The CUAHSI Hydrologic Information System

Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI)

http://www.cuahsi.org
HIS Project Team

- University of Texas at Austin – David Maidment (PI), Tim Whiteaker
- San Diego Supercomputer Center – Ilya Zaslavsky, David Valentine, Tom Whitenack
- Utah State University – David Tarboton, Jeff Horsburgh
- CCNY – Michael Piasecki
- University of South Carolina – Jon Goodall, Tony Castronova
- Idaho State University – Dan Ames, Jiri Kadlec
- Tufts University – Alva Couch, Alex Bedig
Outline

• The HIS Story

• HIS components

• Putting the pieces together
Outline

- *The HIS Story*
- HIS components
- Putting the pieces together
HIS Connects People with Data

The CUAHSI Hydrologic Information System (HIS) provides web services, tools, standards and procedures that enhance access to more and better data for analysis.

his.cuahsi.org
What kind of data does HIS support?

- HIS is designed for *in situ* (e.g., sensor) data collected at a *fixed point*.
- Archetype: Stream flow at a gage on a river
- Sensor can measure any physical, chemical, or biological property
- Relaxing assumptions
  - Moving platform data can be represented (but not as efficiently as custom designed database)
  - Time series can be associated with polygon or arc in WaterML 2.0
  - “Simple” laboratory analyses (aquatic chemistry...
We Collect Lots of Water Data

Water quantity

Rainfall

Soil water

Water quality

Meteorology

Groundwater
The Data Have a Similar Structure

A **point** location in **space**

A **series** of values in **time**
Data Are Collected by Many Organizations
USA has GIS Data Repository

...but nothing equivalent for water data
Currently, the focus is on data from monitoring sites at point locations.
The Result

• WaterML language for describing water data

• Global (?) catalog of water data sources

• Free software for data access
Metadata Catalog, October, 2012

100 public services
32,000+ variables
2.79 million sites
33.9 million series
Referencing 18+ billion data values
HIS Usage

Number of Time Series Downloaded
For more on the HIS Story

his.cuahsi.org
Outline

- The HIS Story
- **HIS components**
- Putting the pieces together
Web Paradigm

Catalog (Google)

Web Server (CNN.com)

Browser (Firefox)

Catalog harvest

Search

Access
Services-Oriented Architecture for Water Data

HIS Central

HydroServer
- Service registration
- Catalog harvest

HydroDesktop
- Data access
- Search
Services-Oriented Architecture for Water Data
HydroDesktop

- Free, open source solution for HIS data access
- [www.hydrodesktop.org](http://www.hydrodesktop.org)
HD-Opening Screen
Select Area of Interest
Select Constituents of Interest
Examine Metadata

