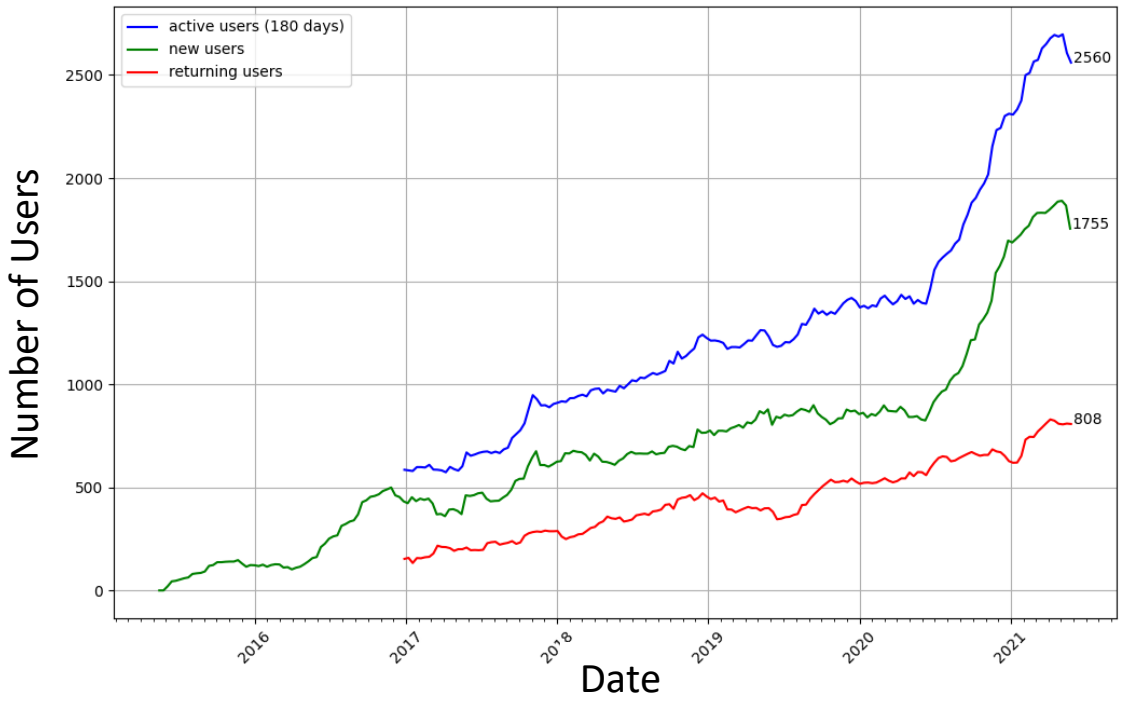
Audience and Use

Primary audience is US Hydrologic Research community (NSF funding) but open to international use and use by water resource professionals, educators and citizen scientists

