

Community Approaches to Hydrologic Data and Model Sharing using Cloud Resources

David Tarboton, Ray Idaszak, Jeffery Horsburgh, Dan Ames, Jon Goodall, Larry Band, Venkatesh Merwade, Alva Couch, Jennifer Arrigo, Rick Hooper, David Valentine

<http://his.cuahsi.org>

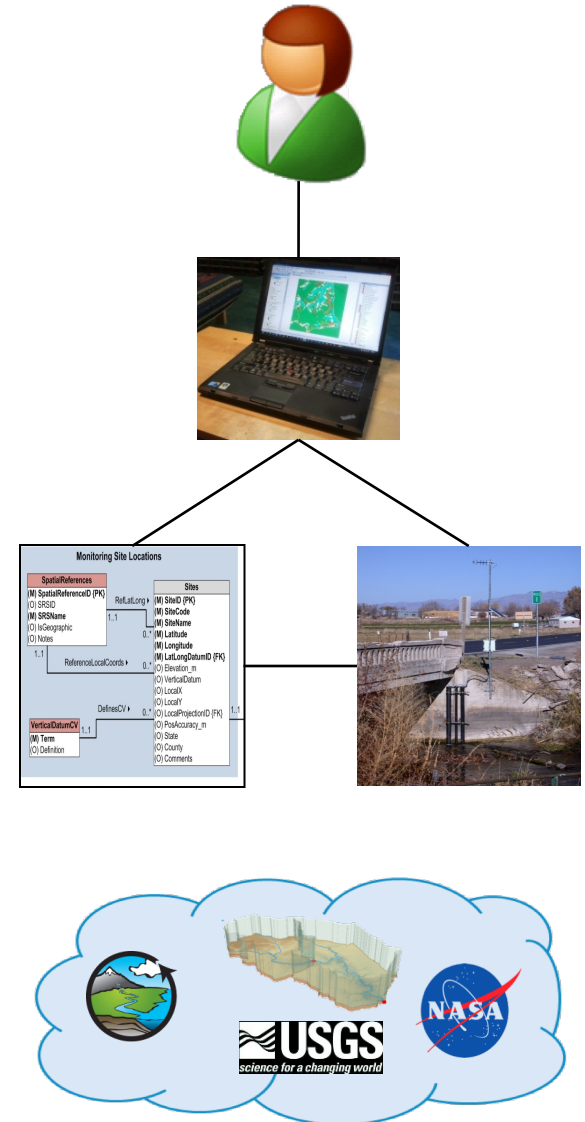
<http://www.hydroshare.org>



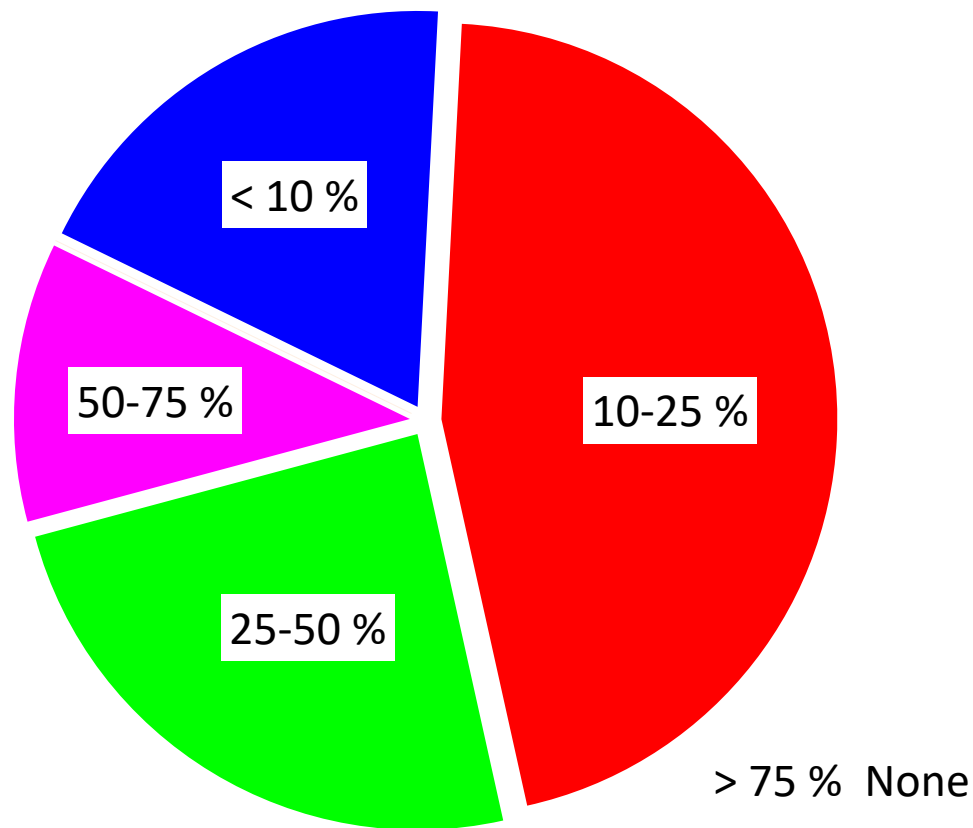
EAR 0622374 (2007-2012)
OCI-1148453 (2012-2017)
OCI-1148090 (2012-2017)

Outline

- Cloud Computing
- The CUAHSI HIS
 - A Services-Oriented Architecture Based System for Sharing Hydrologic Data
- HydroShare
 - A Web-Based Collaborative Environment for the Sharing of Hydrologic Data and Models



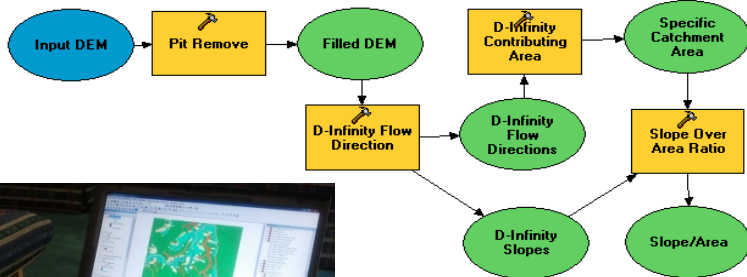
What proportion of your research time do you spend on preparing or preprocessing data into appropriate forms needed for research purposes?



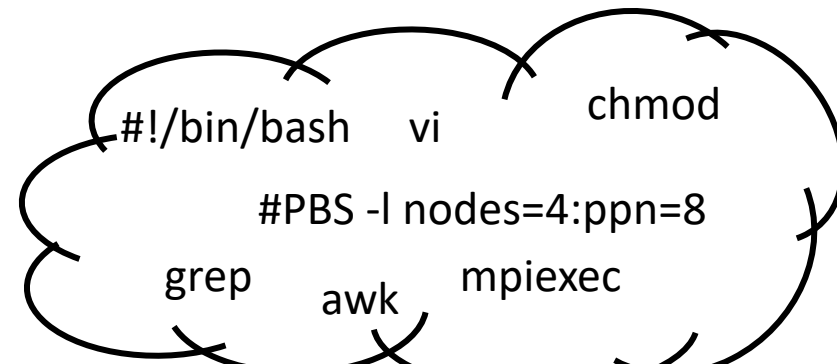
Do you have the access or know how to take advantage of advanced computing capability?

