Sharing Data Using the CUAHSI Hydrologic Information System
Synopsis of Class 18, GIS in Water Resources, Fall 2011

The class has to date focused mostly on spatial geographic data in the context of ArcGIS. Today we expand our view to other types of hydrologic data and the development of capability for publishing, managing and sharing it using Hydrologic Information Systems, and specifically (HIS) the CUAHSI HIS that has been the focus of significant recent research.

Learning objectives

- Describe the components of a Hydrologic Information System and some of the key functionality of each component
- Reflect on the basic attributes needed to quantify data in an unambiguous way and how to organize this information to enhance analysis
- Discuss the placement of information in a relational data model such as ODM and the relationships required to associate attributes with data values
- Discuss the functionality and models for development of HIS components

The CUAHSI HIS

The CUAHSI HIS is an internet based system to support the sharing of hydrologic data (see http://his.cuahsi.org for details). It is comprised of hydrologic databases and servers connected through web services as well as software for data publication, discovery and access. The system that has been developed provides new opportunities for the water research community to approach the management, publication, and analysis of their data systematically. The system's flexibility in storing and enabling public access to similarly formatted data and metadata has created a community data resource from public and academic data that might otherwise have been confined to the private files of agencies or individual investigators. HIS provides an analysis environment for the integration of data from multiple sources and serves as a prototype for the infrastructure to support a network of large scale environmental observatories or research watersheds. As presently implemented HIS focuses on a relatively narrow class of data, namely in situ point time series of observations with some capability for including ex-situ samples and simple geospatial information. Nevertheless many of the concepts developed have broader application than presently implemented.

Two concepts, (1) the services oriented architecture; and (2) the desktop hydrologic information system underlie the architecture of the CUAHSI HIS (Figure 1).

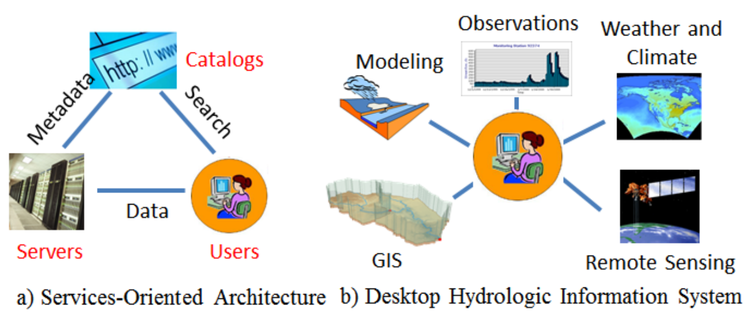

Figure 1. Hydrologic Information System Overarching Vision

The HIS services-oriented architecture can be viewed as: 1) a way of publishing hydrologic data in a uniform way; 2) a way of discovering and accessing remote water information archives in a uniform way; and 3) a way of displaying, synthesizing and analyzing water information and exporting it to other analysis and modeling systems. The connections among components are established by web services.

The concept of HIS desktop application software is somewhat analogous to Geographic Information System (GIS) desktop software that supports storage and analysis of logically linked data (Tomlinson, 2003). Our implementation, "HydroDesktop" provides an analysis environment within which data from multiple sources can be discovered, accessed and integrated. HydroDesktop is available from http://his.cuahsi.org/hydrodesktop.html and will be demonstrated during this class.

The HIS services-oriented architecture is comprised of three classes of functionality: 1) data publication (HydroServer), 2) data cataloging (HydroCatalog), and 3) data discovery, access and analysis (HydroDesktop) (Figure 2). This functionality follows the general paradigm of the Internet. HydroServer publishes data similar to the way Internet web servers publish content. HydroDesktop consumes data published from HydroServer, similar to the way web browsers consume Internet content. HydroCatalog supports data discovery based on indexed metadata similar to the way search engines support the discovery of Internet content. Syntactic (file types and formats) and semantic consistency has been a focus of HIS with an ontology and community controlled vocabulary used to harmonize the terminology used and support thematic key word based data discovery.

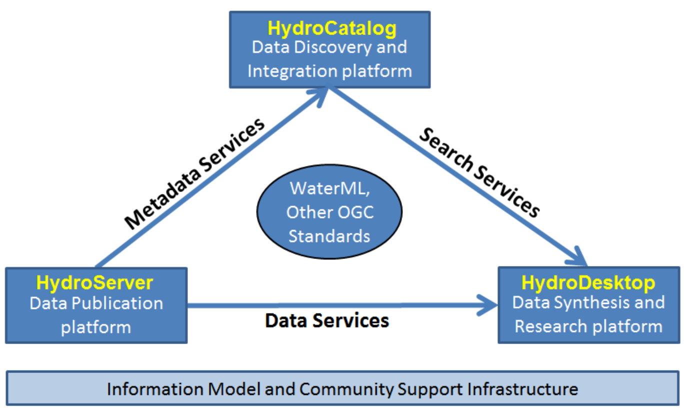

Figure 2. Components of CUAHSI HIS Services Oriented Architecture.