Overall User Accounts



Active User Accounts



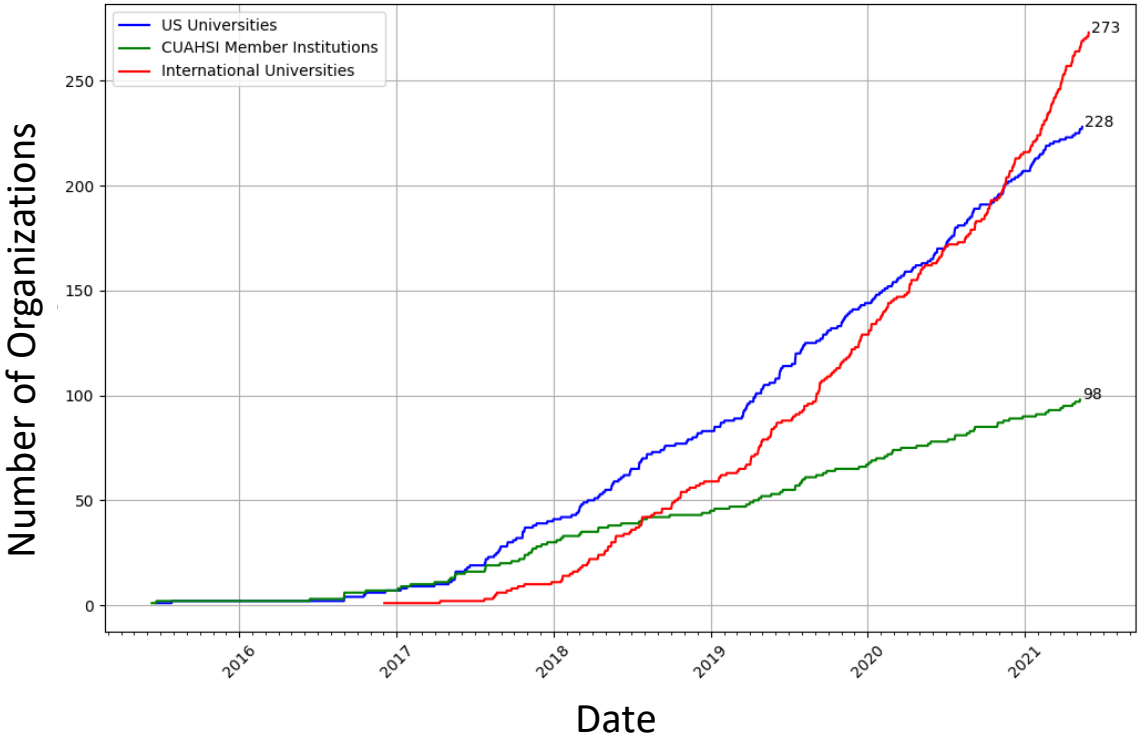
Returning users are users who returned to use (logged in to) HydroShare in the last 180 days after having created their account prior to the last 180 days

Data from: http://public.cuahsi.org/metrics-report/

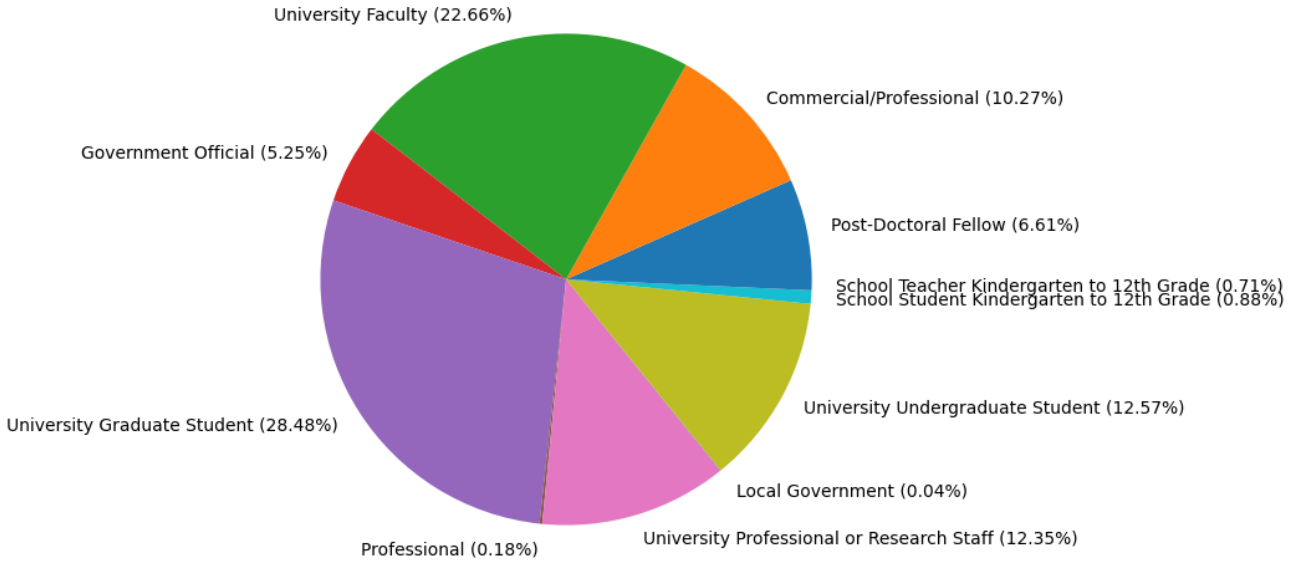
Statistics as of 5/30/2021

HYDROSHARE

Organizations and Users by Types



Distinct organizations that users are affiliated with.