Researchers

- Experimentalists
- Modelers

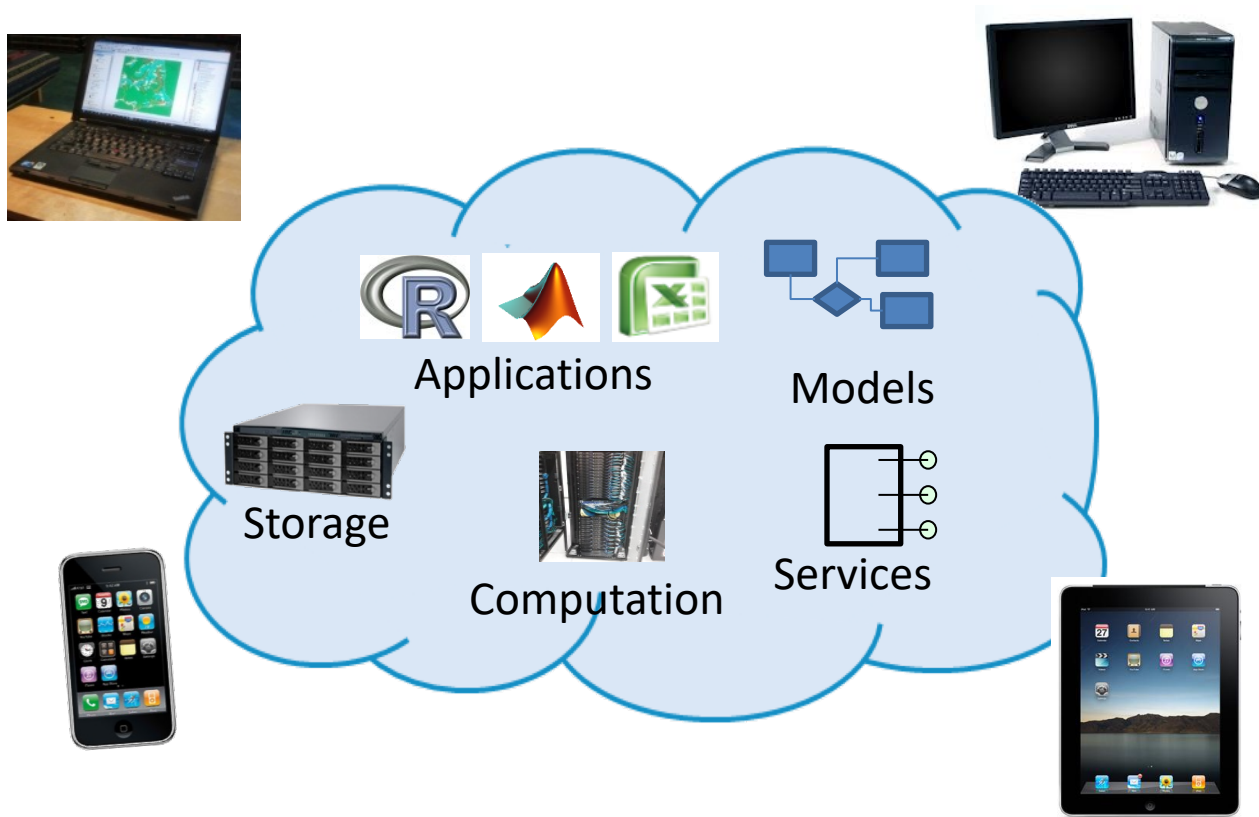


HPC Specialists



```
-bash-3.2$ ls tddata
Logan      LoganOutlet.shn  LoganOutlet.shp      LoganOutlet.shx
LoganOutlet.dbf  LoganOutlet.shx  LoganOutlet.shp.xml
-bash-3.2$ ls tddata/logan
logan.tif
-bash-3.2$ ls
eric  logM5fel  run.bash  taudem.bash  taudem_submit.sh
icowf  run_all.bash  run_taudem.sh  taudem.o41959  tddata
-bash-3.2$ run_taudem.sh pitremove -z logan -fel loganfel
43058.lb-nec
-bash-3.2$
```

Cloud Computing

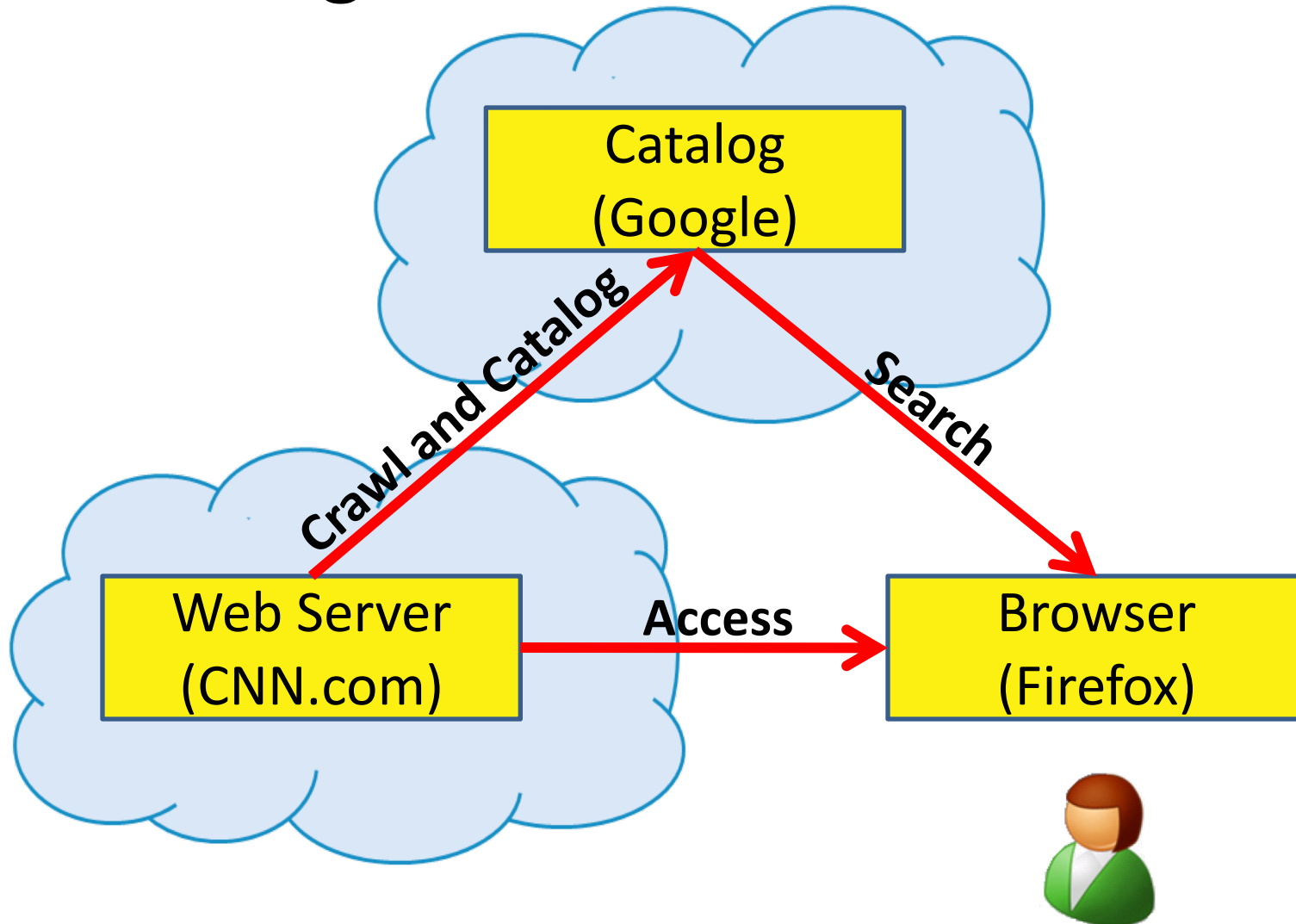


Wikipedia: Cloud computing is the use of **computing resources** (hardware and software) that are delivered as a **service** over a **network** (typically the Internet)

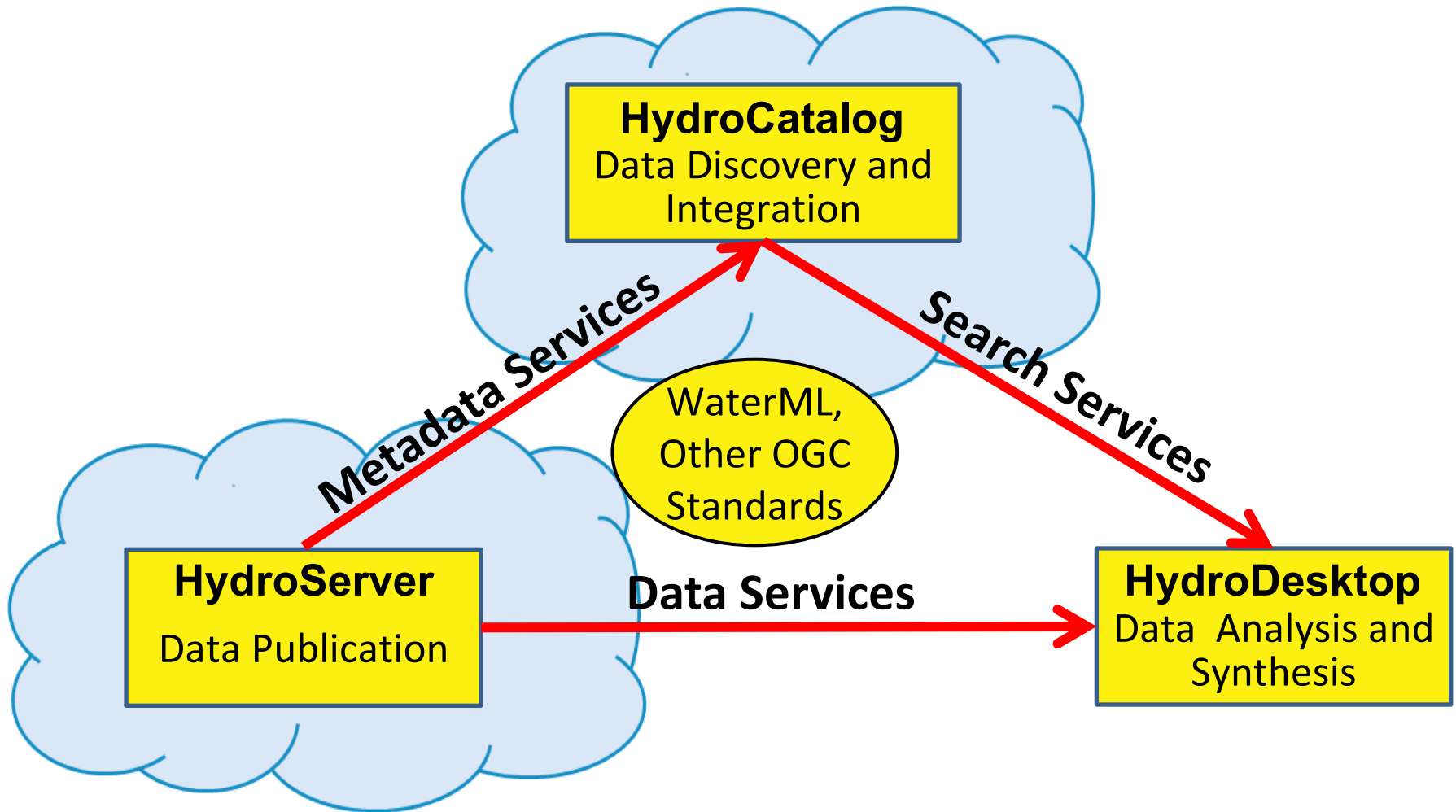
Google, Amazon, Microsoft, Apple, DropBox

XSEDE, Condor, BOINC

Cloud Services-Oriented Architecture Paradigm of the World Wide Web



CUAHSI Hydrologic Information System: A Services-Oriented Architecture Based System for Sharing Hydrologic Data



Hydrologic Data Challenges

- From dispersed federal agencies
- From investigators collected for different purposes
- Different formats
 - Points
 - Lines
 - Polygons
 - Fields
 - Time Series

