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

ACI-1148453
ACI-1148090
2012-2017

HydroShare: A Platform for Collaborative Data and Model Sharing in Hydrology

Access these slides in HydroShare by searching for "GeoDaRRS"

David Tarboton, Ray Idaszak, Jeffery S Horsburgh, Daniel P Ames, Jonathan L Goodall, Alva Couch, Richard Hooper, Shaowen Wang, Martyn Clark, Pabitra Dash, Hong Yi, Christina Bandaragoda, Anthony Castronova, Tian Gan, Zhiyu Li, Mohamed Morsy, Maurier Ramirez, Jeffrey Sadler, Dandong Yin, Yan Liu.

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

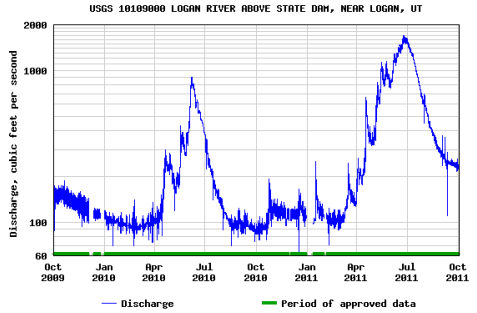
http://www.hydroshare.org

CUAHSI
universities allied for water research

Motivation: Collaborative research

Advancing Hydrologic Understanding

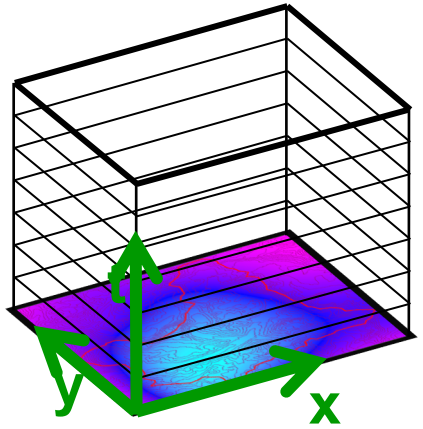
- requires integration of information from multiple sources
- using diverse types of data and models
- may be data and computationally intensive
- requires collaboration and working as a team/community

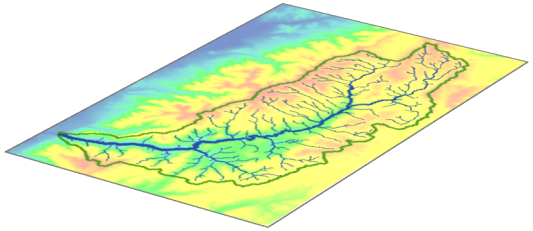


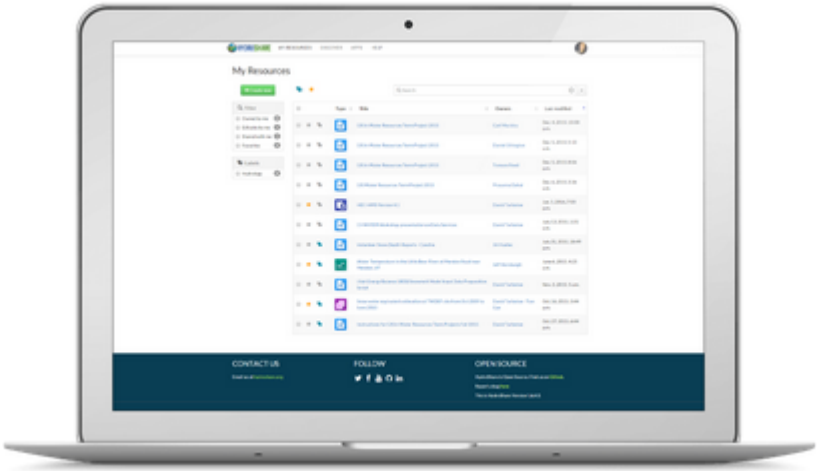
Resources

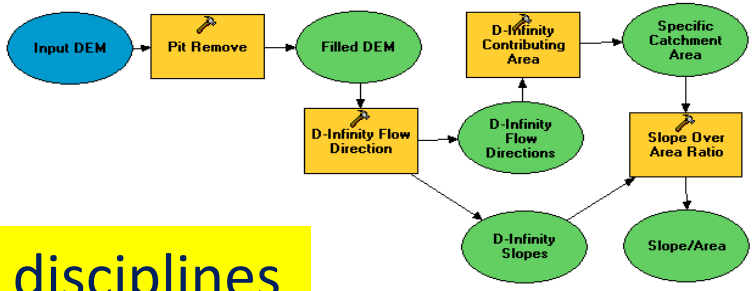
Data and models

Social Objects









This applies for many Geoscience disciplines

“All of the primary datasets collected as part of this project will be made freely and publicly available…”

- iUTAH Proposal Data Management Plan

Many of us put statements like this in our Data Management Plan, but how do we really accomplish this?

From Jeff Horsburgh

Ideal Investigator Data Workflow

- Easily create a digital instance of a dataset or model
- Quickly share it with colleagues (perhaps privately at first)
- Add value through collaboration, annotation, and iteration
- Describe with metadata
- Eventually…share publicly or formally Publish

This is still not as easy as it should be!

