The CUAHSI Community Hydrologic Information System

David Tarboton, David Maidment, Ilya Zaslavsky, Dan Ames, Jon Goodall, Richard Hooper, Jeffrey Horsburgh

http://his.cuahsi.org/
Hydrologic Data Challenges

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

Data Heterogeneity

The way that data is organized can enhance or inhibit the analysis that can be done
The CUAHSI Hydrologic Information System (HIS) is an internet based system to support the sharing of hydrologic data. It is comprised of hydrologic databases and servers connected through web services as well as software for data publication, discovery and access.
CUAHSI Hydrologic Information System

Services-Oriented Architecture

- **HydroCatalog**: Data Discovery and Integration
- Metadata Services
- Search Services
- WaterML, Other OGC Standards
- **Data Services**
- **HydroServer**: Data Publication
- ODM, Geo Data
- **HydroDesktop**: Data Analysis and Synthesis

Information Model and Community Support Infrastructure
What are the basic attributes to be associated with each single data value and how can these best be organized?

<table>
<thead>
<tr>
<th>Basic Attributes</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date/Time</td>
<td></td>
</tr>
<tr>
<td>Interval (support)</td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td></td>
</tr>
<tr>
<td>Quality Control Level</td>
<td></td>
</tr>
<tr>
<td>Sample Medium</td>
<td></td>
</tr>
<tr>
<td>Value Type</td>
<td></td>
</tr>
<tr>
<td>Data Type</td>
<td></td>
</tr>
<tr>
<td>Source/Organization</td>
<td></td>
</tr>
<tr>
<td>Units</td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
</tr>
<tr>
<td>Censoring</td>
<td></td>
</tr>
<tr>
<td>Qualifying comments</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>Feature of interest</td>
<td></td>
</tr>
</tbody>
</table>
Observations Data Model (ODM)

- A relational database at the single observation level
- Metadata for unambiguous interpretation
- Traceable heritage from raw measurements to usable information
- Promote syntactic and semantic consistency
- Cross dimension retrieval and analysis

WaterML and WaterOneFlow

**WaterML** is an XML language for communicating water data. **WaterOneFlow** is a set of web services based on WaterML.

- Set of *query* functions
  - GetSites
  - GetSiteInfo
  - GetVariableInfo
  - GetValues

- Returns data in **WaterML**

```
<timeSeries>
  - <sourceInfo xsi:type="SiteInfoType">
    <siteName>Colorado River at Austin, TX</siteName>
    <siteCode>4619631</siteCode>
    - <geoLocation>
      - <geogLocation xsi:type="LatLonPointType" srs="EPSG:4326">
        <latitude>30.4445429</latitude>
        <longitude>-97.694448</longitude>
      </geogLocation>
    </geoLocation>
  </sourceInfo>
  - <variable>
    <variableCode>NWIS</variableCode>
    <variableName>Discharge, cubic feet per second</variableName>
    <units>feet³/sec</units>
    <unitsAbbreviation>cfs</unitsAbbreviation>
  </variable>
  - <values count="2545">
    <value dateTime="2006-12-31T00:00:00">129</value>
    <value dateTime="2006-12-31T00:15:00">129</value>
    <value dateTime="2006-12-31T00:30:00">129</value>
    <value dateTime="2006-12-31T00:45:00">129</value>
    <value dateTime="2006-12-31T1:00:00">124</value>
    <value dateTime="2006-12-31T1:15:00">129</value>
    <value dateTime="2006-12-31T1:30:00">124</value>
    <value dateTime="2006-12-31T1:45:00">124</value>
  </values>
</timeSeries>
```
HydroServer – Data Publication

Ongoing Data Collection

Point Observations Data

Historical Data Files

GIS Data

ODM Database

Internet Applications

GetSites
GetSiteInfo
GetVariableInfo
GetValues

WaterML

WaterOneFlow
Web Service

OGC Spatial Data Service from ArcGIS Server

Data presentation, visualization, and analysis through Internet enabled applications
HydroCatalog