Data Heterogeneity

Water quality



Water quantity



Rainfall and
Meteorology



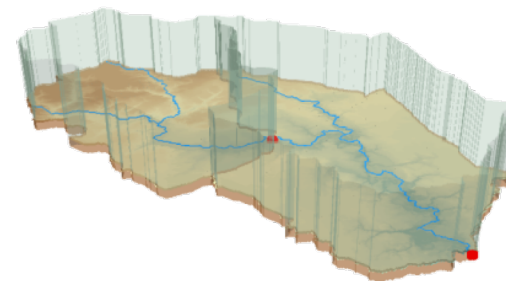
Soil water



Groundwater

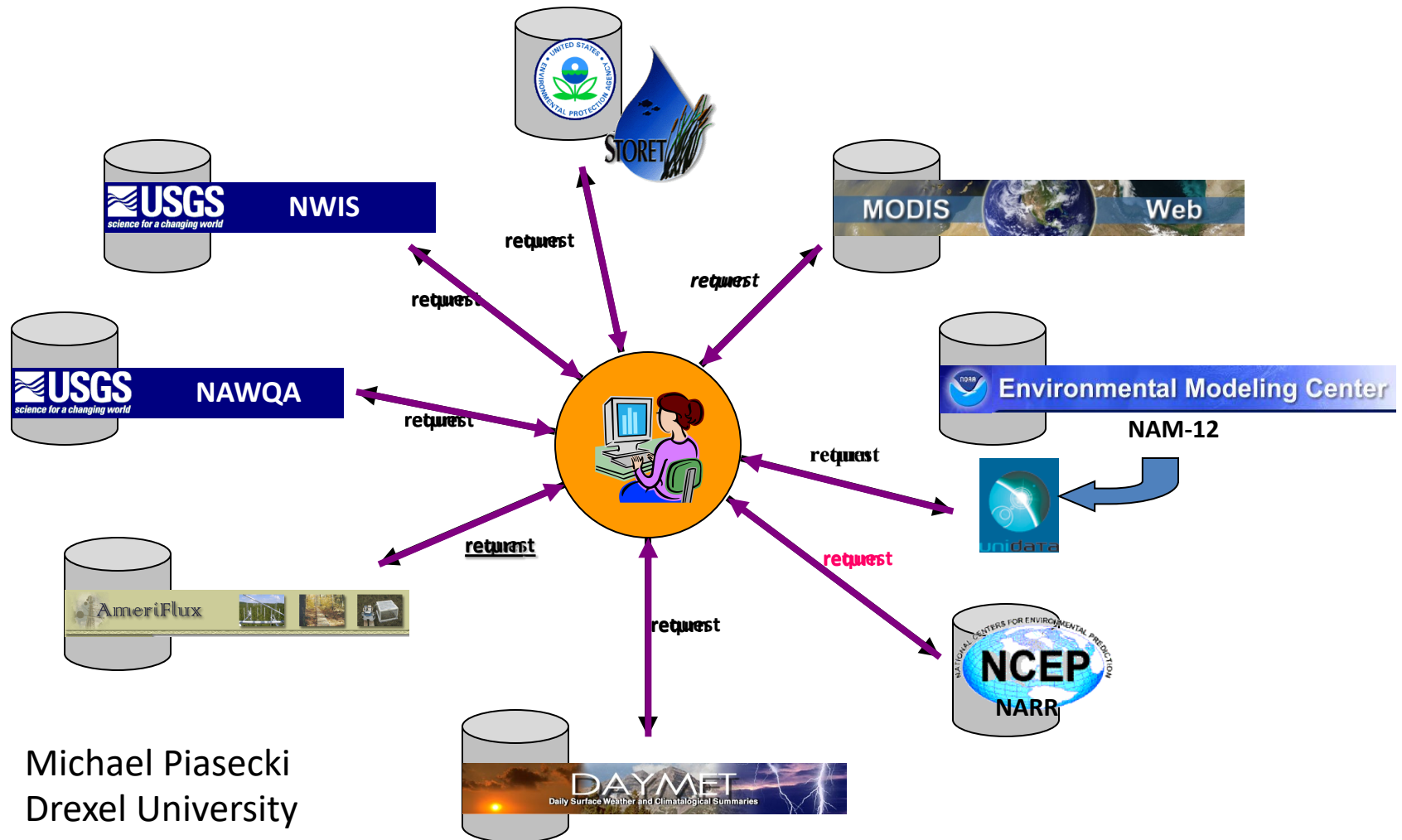


GIS



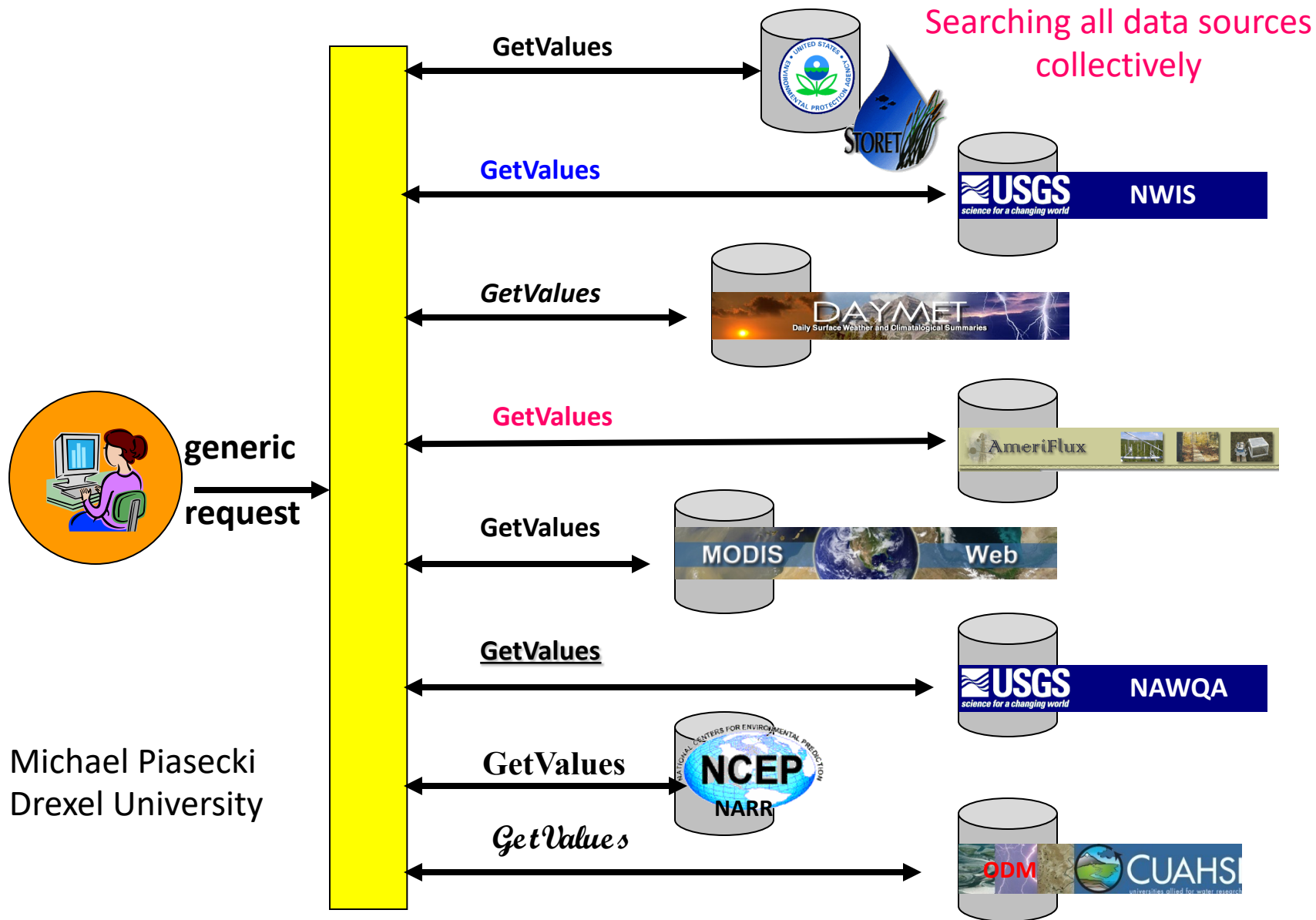
Data Searching – What we used to have to do

Searching each data source separately



Michael Piasecki
Drexel University

What CUAHSI HIS enables



Michael Piasecki
Drexel University

HIS Example.