From Jeff Horsburgh

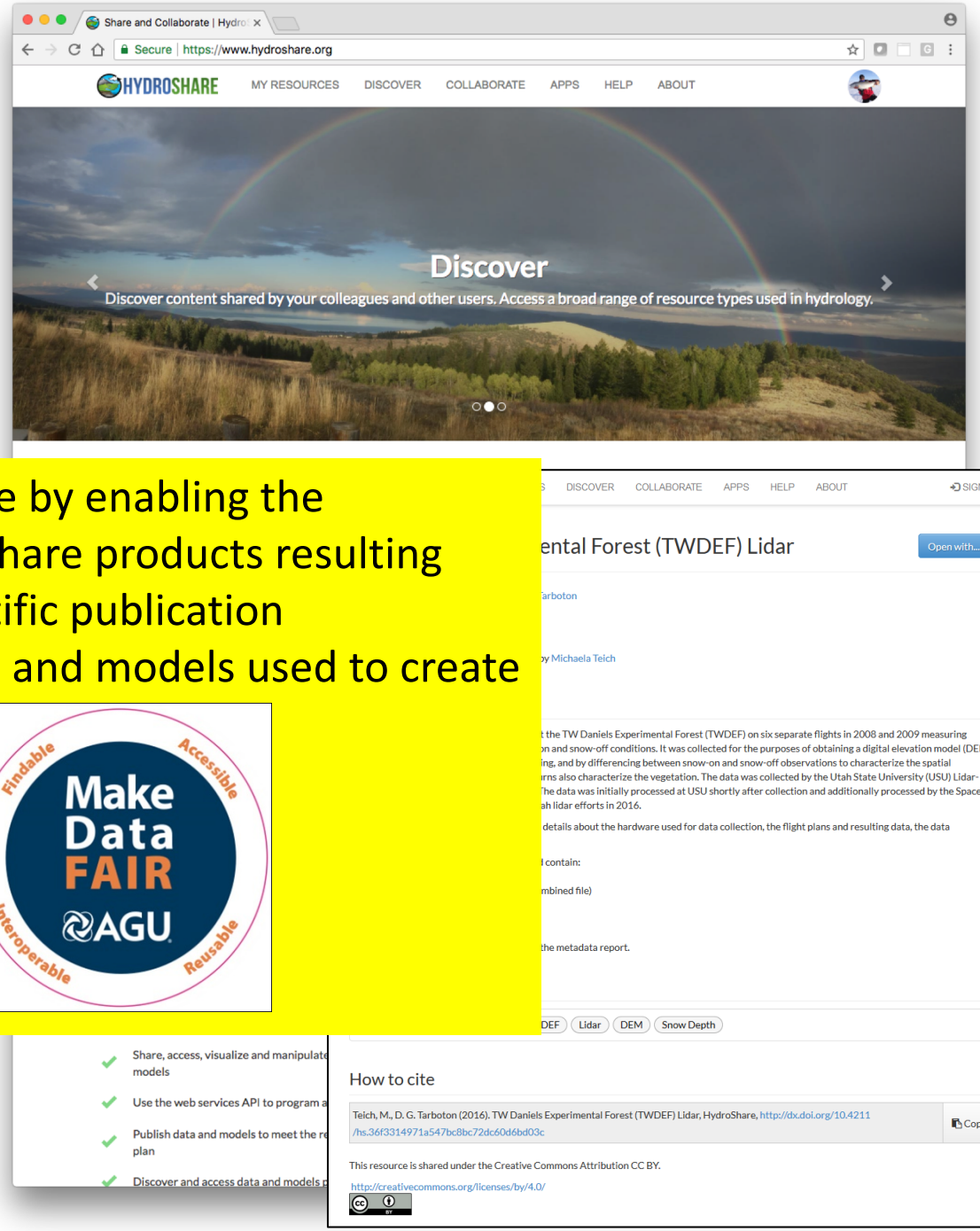
HYDROSHARE

http://www.hydroshare.org

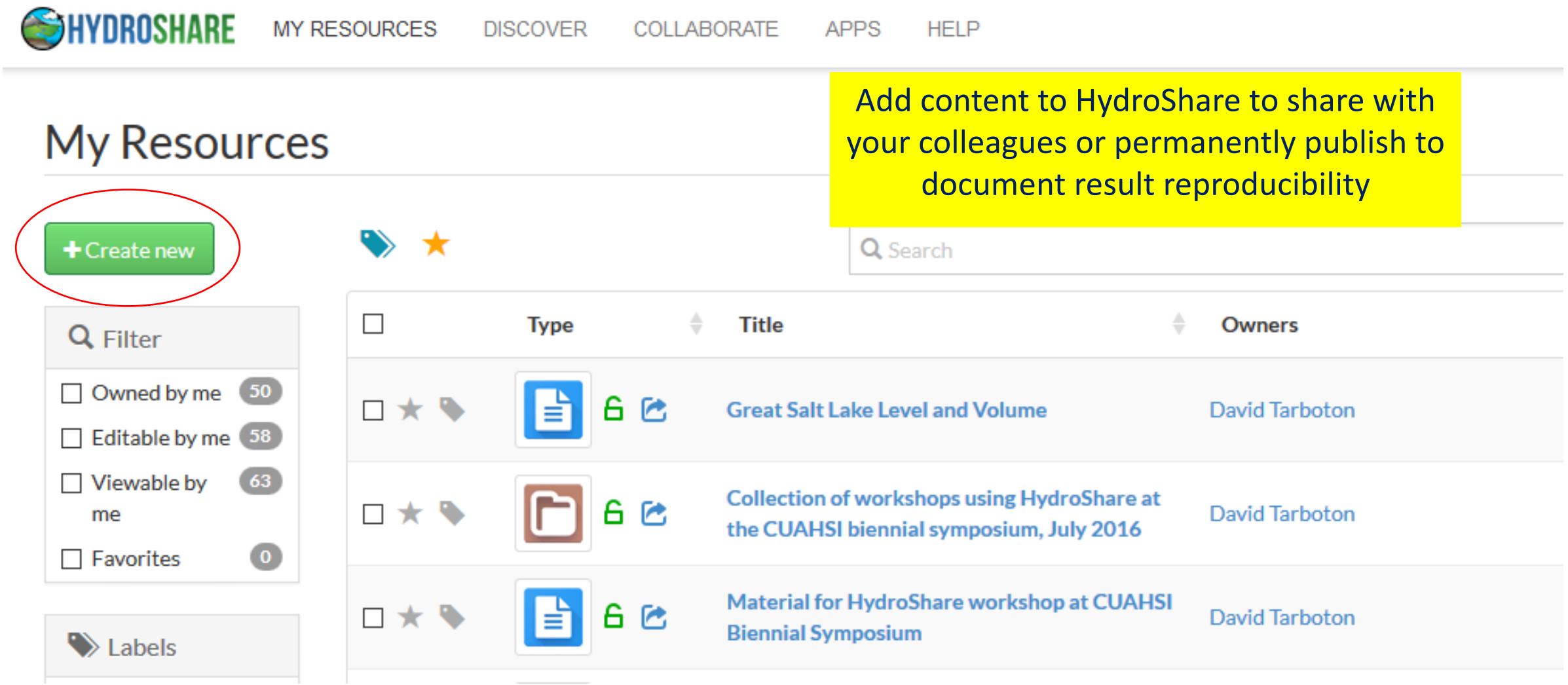
- Web-based Hydrologic Information System operated by the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI)
- Gives you a way to s
  research products
- Has capabilities for c
- Has Links to comput
- Provides permanent
  citable digital object

A system to advance hydrologic science by enabling the 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.

- **F**indable
- **A**ccessible
- **I**nteroperable
- **R**eusable



Collaborative data sharing



Resources (data and models) in HydroShare are objects of collaboration (social objects)

TW Daniels Experimental Forest (TWDEF) Lidar

Open with...

Authors: Michaela Teich · David G. Tarboton
Owners: Michaela Teich
Resource type: Generic
Created: Nov. 17, 2016, 9:11 p.m.
Last updated: Dec. 9, 2016, midnight by Michaela Teich

Abstract

This resource contains lidar data, collected at the TW Daniels Experimental Forest (TWDEF) on six separate flights in 2008 and 2009 measuring surface and canopy properties during snow-on and snow-off conditions. It was collected for the purposes of obtaining a digital elevation model (DEM) to characterize the area for snowmelt modeling, and by differencing between snow-on and snow-off observations to characterize the spatial distribution of snow depth. Canopy lidar returns also characterize the vegetation. The data was collected by the Utah State University (USU) Lidar-Assisted Stereo Imaging (LASSI) laboratory. The data was initially processed at USU shortly after collection and additionally processed by the Space Dynamics Laboratory (SDL) in support of iUtah lidar efforts in 2016.

The metadata report (sdl16-1363-.pdf) gives details about the hardware used for data collection, the flight plans and resulting data, the data processing steps, and a brief error analysis.

Zip files are named by the collection date and contain:
- Terra Scan Binary Files
- LAS Files (one for each flight line and the combined file)
- KML Files (one for each flight line)
- ASC DEM file (1 m resolution)
- PNG Hillshade file

A complete list can be found on pp. 17-22 of the metadata report.