The components shown in Figure 2 either publish or consume information via the following categories of web services:

- Data Services – which convey the actual data.
- Metadata Services – which convey metadata about specific collections or series of data.
- Search Services – which enable search, discovery, and selection of data and convey metadata required for accessing data using data services.

The formats for transmission of information between these systems and the interfaces that enable the communication between them (the connecting arrows in Fig. 2) are critical to the functioning of the system. CUAHSI HIS has developed WaterML, an XML based language for transmitting observation data via web services (Zaslavsky et al., 2007). The web services are referred to as WaterOneFlow web services. CUAHSI HIS also relies on other established standards such as World Wide Web Consortium Simple Object Access Protocol (SOAP) and Open Geospatial Consortium (OGC) Geographic Markup Language (GML) for transmission of information between the three primary components.

At the base of Figure 2 is depicted the information model and community support infrastructure upon which the system is founded. The information model is the conceptual model used to organize and define sufficient metadata about hydrologic observations for them to be unambiguously interpreted and used. Within HydroServer, it is encoded using the Observations Data Model (ODM) (Horsburgh et al., 2008) relational database and the HydroServer Capabilities Database to ensure that data and metadata are stored together. The information model also serves as the conceptual basis for WaterML to ensure that data and associated metadata are transmitted with fidelity when data are downloaded. HydroDesktop implements the information model within its data repository database ensuring that local copies of data retrieved from a server maintain their original context. ODM includes a number of controlled vocabularies for metadata such as units, variable names, sample media etc., where semantic consistency in describing observations is important.

The architecture shown in Figure 2 has evolved as an approach for sharing hydrologic observations data that is general and open to allow broad participation. By relying on standards to define the interactions between components, different organizations can develop components and have them be interoperable. This approach based on standards provides a general foundation and approach for integration and sharing of hydrologic data around the world.

References and additional readings

Horsburgh, J. S., D. G. Tarboton, D. R. Maidment and I. Zaslavsky, (2008), "A Relational Model for Environmental and Water Resources Data," Water Resour. Res., 44: W05406, http://dx.doi.org/10.1029/2007WR006392.

Tarboton, D., D. Maidment, I. Zaslavsky, J. Goodall, D. Ames, R. P. Hooper, J. Horsburgh, K. Schreuders, R. Idaszak and A. Couch, (2011), "Advancing Solutions for an EarthCube Design. What can be learned from the CUAHSI HIS experience? An EarthCube Design Approaches Paper," EarthCube Discussion Forum, http://earthcube.ning.com/, http://api.ning.com/files/eqFq1lLUgULHk74T*w4KjjCoN4IQ3FgDjD*EjSnz6n4satxJX-mGYdvR7GuD6tKTrMORo6kRqN1LIxcd9db*N4jNjyHOr9ZH/EarthCubeDesign_CUAHSIHIS.pdf.

Tarboton, D. G., D. Maidment, I. Zaslavsky, D. Ames, J. Goodall, R. P. Hooper, J. Horsburgh, D. Valentine, T. Whiteaker and K. Schreuders, (2011), "Data Interoperability in the Hydrologic Sciences, The CUAHSI Hydrologic Information System," Proceedings of the Environmental Information Management Conference 2011, 132-137, http://eim.ecoinformatics.org/eim2011/eim-proceedings-2011/view.

Tarboton, D. G., D. Maidment, I. Zaslavsky, D. P. Ames, J. Goodall and J. S. Horsburgh, (2010), "CUAHSI Hydrologic Information System 2010 Status Report," http://his.cuahsi.org/documents/CUAHSIHIS2010StatusReport.pdf.

Tarboton, D. G., J. S. Horsburgh, D. R. Maidment, T. Whiteaker, I. Zaslavsky, M. Piasecki, J. Goodall, D. Valentine and T. Whitenack, (2009), "Development of a Community Hydrologic Information System," 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation, ed. R. S. Anderssen, R. D. Braddock and L. T. H. Newham, Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and

Computers in Simulation, July 2009, 988-994, http://www.mssanz.org.au/modsim09/C4/tarboton_C4.pdf.
Tomlinson, R., (2003), Thinking about GIS, ESRI Press, Redlands CA, 283 p.
Zaslavsky, I., D. Valentine and T. Whiteaker, (2007), "CUAHSI WaterML," OGC 07-041r1, Open Geospatial Consortium Discussion Paper, http://portal.opengeospatial.org/files/?artifact_id=21743.