1. Delineate Watershed using EPA Web Services

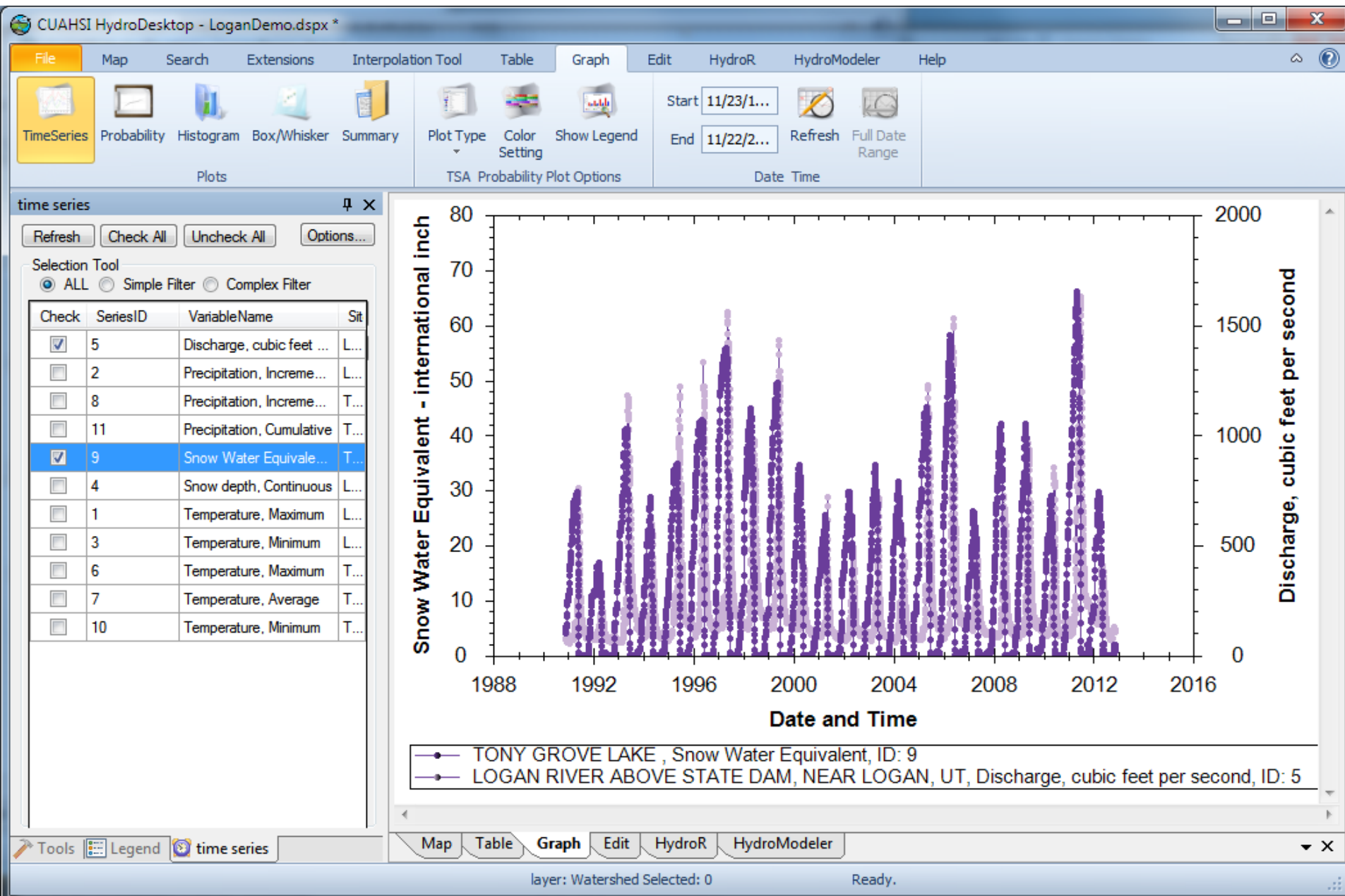
The screenshot displays the CUAHSI HydroDesktop software interface. The title bar reads "CUAHSI HydroDesktop - DotSpatial.SampleProjects.NorthAmerica.dspx *". The menu bar includes File, Map, Search, Extensions, Interpolation Tool, Table, Graph, Edit, HydroR, HydroModeler, and Help. The toolbar contains various tools: Add Layer..., Remove Layer, Pan, In, Out, To Extents, Previous, Next, Zoom To Layer, Select, Deselect All, Identify, and View Attribute Table. The "Delineate" tool, represented by a blue globe icon with a red crosshair, is highlighted with a red rectangle. The map area shows a topographic view of the Logan River watershed in Utah, with a red arrow pointing to the river. The legend on the left lists layers: Map Layers, Base Map Data, Rivers, Online Basemap, Lakes, us_counties, U.S HUC, U.S. States, NAME, Canada Provinces, and Countries. The status bar at the bottom shows "Longitude: 111°45'49\"

Uses EPA WATERS Web, Mapping, and Database Services at <http://www.epa.gov/waters/geoservices/index.html>

2. Search last 22 years for all data in buffer around watershed

The screenshot displays the CUAHSI HydroDesktop software interface. The main window title is "CUAHSI HydroDesktop - LoganDemo.dsp". The top menu bar includes File, Map, Search, Extensions, Interpolation Tool, Table, Graph, Edit, HydroR, HydroModeler, and Help. The toolbar contains various icons, with "Select Features" and "Search" highlighted with red boxes. The "Search" panel is active, showing a "Keyword" field with "All" entered, a "Time Range" section with "Start" set to "11/23/1990" and "End" set to "11/23/2012", and a "Search" button. The map area shows a topographic view of a watershed area, with a cyan buffer around the Logan River watershed. Several data points are marked with red dots and labeled, including "FRANKLIN BASIN", "BEAVER CK AB BEAVER MTN SKI RESORT", "LOGAN R AB BEAVER CK", "TONY GROVE LAKE 01", "SUMMIT CK @ USFS BNDY", "TEMPLE FK AB LOGAN R", "LOGAN RIVER", "LOGAN R AB RIGHT FK LOGAN R", "LOGAN:HYD FK AND SMTHFLD CANAL AT HEAD NR LOGAN UT", "WA-12-1) 3cbd-1", and "(A-11-1) 3dcd-S1". The left sidebar shows a "Legend" panel with "Map Layers" and "Data Sites" sections, listing various data sources like "NIDIS Upper Colorado R", "NCDC Upper Colorado F", "NRCS SNOTEL Standar", "NWIS Daily Values", "Logan River Observation", "NWIS Unit Values", "GHCN Daily Climate Da", "EPA STORET", "NWIS Ground Water Lev", "NWIS Instantaneous Irre", and "Watershed Point". The bottom status bar shows "layer: Loganbuffer Selected: 1" and "Ready."

3. Download and plot data from multiple sources



4. Analyze with R using the HydroR plugin

The screenshot displays the CUAHSI HydroDesktop interface with the HydroR plugin active. The 'HydroR Tools' section includes 'Start R' and 'Generate R Code' buttons, both highlighted with red boxes. The 'time series' panel shows a table of data series with 'Snow Water Equivale...' selected. The R script window shows the following code:

```
data0 <- getDataSeries(
  seriesID=5,
  SQLite=TRUE,
  startDate="1990-
  endDate="2012-11-
data1 <- getDataSeries(
  seriesID=9,
  SQLite=TRUE,
  startDate="1990-
  endDate="2012-11-
```

The R Graphics window shows a scatter plot with a regression line. The y-axis is labeled '1000' and '1500', and the x-axis is labeled '10', '20', '30', '40', '50', '60' with the title 'smax'.

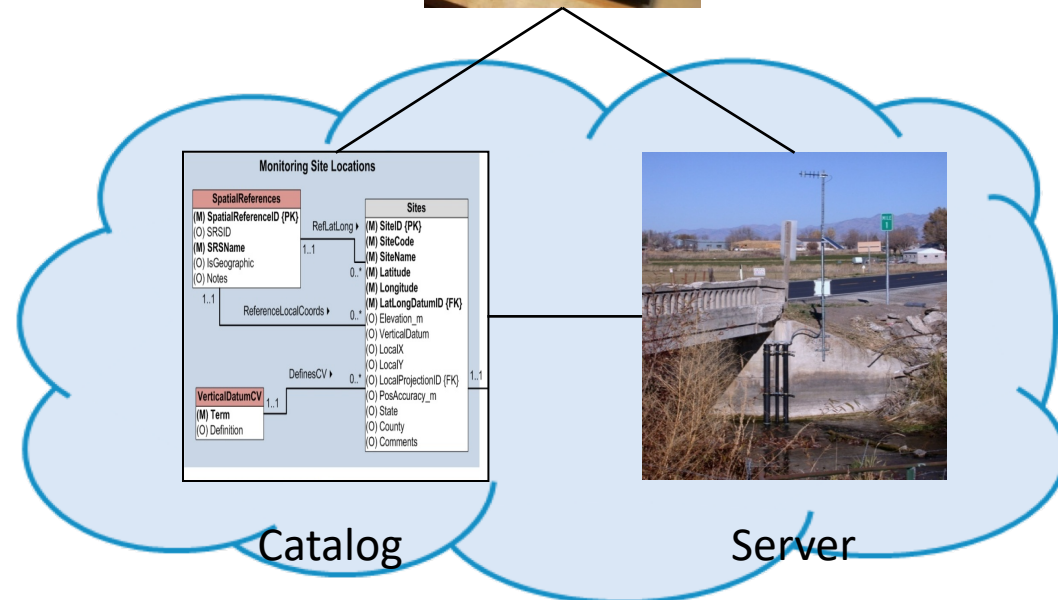
You have the full analysis capability of R at your fingertips for data from multiple sources accessed using HydroDesktop using the HIS cloud data resources.