How to cite

Teich, M., D. G. Tarboton (2016). TW Daniels Experimental Forest (TWDEF) Lidar, HydroShare, http://dx.doi.org/10.4211/hs.36f3314971a547bc8bc72dc60d6bd03c

This resource is shared under the Creative Commons Attribution CC BY. http://creativecommons.org/licenses/by/4.0/

For each resource you can
- Manage who has access
  - To edit
  - To view
- Comment or rate
- Obtain unique identifier
- Describe with metadata
- Organize into collections
- Permanently publish with DOI
- Version
- Open with compatible web app

Publishing data and models for reproducibility and trust

Journal of Hydrology 559 (2018) 43–55

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Modeling urban coastal flood severity from crowd-sourced flood reports using Poisson regression and Random Forest

J.M. Sadler [a], J.L. Goodall [a,*], M.M. Morsy [a,b], K. Spencer [c]

[a] Dept. of Civil and Environmental Engineering, Univ. of Virginia, 351 McCormick Rd., P.O. Box 400742, Charlottesville, VA 22904, United States
[b] Irrigation and Hydrau
[c] Deputy Resilience Offi

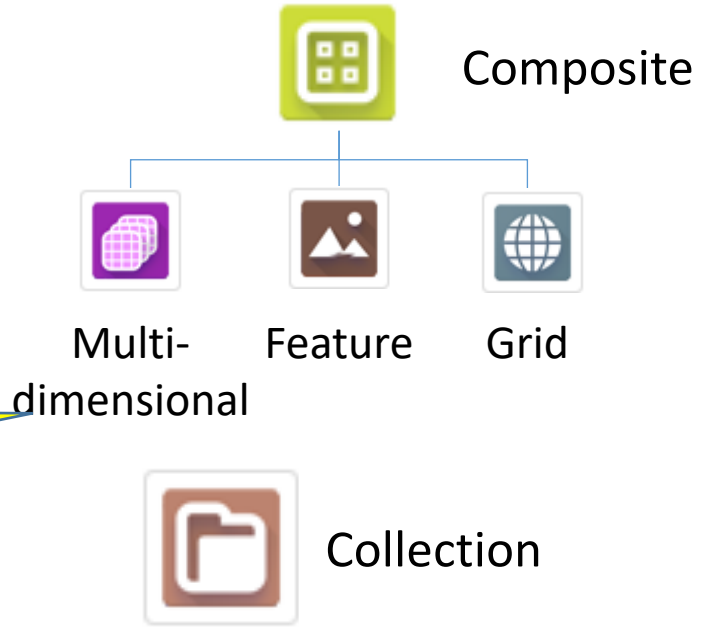
References cite **input data**, **output data**, and **models** and **scripts** published in HydroShare

Sadler, J., 2018a. Input data for flood severity modeling in Norfolk, VA. HydroShare. https://doi.org/10.4211/hs.ff8be5aea3224c15b262bfddd5fb6033.
Sadler, J., 2018b. Output from data-driven model of flood severity in Norfolk, VA. HydroShare. https://doi.org/10.4211/hs.54df00b15c02458685fa3b622f2ecc7b.
Sadler, J., 2018c. Data-driven model script for flood severity modeling in Norfolk, VA. HydroShare. https://doi.org/10.4211/hs.712cd2ce8f604c8f824d6836ee3fcb53.

Resource Organization Concepts

- A composite resource can hold multiple aggregations
  - Each being a different type of data
  - Managed as one discoverable resource
  - One set of access controls (Owners, Editors etc.)
  - One unique identifier
  - One set of resource level metadata
- A collection can hold multiple r
  - Each has own unique identifier
  - Own access control (separate owne and editors etc.)
  - Separate resource level metadata and landing page
- Collections and their members may each be discovered separately
- Composite resources may be members of Collections
- Unique keyword tags form informal collections
  (e.g. “geodarrs”)