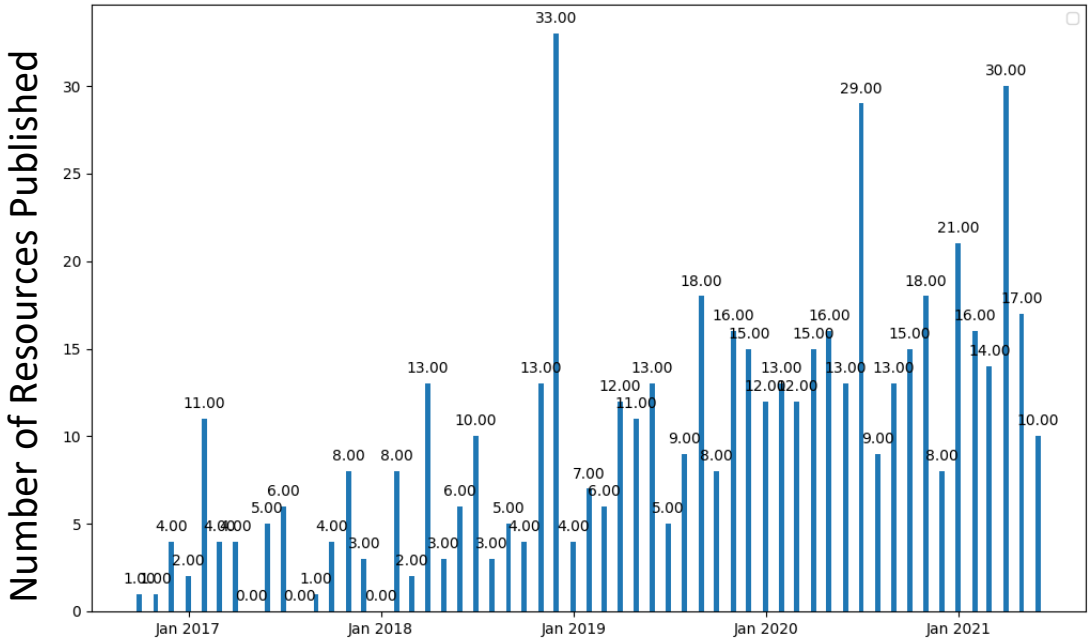


Users by type (from 54 % of users reporting type)