CUAHSI HIS: A common window on water observations data for the United States unlike any that has existed before

- **Storage** in a community data model
- **Publication** from a server
- Data **access** through internet-based services using consistent language and format
- Tools for **access and analysis**
- **Discovery** through thematic and geographic search functionality
- **Integrated modeling and analysis** combining information from multiple sources

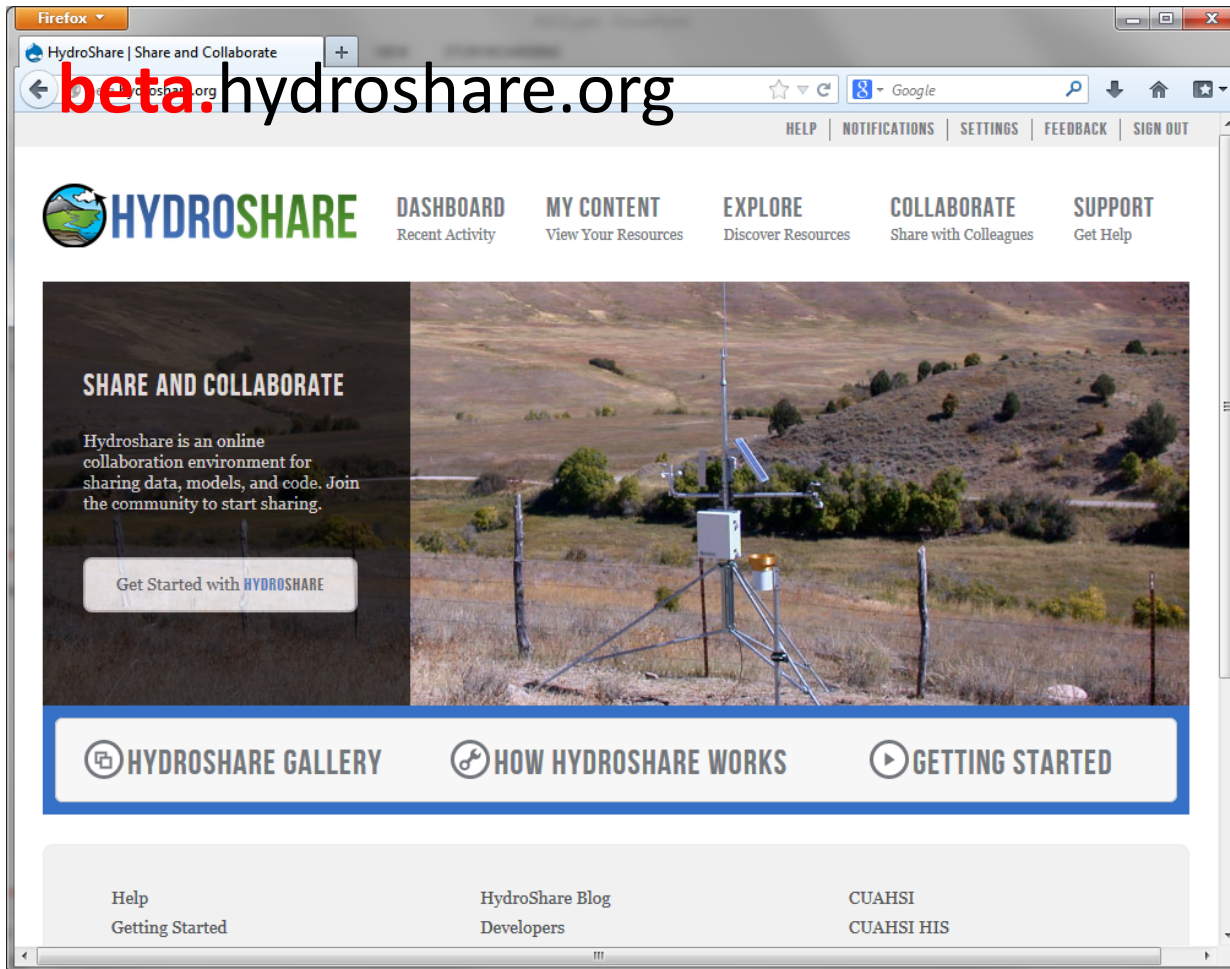


Desktop



Looking to the Future

HydroShare - A web-based collaborative environment for the sharing of hydrologic data and models



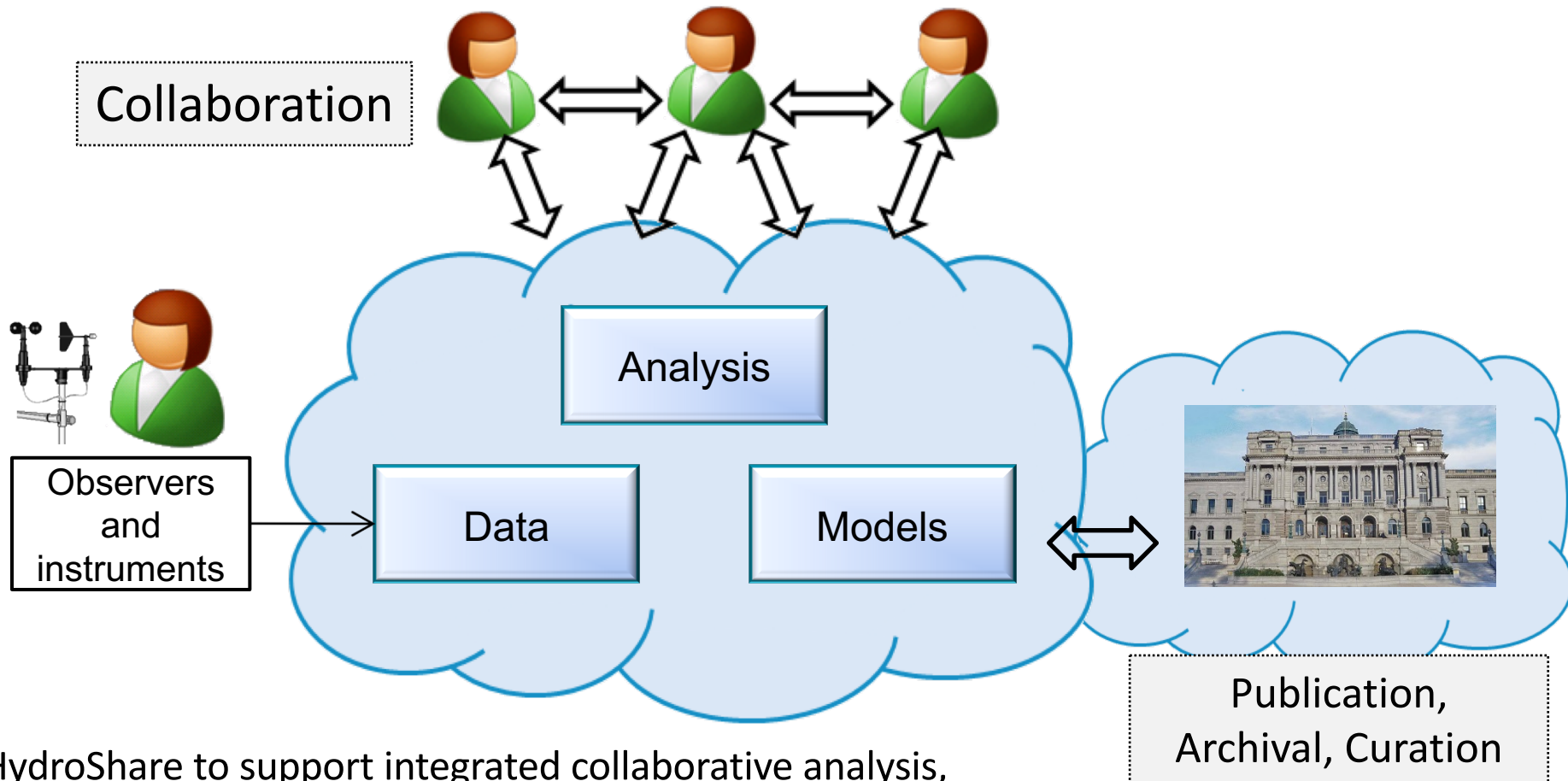
Can sharing data and models be as easy as sharing photos on Facebook or videos on YouTube?

Can finding data and models be as easy as shopping on Amazon?

HydroShare Functionality to be Developed

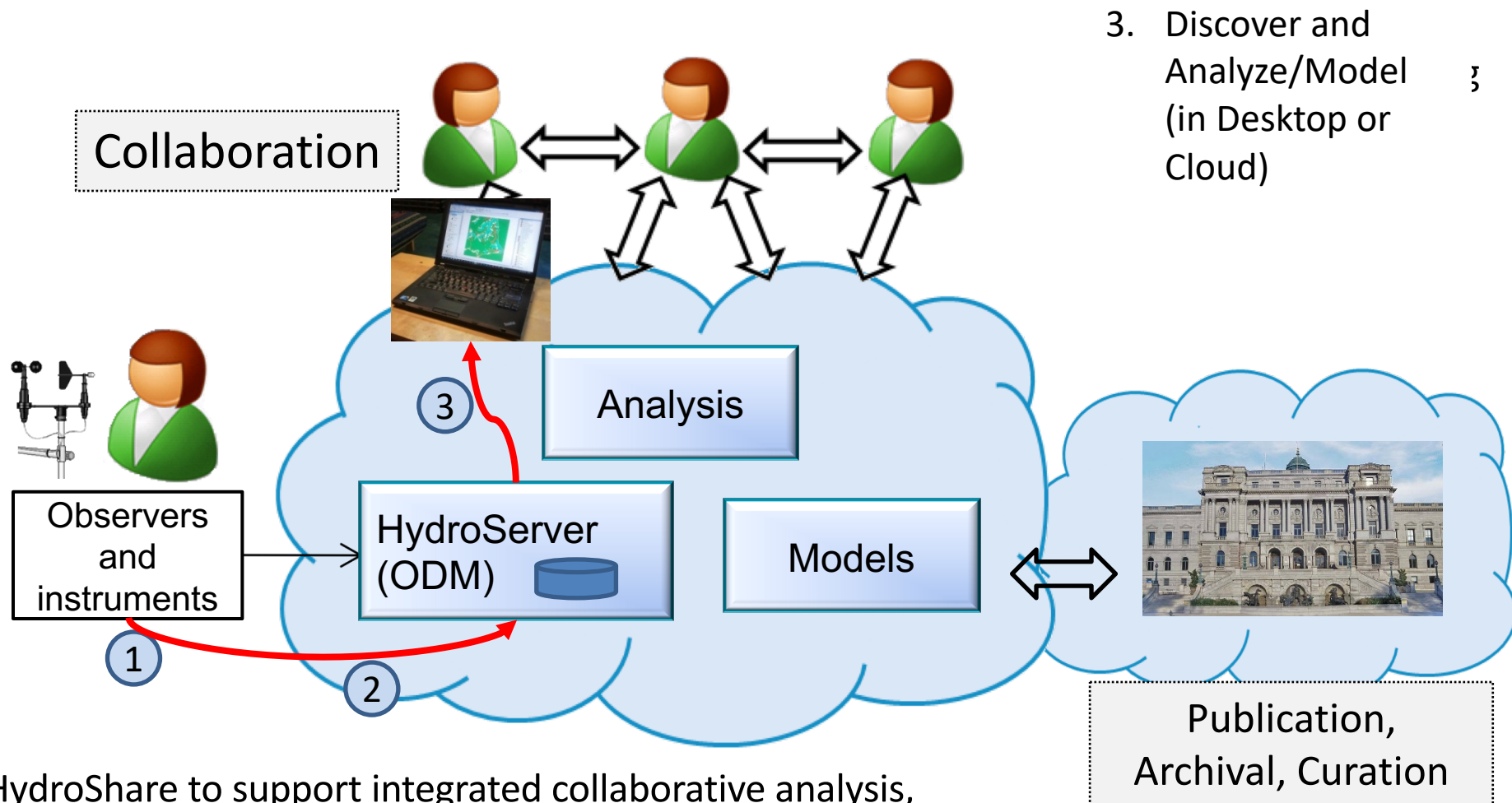
1. A new, **web-based system** for advancing model and data sharing
2. **Sharing** features to HydroDesktop
3. Access **more types of hydrologic data** using **standards** compliant data formats and interfaces
4. Enhance **catalog** functionality that broadens **discovery** functionality to **different data types**
5. New **model** sharing and discovery functionality
6. Facilitate and ease access to use of **high performance computing**
7. New social media and **collaboration** functionality
8. **Links** to other data and modeling systems