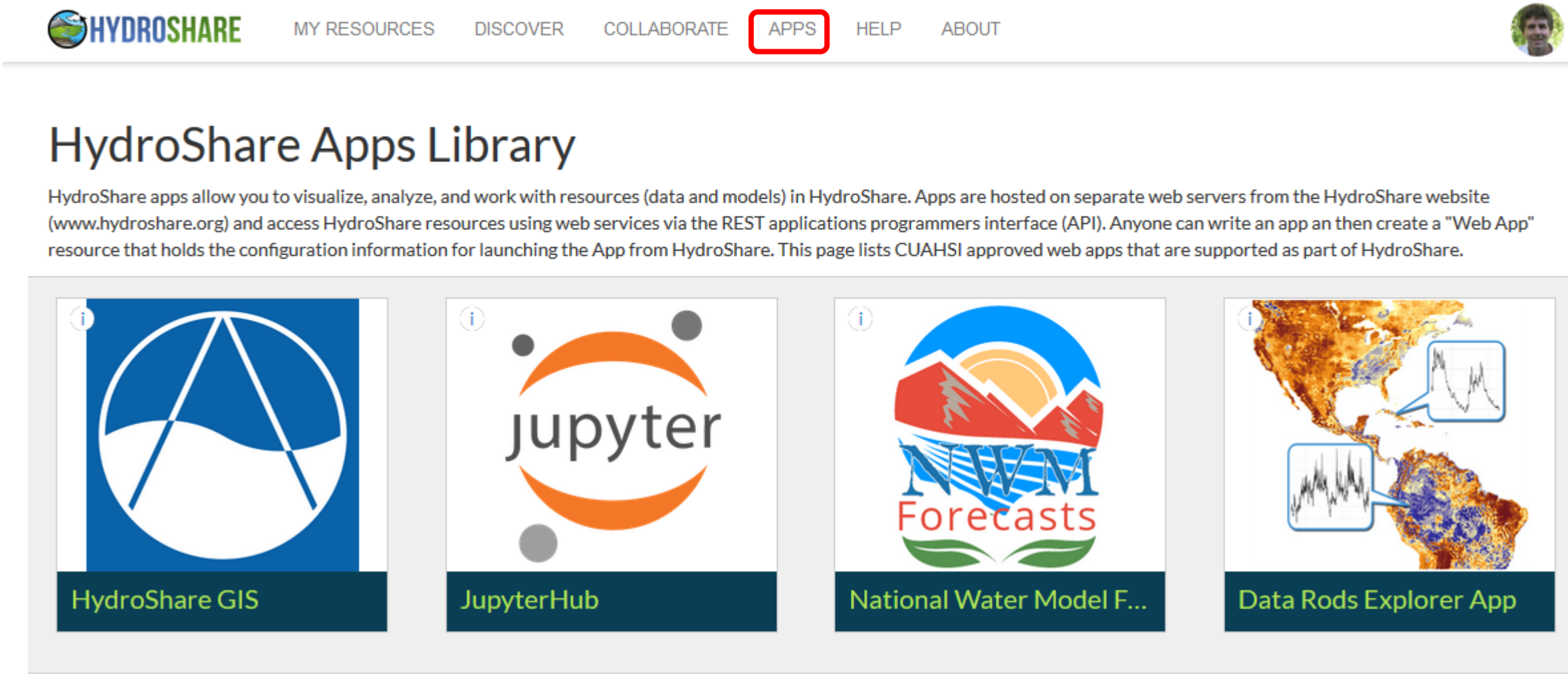
But there is more …



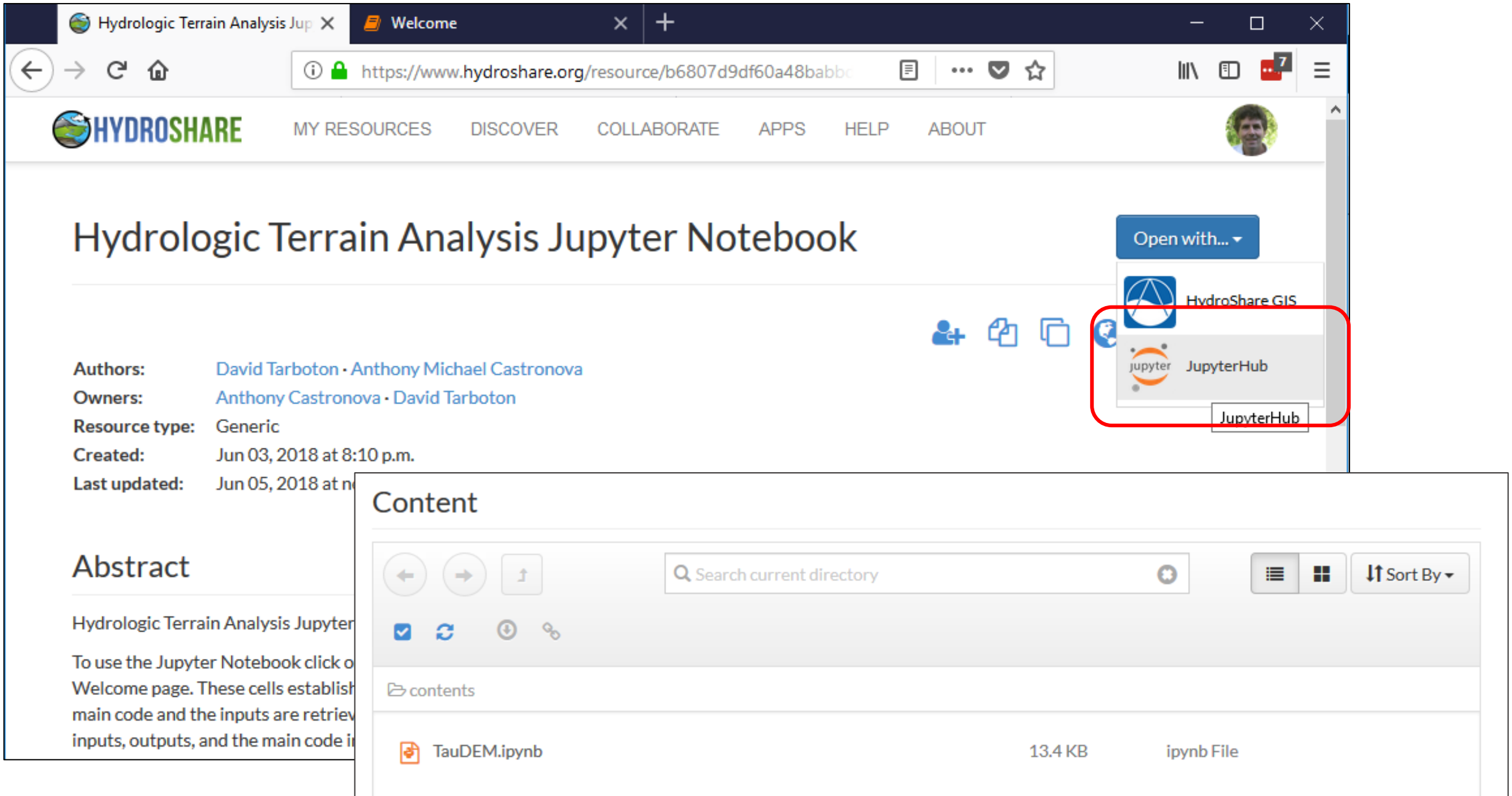
Collection Contents

Title
Great Salt Lake Area Volume Data
Great Salt Lake Level and Volume
Great Salt Lake Basin Digital Elevation Model

Apps act on resources to support web based visualization and analysis http://www.hydroshare.org/apps



JupyterHub App



JupyterHub App Analysis

CUAHSI TauDEM (autosaved) Logout Control Panel Welcome dtarb

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 3

Run Markdown

Hydrologic Terrain Analysis Using TauDEM

The purpose of this notebook is to introduce
Digital Elevation Models (TauDEM) softwa
for watershed delineation and extraction and

This notebook is intended as a brief introduc
functions required to delineate a stream netw
documentation on the use of each TauDEM f
and construct other analyses to meet your ne

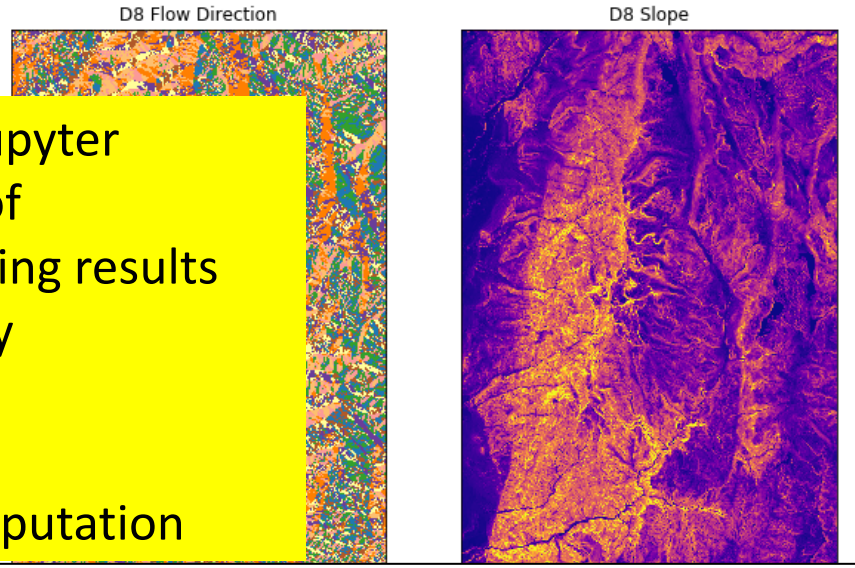
The notebook is organized into the following

1- Preparation, libraries and

2- Digital Elevation Model an

3- Save the Results back into

```
In [7]: grid.plot(raster=['demp.tif', 'demsd8.tif'],
                  title=['D8 Flow Direction', 'D8 Slope'],
                  cm=['Paired','plasma'],
                  cm_scale=[(None,None), (0,1)])
```



Write and execute code in a Jupyter Notebook, acting on content of HydroShare resources and saving results back to HydroShare Repository

- Reproducibility
- Collaboration
- Access to enhanced computation

```
In [11]: # Find the files that are not folders.  (The initial folders are already there)
files = !find . -maxdepth 1 -type f
print(files)
hs.addContentToExistingResource(resid, files)
```

```
['./demfel.tif', './demsd8.tif', './demp.tif', './demad8o.tif', './demsrc.tif', './demtree.dat', '
./demcoord.dat', './demnet.shp', './demnet.shx', './demnet.dbf', './demnet.prj', './demw.tif', './
demord3.tif']
Successfully Added Content Files
```

What I would like and why?