<table>
<thead>
<tr>
<th>DataSource</th>
<th>SiteName</th>
<th>VarName</th>
<th>SiteCode</th>
<th>VarCo</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPA</td>
<td>CAYUGA CREEK IN Cheektowaga</td>
<td>Inorganic nitrogen (nitrate and nitrite) as N</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:3</td>
</tr>
<tr>
<td>EPA</td>
<td>CAZENOVIA CREEK In Buffalo</td>
<td>Inorganic nitrogen (nitrate and nitrite) as N</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:3</td>
</tr>
<tr>
<td>EPA</td>
<td>BUFFALO R. IN BUFFALO</td>
<td>Inorganic nitrogen (nitrate and nitrite) as N</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:3</td>
</tr>
<tr>
<td>EPA</td>
<td>MUD CREEK IN RAYMOND</td>
<td>Inorganic nitrogen (nitrate and nitrite) as N</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:3</td>
</tr>
<tr>
<td>NWISIID</td>
<td>Ni1093</td>
<td>NH3+orgN, wf</td>
<td>NWISIID:431500</td>
<td>NWISI</td>
</tr>
<tr>
<td>NWISIID</td>
<td>E2642</td>
<td>NH3+orgN, wf</td>
<td>NWISIID:425749</td>
<td>NWISI</td>
</tr>
<tr>
<td>EPA</td>
<td>BUFFALO R. IN BUFFALO</td>
<td>Nitrite as N</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:1</td>
</tr>
<tr>
<td>NWISIID</td>
<td>Ni1093</td>
<td>Nitrite, wf</td>
<td>NWISIID:431500</td>
<td>NWISI</td>
</tr>
<tr>
<td>NWISIID</td>
<td>E2642</td>
<td>Nitrite, wf</td>
<td>NWISIID:425749</td>
<td>NWISI</td>
</tr>
<tr>
<td>NWISIID</td>
<td>E2642</td>
<td>Nitrite, wf</td>
<td>NWISIID:425749</td>
<td>NWISI</td>
</tr>
<tr>
<td>EPA</td>
<td>CAYUGA CREEK IN Cheektowaga</td>
<td>Nitrogen, Kjeldahl</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:3</td>
</tr>
<tr>
<td>EPA</td>
<td>MUD CREEK IN RAYMOND</td>
<td>Nitrogen, Kjeldahl</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:3</td>
</tr>
<tr>
<td>EPA</td>
<td>TONAWANDA CREEK IN Rapids</td>
<td>Nitrogen, Kjeldahl</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:3</td>
</tr>
<tr>
<td>EPA</td>
<td>CAZENOVIA CREEK In Buffalo</td>
<td>Nitrogen, Kjeldahl</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:3</td>
</tr>
<tr>
<td>EPA</td>
<td>BUFFALO CREEK In Gardenville</td>
<td>Nitrogen, Kjeldahl</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:3</td>
</tr>
<tr>
<td>EPA</td>
<td>BUFFALO RIVER in Buffalo</td>
<td>Nitrogen, Kjeldahl</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:3</td>
</tr>
<tr>
<td>EPA</td>
<td>RANSOM CREEK IN CLARENCE</td>
<td>Nitrogen, Kjeldahl</td>
<td>EPA:21NYDECA:....</td>
<td>EPA:3</td>
</tr>
<tr>
<td>NPCA</td>
<td>FR001</td>
<td>Nitrogen, nitrate (NO3)</td>
<td>NPCA:FR001</td>
<td>NPCA:</td>
</tr>
<tr>
<td>NPCA</td>
<td>WR010</td>
<td>Nitrogen, nitrate (NO3)</td>
<td>NPCA:WR010</td>
<td>NPCA:</td>
</tr>
<tr>
<td>NPCA</td>
<td>WW001</td>
<td>Nitrogen, nitrate (NO3)</td>
<td>NPCA:WW001</td>
<td>NPCA:</td>
</tr>
</tbody>
</table>
Select Series for Download
View Data
GIS fully integrated with HIS

- Metadata catalog
- Ontology keywords
- WaterOneFlow/WaterML

Discovery
Access
Analysis
Built-in Analysis

- Tables
- Graphs
- Editing
- Export
Prototype Web-based Client

Available Data
- Sample Medium (5 options)
- Organization (4 options)
- Variable Name (1287 options)
- Data Type (3 options)
- Value Type (4 options)
- Network (7 options)

Search

Showing sites of data where Data Type equals Instantaneous, or Sporadic, or Continuous on the map.

https://data.cuahsi.org
HIS System Overview

Users (HydroDesktop)

USGS

University

Data Discovery Data Access

Web Service

Data Registration

HIS Central

Metadata Catalog

Hydrologic Ontology

Data Registration

Data Access

Data Discovery

HydroServers
Getting Water Data (the old way)

Different Query Pages  Different Query Responses
WaterML includes location, variables, and time series