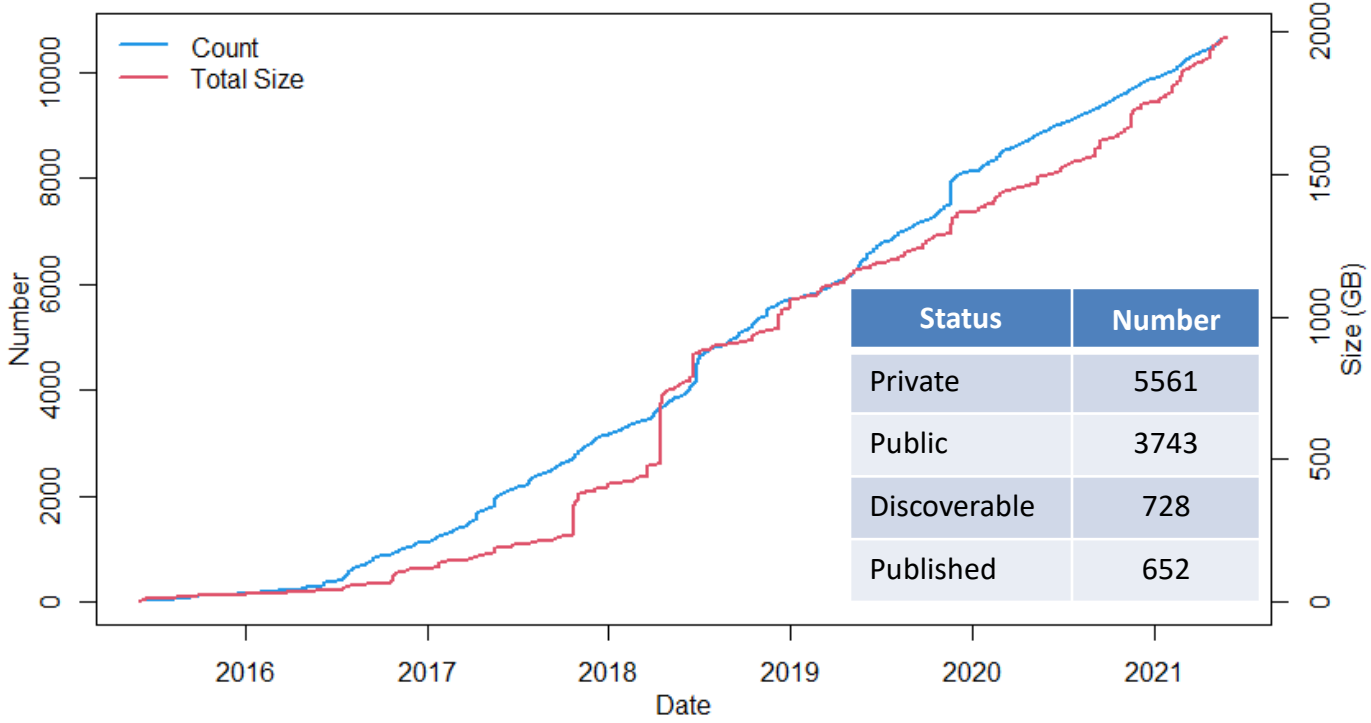
Statistics as of 5/30/2021

HYDROSHARE

Resources



Number of resources permanently published with digital object identifier (DOI) issued per month



Status	Number
Private	5561
Public	3743
Discoverable	728
Published	652

Cumulative number and size of resources in HydroShare

Statistics as of 5/30/2021

Lessons Learned

- Having a community is critical. Listen and address community needs
- Provide immediate value
  - (fulfilling open data mandate is not immediate value)
  - DOI's for citation and credit
  - Simplified access to computational capacity
- Standards
- Be Interoperable – participate in services oriented architecture Ecosystem
- Partnering to expand use and funding base

HYDROSHARE

Summary

HydroShare is a web based collaboration environment to enable more rapid advances in hydrologic understanding through data sharing, analysis and modeling

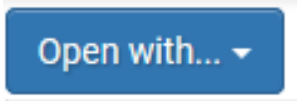
- Sharing and publication of data (DOI)
- Social discovery and added value
- Model and workflow sharing

Collaboration, Reproducibility, Credit, Transparency

- Model input data preparation
- Model execution
- Visualization and analysis (best of practice tools)

Server/Cloud Computation

- Platform independence
- Big data
- Reproducibility
- Interoperability
- Reduced needs for software installation and configuration



HYDROSHARE

Thanks to the HydroShare team!

OAC-1664061
OAC-1664018
OAC-1664119
2017-2021

ACI-1148453
ACI-1148090
2012-2017

HydroShare is operated by CUAHSI with ongoing development through a collaborative project among Utah State University, RENCI University of North Carolina, CyberGIS Center University of Illinois, Tufts, University of Virginia, Brigham Young University, National Center for Atmospheric Research and the University of Washington.

We are looking for developers. If you are interested email dtarb@usu.edu

To learn more

- Publications https://help.hydroshare.org/about-hydroshare/publish/
- Online Help https://help.hydroshare.org/

Access these slides in HydroShare by searching for "niehs2020"

HYDROSHARE

http://www.hydroshare.org

CUAHSI
universities allied for water research

What is CUAHSI?

- **CUAHSI is a 501(c)3 Non-Profit Consortium of about 130 U.S. Academic Institutions, Non-Profits, and International Universities**
- **Mission is to shape the future of water science by:**
  - **Strengthening interdisciplinary collaboration in the water-science community**
  - **Empowering the community by providing research and education services**
  - **Promoting education in the water sciences at all levels**
- **Key Activities**
  - **Community Services, such as workshops, community meeting, training, etc.**
  - **Data and Model Services, including HydroShare and time-series services**

CUAHSI
universities allied for water research

NSF