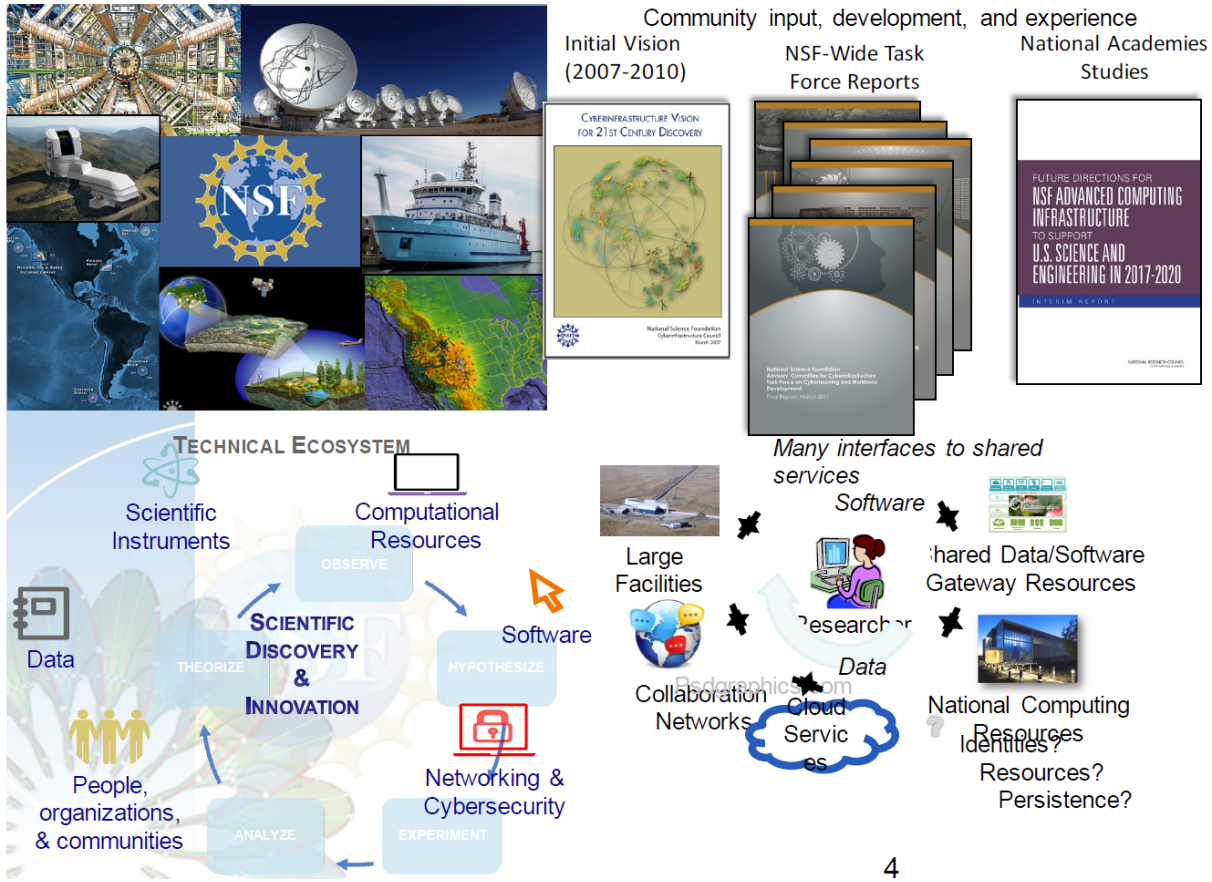
A platform for collaboration and computation that integrates data storage, organization, discovery, and programmable actions through web applications (web apps) and that allows researchers to easily employ services beyond the desktop to make data storage and manipulation more reliable and scalable, while improving ability to collaborate and reproduce results.

A complete web (cloud/server) based innovation environment

- Ability to work with large datasets
- Ability to hold data public and private
- Integrated data and compute
- Support for the complete data life-cycle
- Easy to use tools and coding environment
- Compatibility and platform independence

Interoperability

- 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

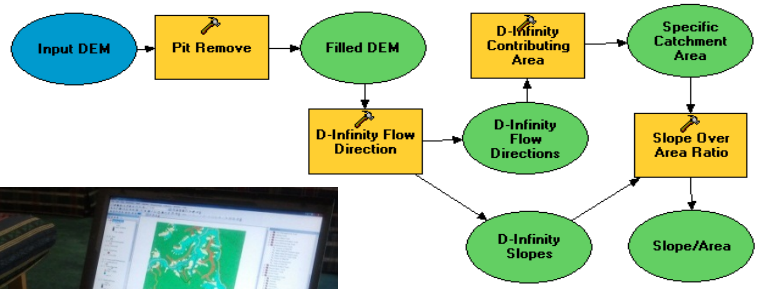


NSF vision for a cyberinfrastructure of many interfaces to shared services [Rajiv Ramnath, NSF Division of Advanced Cyberinfrastructure https://doi.org/10.6084/m9.figshare.4676173]

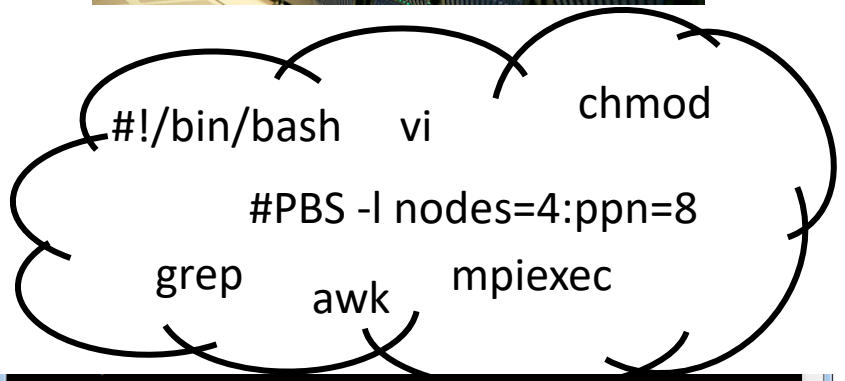
Compute

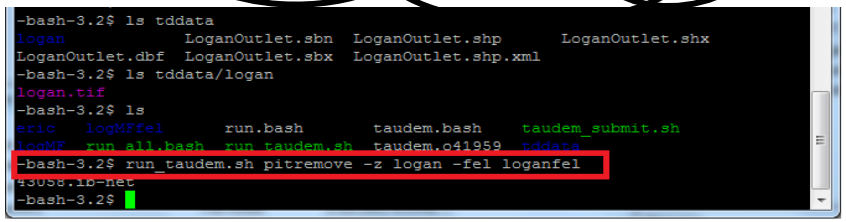
A digital divide

Hydrologic Experimentation and Modeling



Data Intensive High Performance Computing





Enable, for non-HPC specialists, the capability to use HPC Resources

Data Management

- Not a separate topic (last part of proposal)
- Is actually about doing good science
- Knowing what the data means (metadata)
- Knowing suitability for purpose (metadata)
- Needs to be integrated into research workflows

Summary

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

- Sharing and publication of data (DOI)
- Social discovery and added value
- Model 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
- Reduce needs for software installation and configuration

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

Thanks to the HydroShare team!

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.