• Search over data services from multiple sources
• Supports concept based data discovery

Service Registry

Hydrotagger

Harvester

Water Metadata Catalog

Search Services

Discovery and Access

CUAHSI Data Server

3rd Party Server e.g. USGS

WaterML

GetSites
GetSiteInfo
GetVariableInfo
GetValues

WaterOneFlow
Web Service

http://hiscentral.cuahsi.org
Integration from multiple sources

Thematic keyword search

Search on space and time domain

HydroDesktop – Data Access and Analysis
Integration with “R” Statistics Package
Open Geospatial Consortium
Web Service Standards

• Map Services
  • Web Map Service (WMS)
  • Web Feature Service (WFS)
  • Web Coverage Service (WCS)
  • Catalog Services for the Web (CS/W)

• Observation Services
  • Observations and Measurements Model
  • Sensor Web Enablement (SWE)
  • Sensor Observation Service (SOS)

These standards have been developed over the past 10 years .... .... by 400 companies and agencies working within the OGC

OGC Hydrology Domain Working Group evolving WaterML into an International Standard
http://www.opengeospatial.org/projects/groups/waterml2.0swg
A growing collection of HydroServers and community of users

- University of Maryland, Baltimore County
- Montana State University
- University of Texas at Austin
- University of Iowa
- Utah State University
- University of Florida
- University of New Mexico
- University of Idaho
- Boise State University
- University of Texas at Arlington
- University of California, San Diego
- Idaho State University

Dry Creek Experimental Watershed (DCEW) (28 km\(^2\) semi-arid steep topography, Boise Front)

68 Sites
24 Variables
4,700,000+ values

Published by Jim McNamara, Boise State University
Open Development Model

- http://hydrodesktop.codeplex.com
- http://hydroserver.codeplex.com
- http://hydrocatalog.codeplex.com
General aspects of the approach

- **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

![Diagram of common functional components]

- **Catalog**: Discovery and Integration
- **Server**: Publication
- **Desktop/Client**: Synthesis and Research
- **Standards**: Metadata Services
- **Search Services**: Data Services
- **Accessibility**: Information Model and Community Support Infrastructure
- **Consumability**
Looking to the Future

- Move from prototype to operations
  - Operational support of software and systems
  - User support and training
  - Repositories
  - CUAHSI Data Center (User Support Specialist)
  - NSF Data Management Requirements

- Research and development of new functionality
  - data and model sharing “hub” to enhance interactive collaboration (pending)

- Community
  - HIS has become bigger than one project (emerging software ecosystem)
  - Open Development Model (inspire, enable and incorporate broad contributions)
  - The community is the infrastructure that persists (is sustainable)
Thanks! HIS Project Team and Sponsors

- **University of Texas at Austin** – David Maidment, Tim Whiteaker, James Seppi, Fernando Salas, Jingqi Dong, Harish Sangireddy
- **San Diego Supercomputer Center** – Ilya Zaslavsky, David Valentine, Tom Whitenack, Matt Rodriguez
- **Utah State University** – Jeff Horsburgh, Kim Schreuders, Stephanie Reeder, Edward Wai Tsui, Ravichand Vegiraju, Ketan Patil
- **University of South Carolina** – Jon Goodall, Anthony Castronova
- **Idaho State University** – Dan Ames, Ted Dunsford, Jiří Kadlec, Yang Cao, Dinesh Grover
- **Drexel University/CUNY** – Michael Piasecki
- **WATERS Network** – Testbed Data Managers
- **CUAHSI Program Office** – Rick Hooper, Yoori Choi, Conrad Matiuk
- **ESRI** – Dean Djokic, Zichuan Ye

CUAHSI

[http://his.cuahsi.org/](http://his.cuahsi.org/)

Support EAR 0622374