Imagine the Possibilities Enabled by Cloud-Based Systems



HydroShare to support integrated collaborative analysis, modeling and data publication

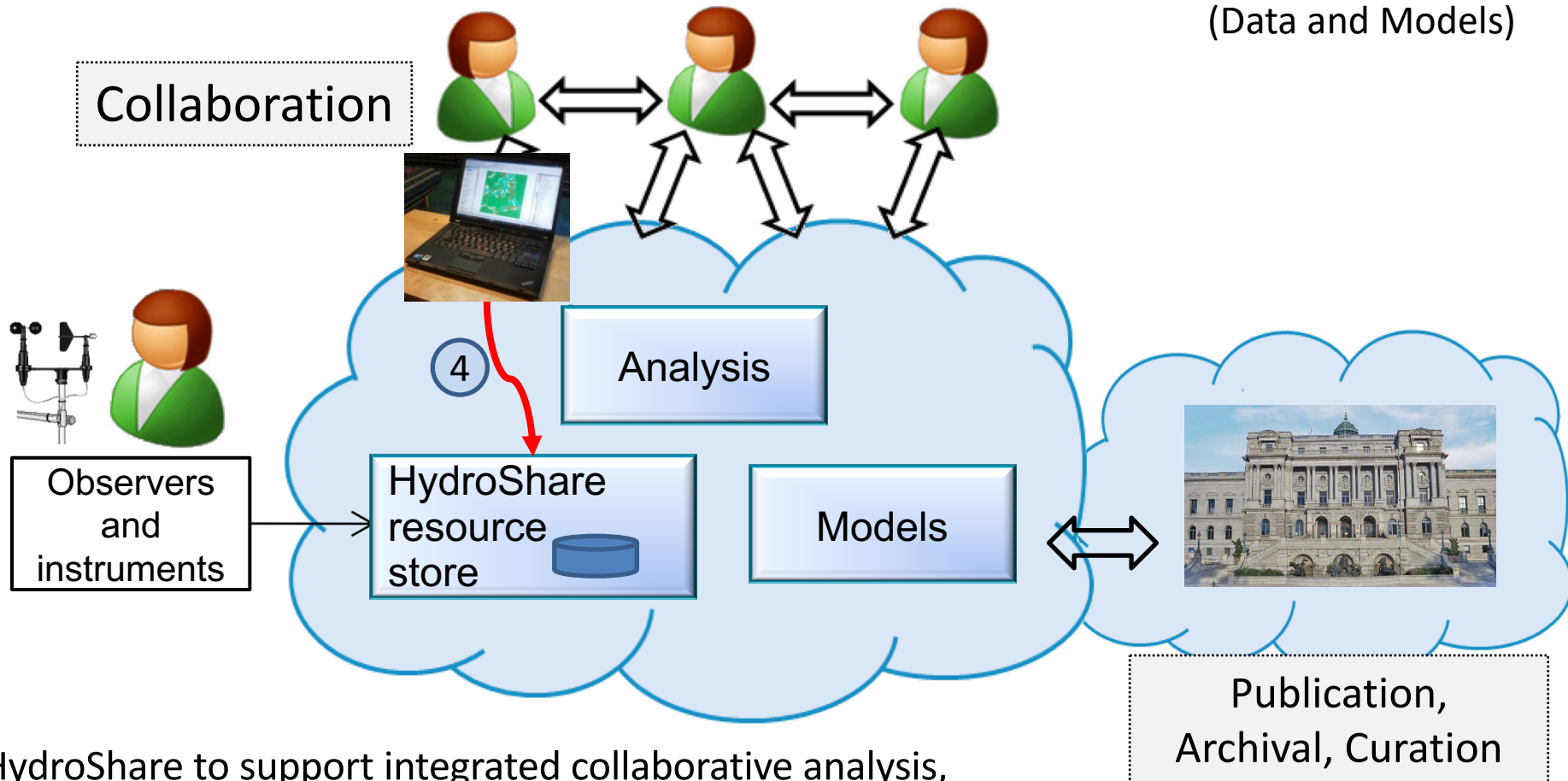
Imagine the Possibilities Enabled by Cloud-Based Systems



HydroShare to support integrated collaborative analysis, modeling and data publication

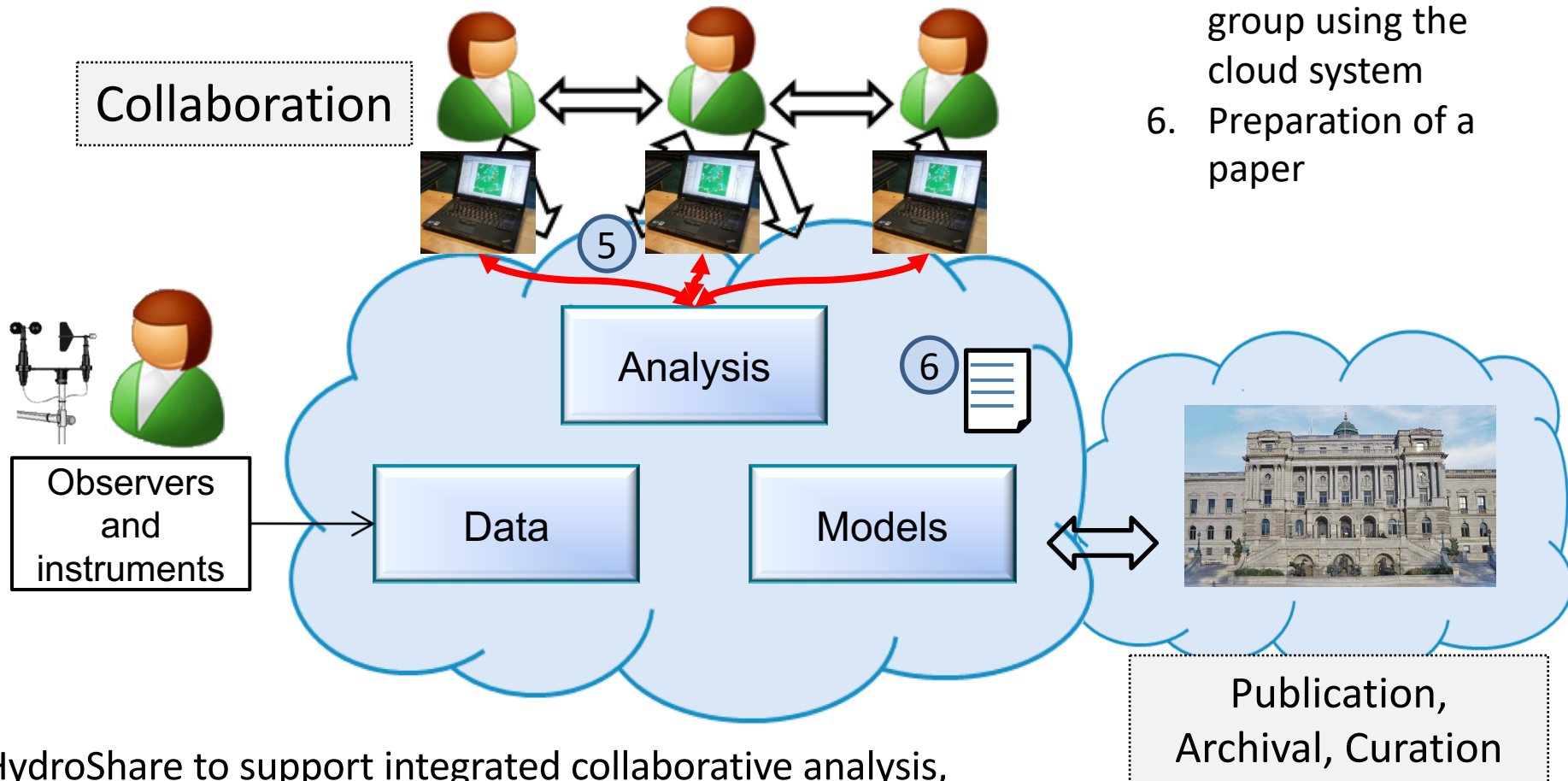
Imagine the Possibilities Enabled by Cloud-Based Systems