To learn more

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

HYDROSHARE

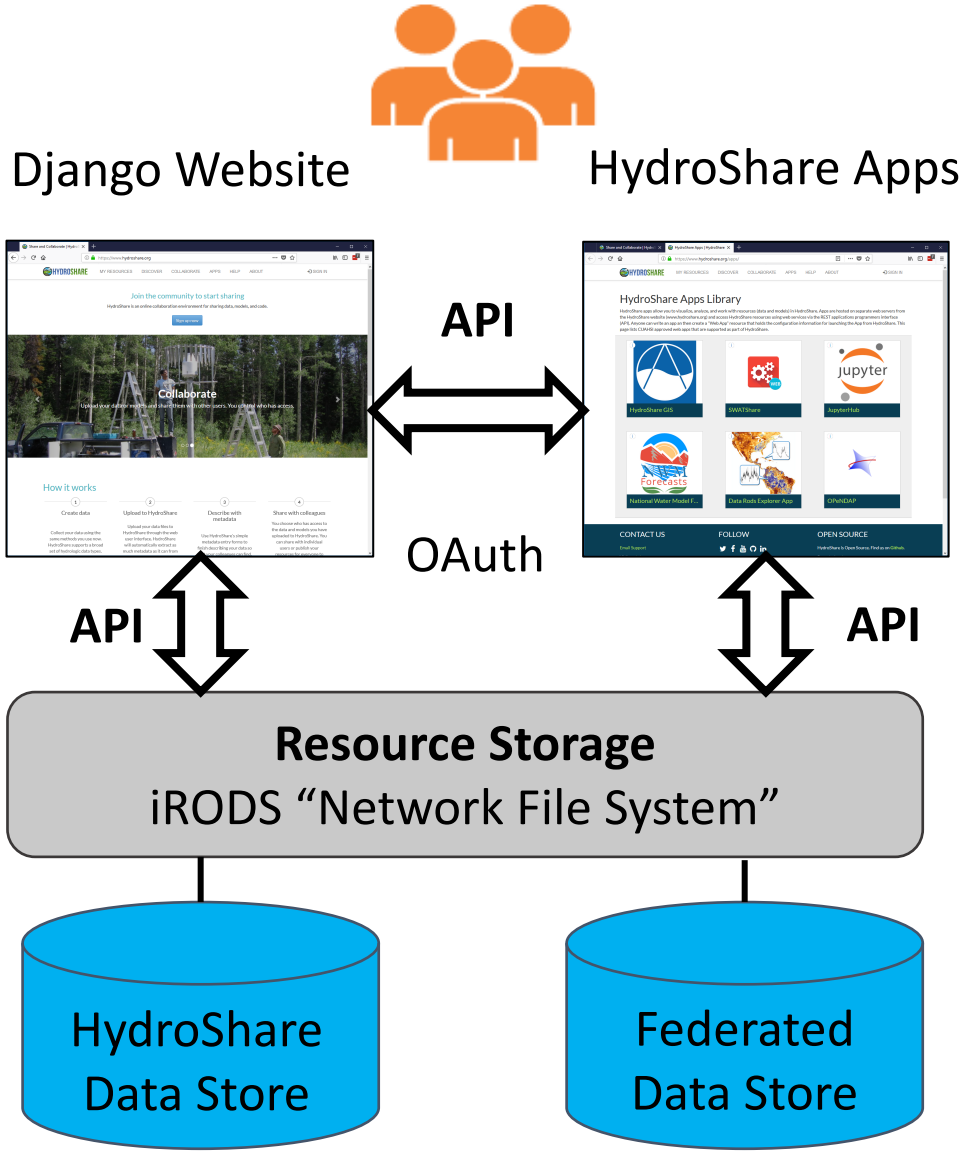
http://www.hydroshare.org

CUAHSI
universities allied for water research

Conceptual Architecture

Resource exploration

- Organize and annotate your content
- Manage access



Distributed file storage

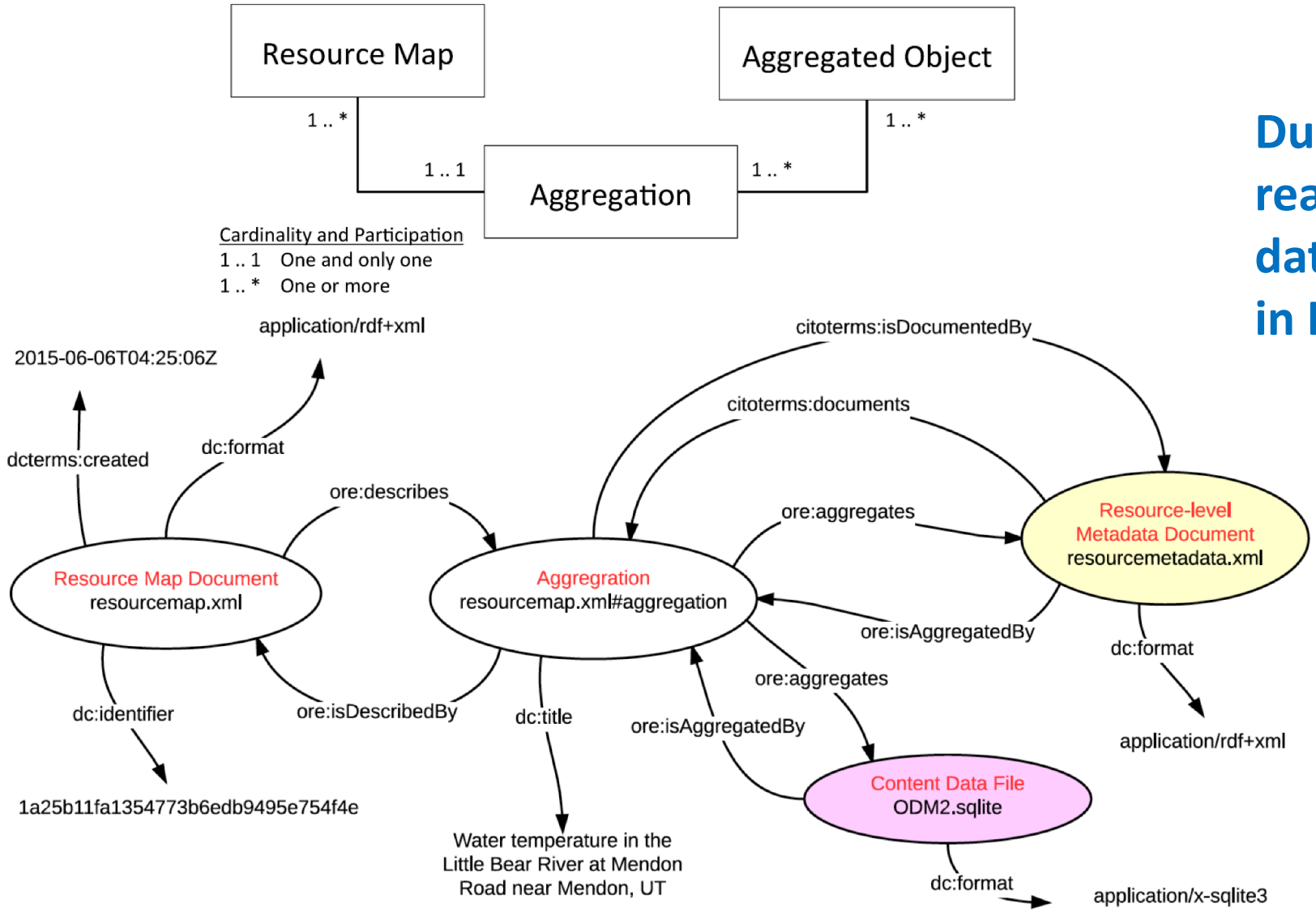
Actions on Resources

- Web software to operate on content you have access to (Apps)
- Extensibility

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

- SWATShare (Hubzero)
- JupyterHub
- Unidata – THREDDS
- NWM Viewer
- Multiple other Tethys Apps