<timeSeriesResponse xmlns:gm1="http://www.opengis.net/gml" xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xn:wtr="http://www.cuahsi.org/waterML/
xn:ns="http://www.cuahsi.org/waterML/1.0/">
  <queryInfo>...
  <timeSeries>
    <sourceInfo xsi:type="SiteInfoType">
      <siteName>Colorado Rv at Austin, TX</siteName>
      <siteCode siteID="1389">08158000</siteCode>
      <timeZoneInfo>...
      <geoLocation>...
      <note>Agency: USGS</note>
    </sourceInfo>
    <variable>
      <variableCode vocabulary="USGS">00060</variableCode>
      <variableName>Discharge</variableName>
      <dataType>Average</dataType>
      <units>cfs</units>
      <options>...</option>
    </variable>
    <values>
      <value qualifiers="A" dateTime="2007-01-01T00:00:00">143</value>
      <value qualifiers="A" dateTime="2007-01-02T00:00:00">231</value>
      <value qualifiers="A" dateTime="2007-01-03T00:00:00">112</value>
      <value qualifiers="A" dateTime="2007-01-04T00:00:00">398</value>
      <value qualifiers="A" dateTime="2007-01-05T00:00:00">152</value>
      <value qualifiers="A" dateTime="2007-01-06T00:00:00">212</value>
      <value qualifiers="A" dateTime="2007-01-07T00:00:00">120</value>
      <value qualifiers="A" dateTime="2007-01-08T00:00:00">151</value>
      <value qualifiers="A" dateTime="2007-01-09T00:00:00">132</value>
      <value qualifiers="A" dateTime="2007-01-10T00:00:00">133</value>
      <value qualifiers="A" network="USGS" vocabulary="dv_rnk_cd">Approved for publication -- Processing and review completed.</value>
    </values>
  </timeSeries>
</timeSeriesResponse>
Web Pages and Web Services

http://www.safl.umn.edu/

http://his.safl.umn.edu/SAFLMC/cuahsi_1_0.asmx

Uses Hypertext Markup Language (HTML)

Uses WaterML

(a Markup Language for water data)
WaterOneFlow Web Service

• Set of **query** functions
  – Get Sites
  – Get Site Info
  – Get Variable Info
  – Get Values

• Returns data in **WaterML**
WaterML and WaterOneFlow

WaterOneFlow is how you ask for data
WaterML is the format of what comes back
HydroServer Goals

• A platform for publishing space-time hydrologic datasets that:
  – Provides local control of data
  – Makes data universally available
  – Is open source (hydroserver.codeplex.com)
GIS Data

Point Observations Data

Ongoing Data Collection

Historical Data Files

HydroServer

GIS Data

GetSites

GetSiteInfo

GetVariableInfo

GetValues

WaterOneFlow

Web Service

WaterML

ArcGIS Server

ODM Database

Data presentation, visualization, and analysis through Internet enabled applications
Operational HydroServers

http://icewater.usu.edu/

http://www.his.npca.ca/hydroserver/
HydroServer Alternatives - KISTERS

External Client Application
HydroGET, ....

KISTERS Telemetry

CUAHSI

CUAHSI

SOS

APIs

SOS

OpenMI

APIs

WISKI / Hydstra

internal

CUAHSI

Import/Export Framework

http
(s)ftp

DB

Files

Vendor driver

XHydro
Why Publish Data with HIS

- Recognition
- Collaboration
- Public service
- Cost savings
HIS Central

- Publishers
  - Register a data service
- Users
  - Find a data service
- Supported by
  - Metadata Catalog
  - Hydrologic Ontology

http://hiscentral.cuahsi.org
Metadata Catalog

Stores description of time series, e.g.,

*The USGS measures streamflow at Waller Creek & Koenig with data from 7/31/1968 to the present...*

...and you can get the data from [here](#).