4. Share the results
(Data and Models)



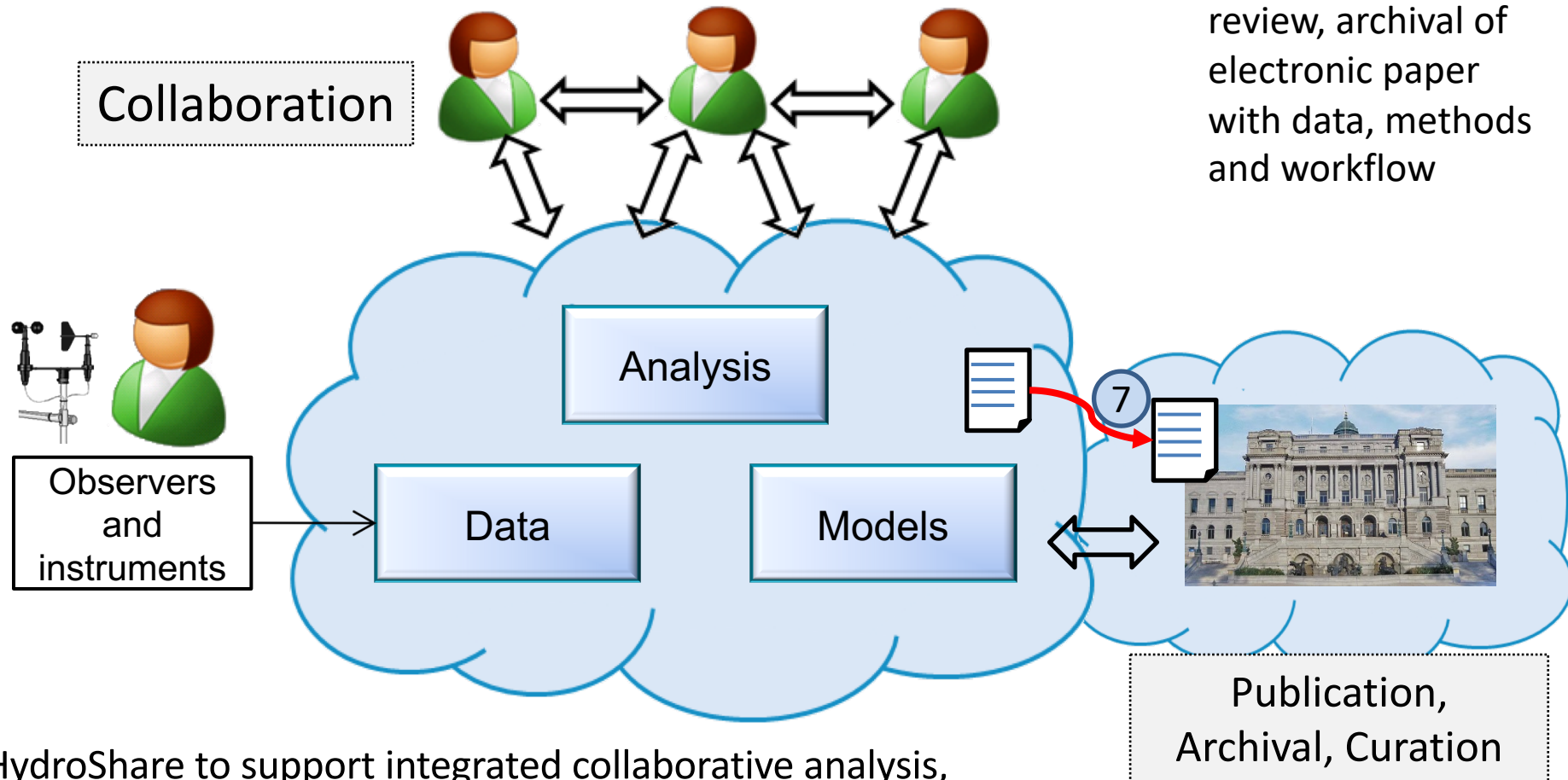
HydroShare to support integrated collaborative analysis,
modeling and data publication

Imagine the Possibilities Enabled by Cloud-Based Systems



HydroShare to support integrated collaborative analysis, modeling and data publication

Imagine the Possibilities Enabled by Cloud-Based Systems

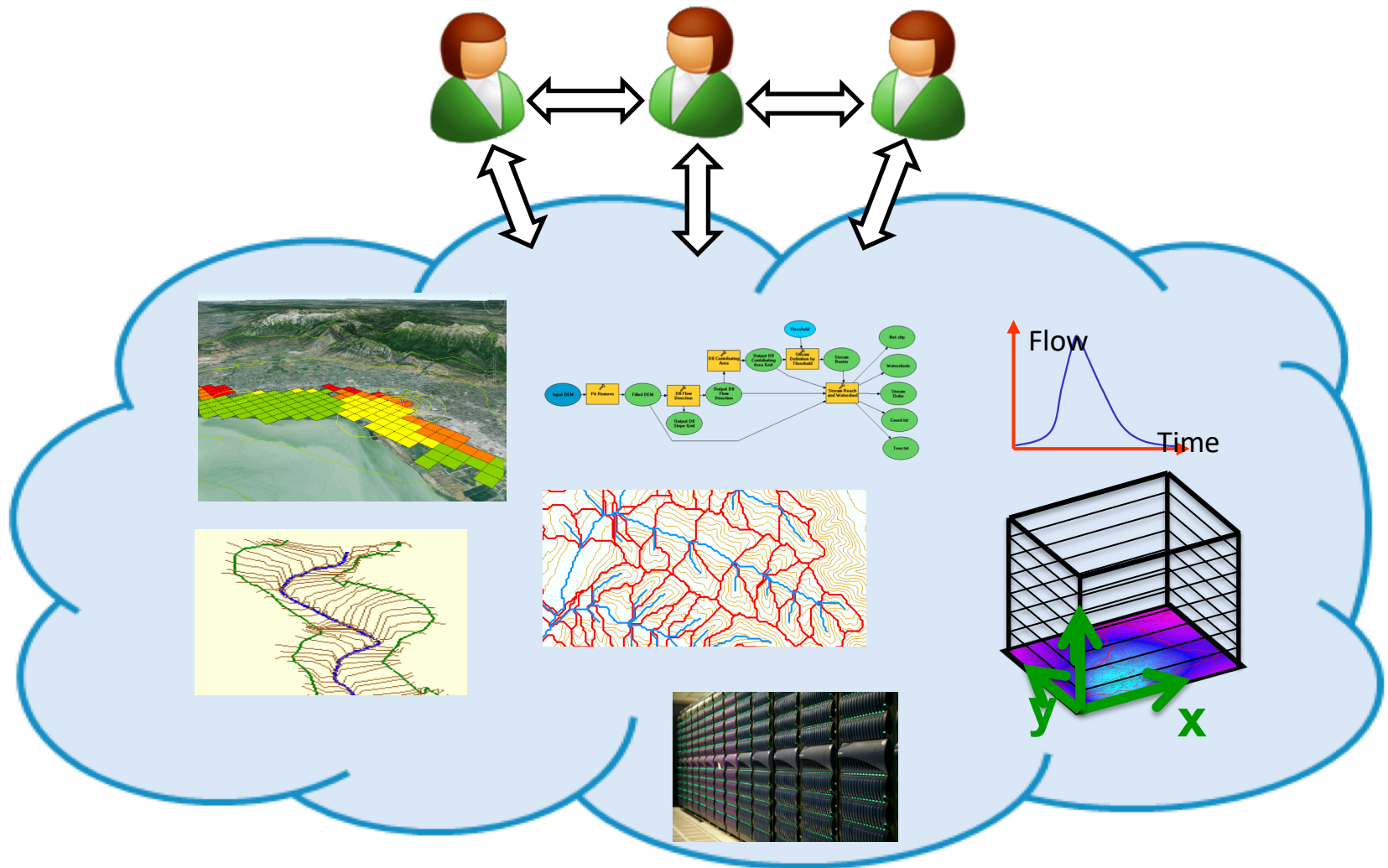


7. Submittal of paper, review, archival of electronic paper with data, methods and workflow

HydroShare to support integrated collaborative analysis, modeling and data publication

DataOne, EarthCube, ...