OAI-ORE standard based Resource Data Model

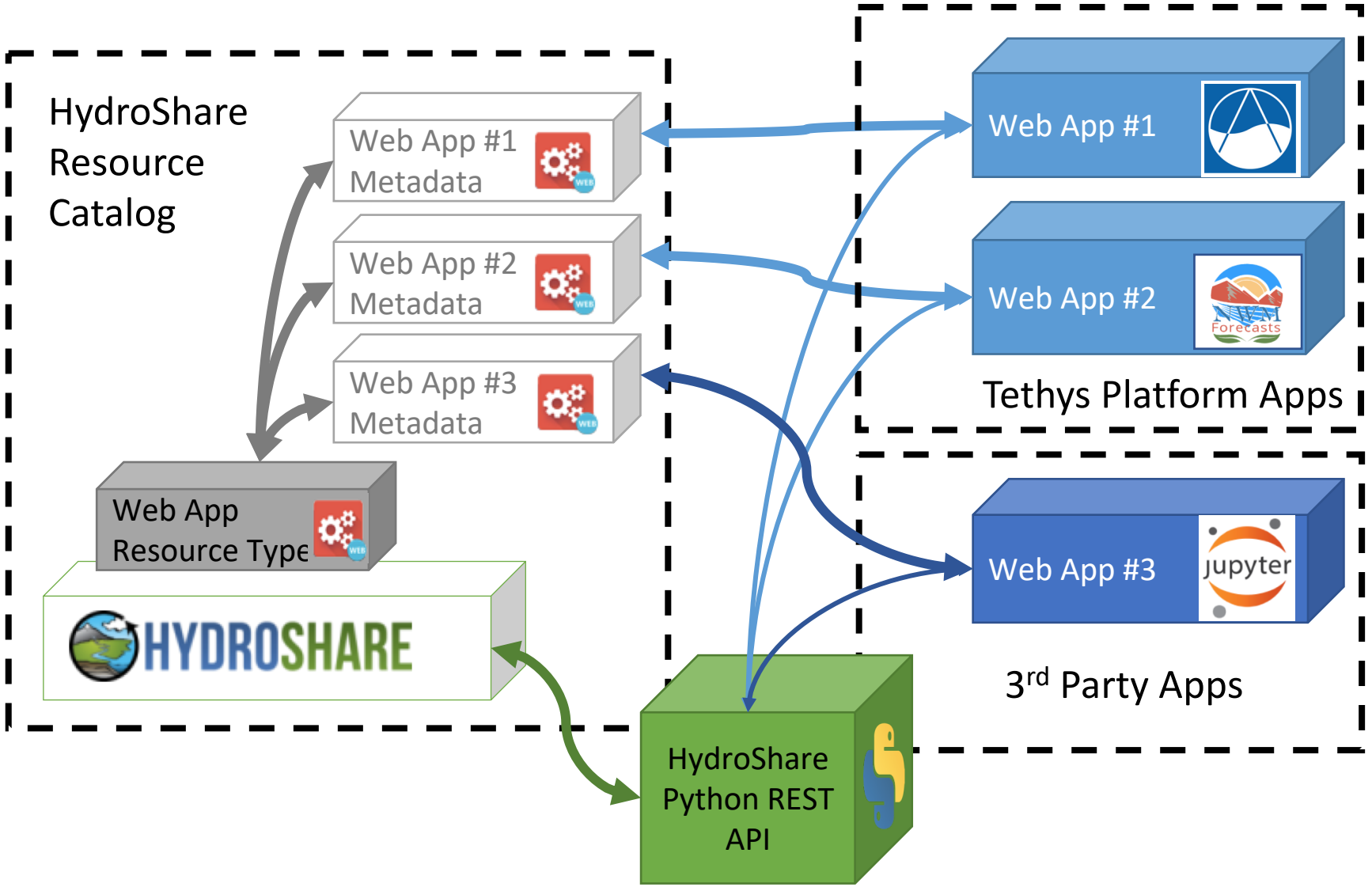


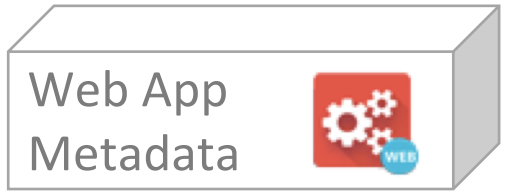
Dublin Core machine readable metadata and data model to make data in HydroShare FAIR

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.

HydroShare Web Apps Architecture

Anybody can create a web app on any web server and configure a web app resource 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:

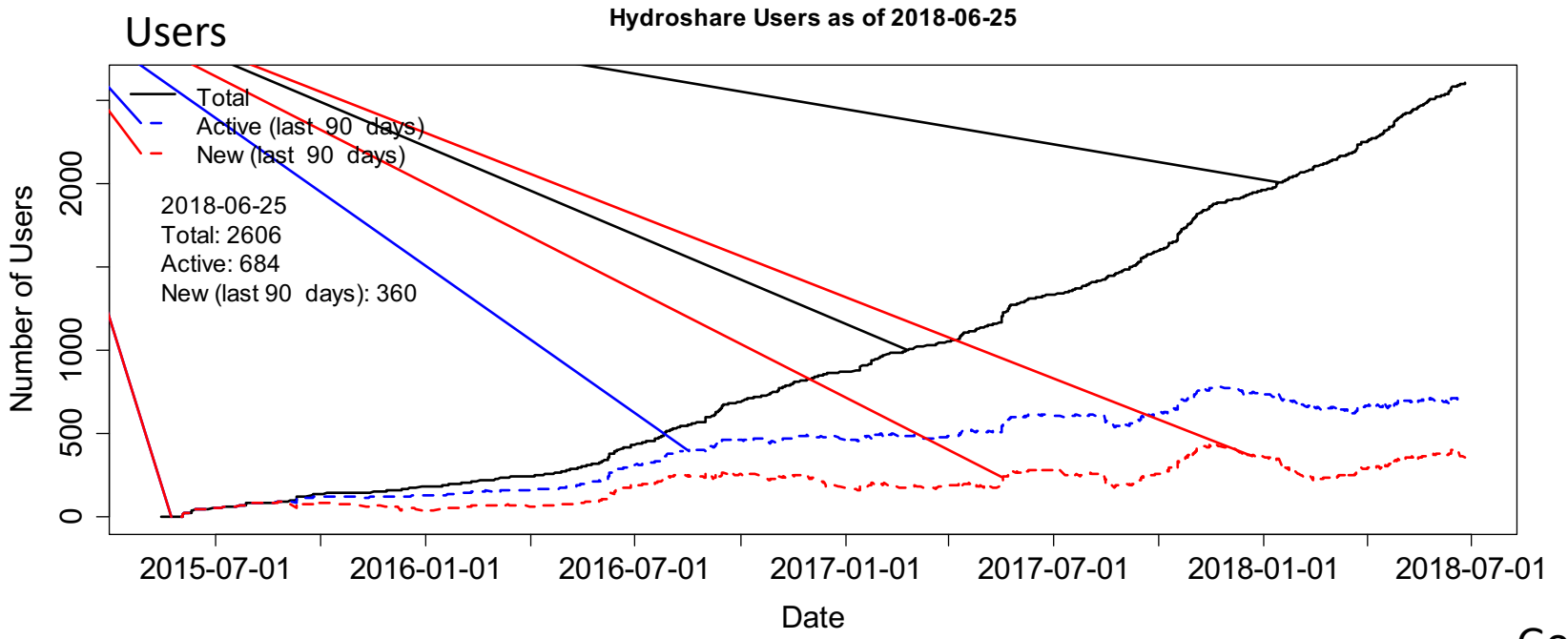
https://apps.hydroshare.org/apps/hydroshare-gis/?res_id=${HS_RES_ID}