Registered services harvested weekly
100 public services
32,000+ variables
2.79 million sites
33.9 million series
Referencing 18+ billion data values
Ontology: Conceptual Framework

• **Chemical** descriptions from **EPA/USGS Substance Registry System**
  
  [http://www.epa.gov/srs/](http://www.epa.gov/srs/)

• **Physical** descriptions from **CF Conventions**
  – NetCDF Climate & Forecast; 137 variables


• **Biological** descriptions from **Integrated Taxonomic Information System**

Thematic Concepts

Core Concept
- Hydrosphere
- Biological
- Chemical
- Physical

Property
- Biological
- Chemical
- Physical

Branch
- Chlorophyll
- Indicator Organisms
- Dissolved Oxygen
- Suspended Solids
- Discharge
- Temperature

Leaf
- Chlorophyll A
- Pheophytin
- E. Coli
- Fecal Coliform
- DO Concentration
- DO Saturation
- Fixed Suspended Solids
- Total Suspended Solids
- Groundwater Flow
- Stream Discharge
- Air Temperature
- Water Temperature

Variables
Each **Variable** in your data is connected to a corresponding **Concept**
All Registered Data Services

<table>
<thead>
<tr>
<th>Data Service Title</th>
<th>Observation Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baltimore Precipitation</td>
<td>BaltPrecip</td>
</tr>
<tr>
<td>Baltimore Ecosystem Study Stream Chemistry Data</td>
<td>BESOD</td>
</tr>
<tr>
<td>Baltimore Ecosystem Study Soils Data</td>
<td>BESSoil</td>
</tr>
<tr>
<td>Baltimore Waters Test Bed Ground Water Level Data</td>
<td>BaltimoreGW</td>
</tr>
<tr>
<td>Beacon Institute for River and Estuary</td>
<td>BEACON_IBM</td>
</tr>
<tr>
<td>Dry Creek Experimental Watershed, SW Idaho</td>
<td>ODMDCE</td>
</tr>
<tr>
<td>Chesapeake Bay Information Management System</td>
<td>CIMS</td>
</tr>
</tbody>
</table>

Dry Creek Experimental Watershed, SW Idaho

Boise State University, Hydrologic Sciences Department

ODMDCE-W2
http://icewater.boisestate.edu/bicsw2/datasets/cuahsi_1_0.asmx?WSDL

Contact:
Pam Aishlin
pamaishlin@boisestate.edu
208-426-2220

Service Statistics:
- Sites: 69
- Variables: 24
- Values: 4730590
- Geographic Extent: 43.74071, -116.1766 to 43.86804, -110.098

Last Harvested on 7/25/2010 1:12:50 PM

Abstract

Dry Creek Experimental Watershed was established by Dr. Jim McNamara in 1998 as an outdoor laboratory for student and faculty research toward improving understanding of hydrologic processes in semi-and mountainous terrain and testing and improving data integration and hydrologic modeling. Continuous and discrete data collection includes climate, surface water, groundwater and soil

http://hiscentral.cuahsi.org
HIS Central Web Service

• Programmatic methods to query the national metadata catalog

• Search by:
  – Location
  – Variable (concept)
  – Date Range
  – Data source (WaterOneFlow service)

http://hiscentral.cuahsi.org/webservices/hiscentral.asmx
Services-Oriented Architecture for Water Data

Catalog

HIS Central

WaterML

Service registration
Catalog harvest
Search

Data access

HydroServer
Data Publisher

HydroDesktop
User
The Road Ahead – WaterML 2.0

• CUAHSI Water Data Center
  – Recommended for funding at NSF
  – $800K/yr for 3 yr start-up

• Broader collaboration
  – Hydrology Domain Working Group
    • World Meteorological Organization
    • Open Geospatial Consortium (OGC)

• Towards an international standard
  – WMO initiate process to accept as standard Nov., 2012
Currently, the focus is on data from monitoring sites at point locations.
Multiple Services, Multiple Clients within a Standards-Oriented Environment

Standards-based interface to serve multiple clients

Structural and Semantic mediation within community and data type
Start Using HIS!

• HIS Website
  – his.cuahsi.org

• Hydrodesktop
  – www.hydrodesktop.org

• User Support available through CUAHSI

• Contribute to software development