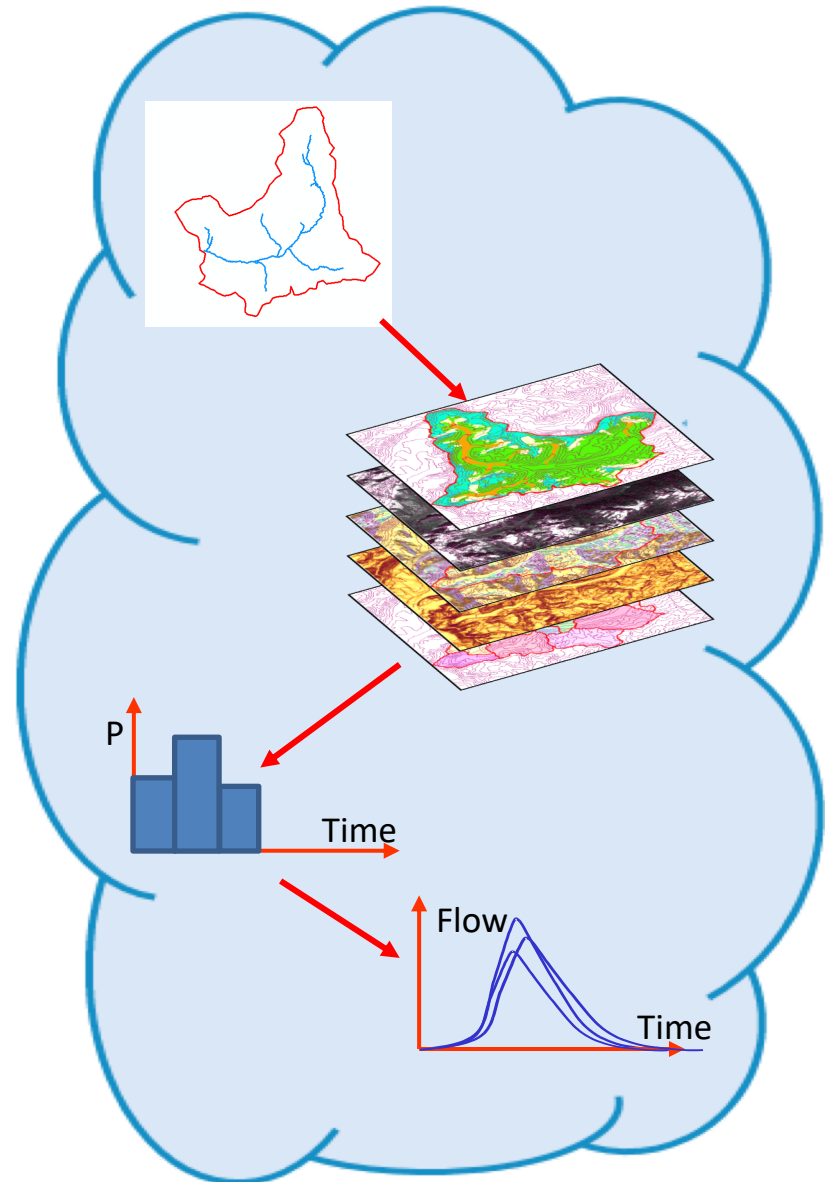
Imagine Cloud Modeling



- Automated reasoning to couple models based on purpose, context, data and resources

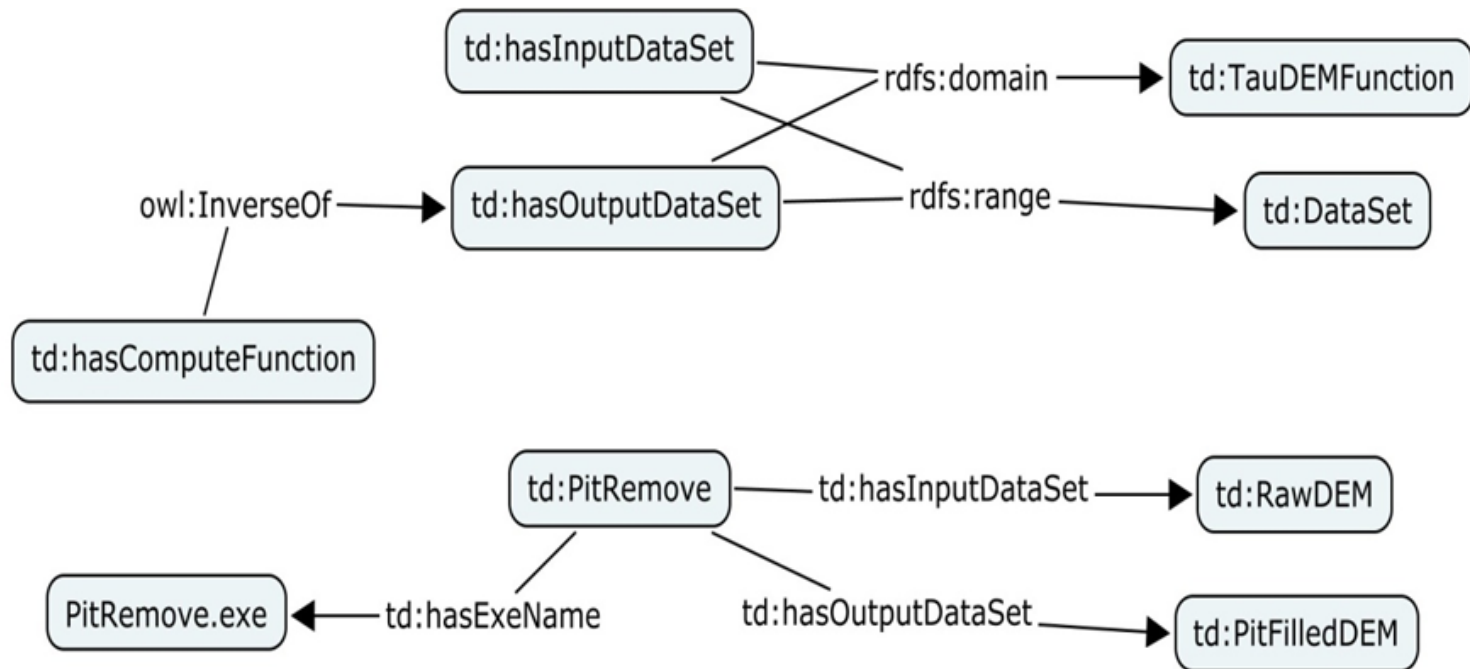
Another Example

- Big snow year
- Will my city flood?
- Click to delineate watershed (model domain)
- Generate model package from Essential Terrestrial Variables
- Generate suite of input scenarios
- Execute model and view results



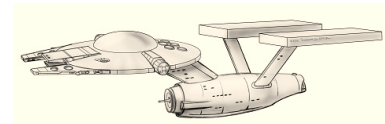
But there is more...

What if I could express my decision needs to the system and have it reason and deduce which models need to run, then configure and run them based on the inputs available, precision needs and resources and time available.



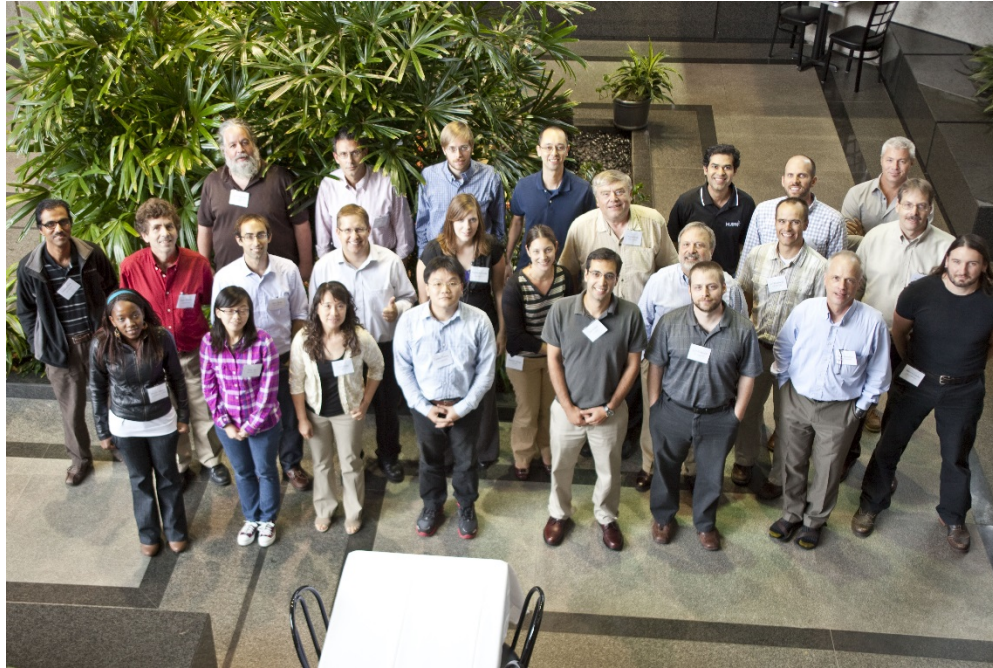
Summary and Conclusions

- Cloud Computing and Data Services
 - Community Participation
 - Interoperability
 - Standards
 - Open Development
- CUAHSI HIS
 - Enhanced Access to Hydrologic Data
 - Combining information from multiple sources
- HydroShare
 - A collaborative website for the sharing of hydrologic data and models
 - To expand data sharing capability of CUAHSI HIS
 - Additional data classes
 - Models, scripts, tools and workflows
 - To boldly go where no one has gone before



Thanks to a lot of people

- USU
- RENCI/UNC
- CUAHSI
- BYU
- Tufts
- USC
- Texas
- Purdue
- SDSC



HydroShare team, CUAHSI HIS team, CI-WATER team, Students and Colleagues (Ray Idaszak, Dan Ames, Jeff Horsburgh, Jon Goodall, Larry Band, Venkatesh Merwade, Carol Song, Alva Couch, David Valentine, Rick Hooper, Jennifer Arrigo, David Maidment, Tim Whiteaker, Alex Bedig, Jason Coposky, Pabitra Dash, Tian Gan, Karl Gustafson, Taehee Hwang, Stephen Jackson, Harry Johnson, Yuri Kim, Phyllis Mbewe, Brian Miles, Jon Pollak, Stephanie Reeder, Terrell Russell, Tom Whitenack, Ilya Zaslavsky, Lan Zhao, Michael Piasecki, Kim Schreuders, Jiri Kadlec, Scott Peckham, Chris Duffy, Lorne Leonard, Tseganeh Gichamo, Norm Jones, Fred Ogden, Aaron Byrd)



HYDROSHARE

<http://www.cuahsi.org/hydroshare.aspx>



EAR 0622374 (2007-2012)
OCI-1148453 (2012-2017)
OCI-1148090 (2012-2017)



CUAHSI Conference on Hydroinformatics and Modeling

July 17-19, 2013

Utah State University, Logan, Utah

NOW OPEN! Conference **Registration** and **Abstract** Submission

Abstracts Due: June 15, 2013

Registration Closes: June 30, 2013

For conference details, schedule, costs, presenter information, and logistics

<http://www.cuahsi.org/WDCconf2013>

Conference format includes

- Technical sessions on **hydrologic information systems and hydrologic modeling**
- Hands-on training and workshops to introduce community to services provided by newly established **Water Data Center (WDC)**
- Planned sessions on **using water data and information systems in the classroom**

Questions on conference or registration? Contact kberry@cuahsi.org