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

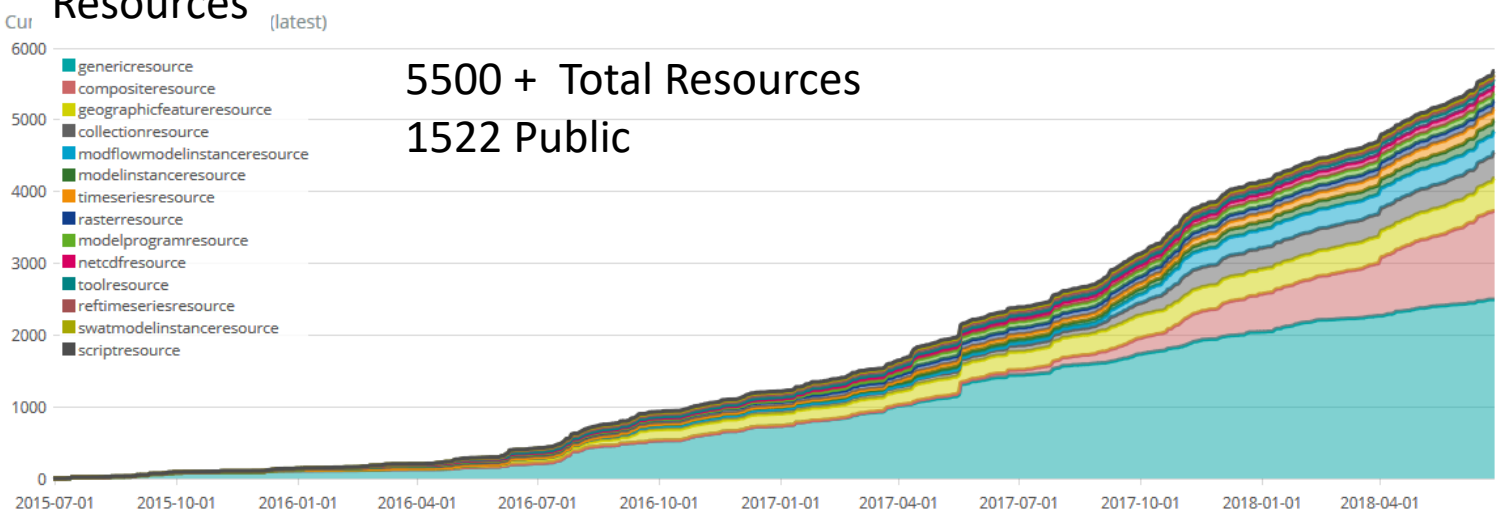
http://hyrax.hydroshare.org/opendap/${HS_RES_ID}/data/contents/

Audience and User base

Users



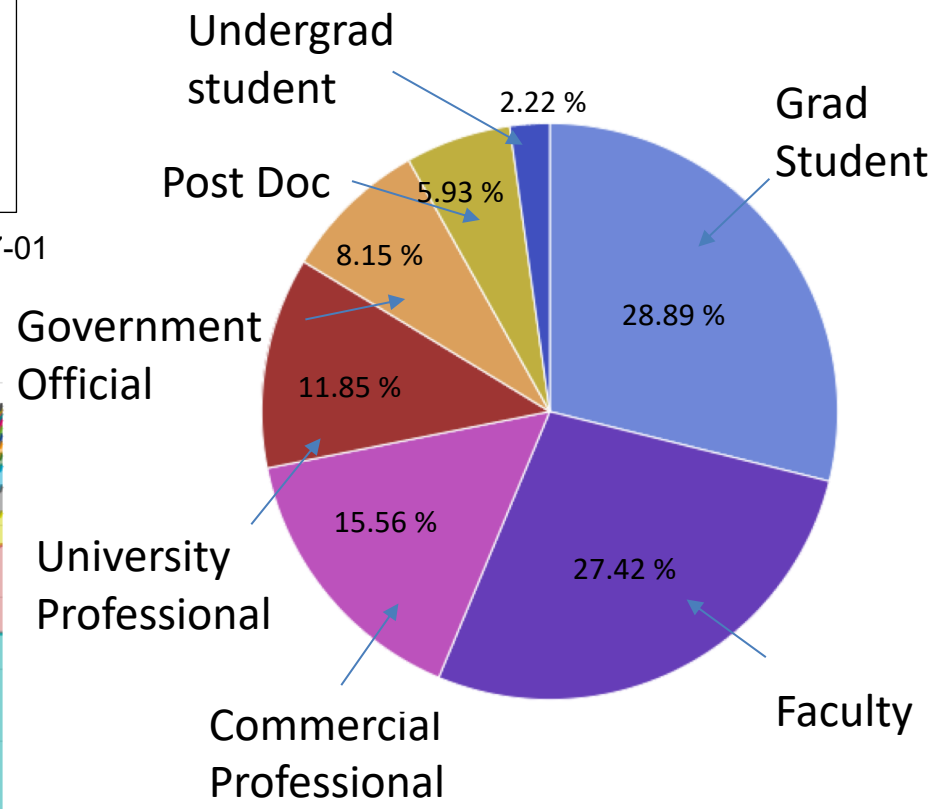
Resources



5500 + Total Resources

1522 Public

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



For users who have indicated type in their profile

Statistics as of 6/25/18 , see https://help.hydroshare.org/about-hydroshare/usage-metrics/

Example: Data from 2017 US Hurricanes

HYDROSHARE MY RESOURCES DISCOVER COLLABORATE APPS HELP ABOUT

Hurricane Harvey 2017 Collection

Authors:
Owners:
Resource typ
Created:
Last updated

Abstrac

HYDROSHARE MY RESOURCES DISCOVER COLLABORATE APPS HELP ABOUT

Hurricane Irma 2017 Collection

Authors:
Owners:
Resource type:
Created:
Last updated:

HYDROSHARE MY RESOURCES DISCOVER COLLABORATE APPS HELP ABOUT

Hurricane Maria 2017 Collection

Authors:
Owners:
Resource typ
Created:
Last updated

HYDROSHARE MY RESOURCES DISCOVER COLLABORATE APPS HELP ABOUT

Find Groups My Groups

CUAHSI 2017 Hurricane Data Community

To share data from Hurricanes Harvey, Irma and Maria that impacted the US and Caribbean region in 2017.

Following Hurricanes Harvey, Irma and Maria that had significant impacts to parts of the US and islands in the Caribbean there has been much activity to assemble, document and archive data from these events. This data is intended to support research to improve our understanding of and capability to prepare for and respond to such extreme events in the future. This group has been created as a community within HydroShare for users interested in these data archives. This is open for anyone to join. Anyone who joins can share data (HydroShare resources) with this group. Sharing resources with this group makes them

Submerged I-10 in Houston,Tx

I-10, Houston, Texas, 77079

Portions of the Interstate 10 remains flooded in the wake of Hurricane Harvey after it dumped up to 50 inches of rain in Houston, Texas, on Aug. 29, 2017. (Photo: Marcus Yam / Los Angeles Times via Getty Images