Water Data Sharing – an Update

By David R. Maidment
Leader, CUAHSI Hydrologic Information System Project
and Director, Center for Research in Water Resources, University of Texas at Austin

Prepared for Subcommittee for Water Availability and Quality,
Office of Science and Technology Policy, Executive Office of the President

12 February 2011

Introduction

On January 19, 2011, I made a presentation to the Subcommittee on Water Availability and Quality (SWAQ) on "Water Data Sharing in the United States", which can be seen at: http://www.caee.utexas.edu/prof/maidment/meetings/SWAQJan19/WaterDataSharing.pptx and I prepared an introductory statement (Maidment, 2011) for the SWAQ introducing CUAHSI water data services in the context of the SWAQ's mission under the Secure Water Act (Section 9509 of Public Law 111-11) "to establish data management and communication protocols and standards to increase the quality and efficiency by which each agency acquires and reports relevant data; to consider options for the establishment of a data portal to enhance access to water resources data … that is collected by each Federal agency and any other private or public entity for each nationally significant freshwater watershed and aquifer located in the United States".

During the discussion of this mission at the January 19 SWAQ meeting, I suggested that there are two parts to this subject, a "why" part and a "how" part, and that it is appropriate for the SWAQ to focus on the "why" part of this mission and allow other institutions of the government to address the "how" part, such as the Advisory Committee for Water Information (ACWI) and the Federal Geographic Data Committee (FGDC). In particular, I suggested that the ACWI/FGDC Subcommittee for Spatial Water Data as one group that is interested in implementing water observations data as a class of national water geospatial information along with the National Hydrography Dataset and Watershed Boundary Dataset that they have previous fostered.

Concept Development Study

I also mentioned an NSF-supported Water Information Services Concept Development Study that the Consortium of Universities for the Advancement of Hydrologic Science, Inc (CUAHSI) initiated in January 2011 with the Open Geospatial Consortium (OGC), whose purpose is to understand how to translate the national catalog of about 70 water observations data services that CUAHSI has developed with its partners such as the USGS into a structure that conforms to the OGC information reference model. This translation will make water information conformal with other forms of geospatial and geoscience information that is already conformal with this reference model.

The Concept Development Study is being prepared using a framework illustrated in Figure 1. This begins with a "Community Objectives" section that describes the "What for, Why, Who, When" of the effort. That is followed by an Abstract/Best Practices section that describes "what is it about", a model for the information elements, and presents "how does each bit work" to describe how web services produce requests and receive responses for the information. Finally, there is an Implementation/Development section that summarizes "how do the components work together" in a connected information architecture, and a prescription for the various technology combinations that can be selected to actually implement the methodology.

Figure 1. Viewpoints of the Open Geospatial Consortium Reference Model (OGC, 2008)

The key part of the "Enterprise Viewpoint" which outlines the Community Objectives is a set of "Use Cases" that describe who the actors are in this information community and what functions they perform. In this instance we see three actors – data *Providers* such as water agencies or scientific research groups who publish water data and metadata describing their information, data *Catalogs* such as a national water data portal or data.gov, which index and provide search capability for the information, and data *Consumers* who search metadata in the catalogs to identify the information they need, and then directly access that data from the water data service Providers. This triangular arrangement of Providers, Catalogs and Consumers is illustrated in Figure 2, and it follows the classic pattern of internet use where users first search a catalog like Google or Bing to find summaries and links, then go to the source to get the information.

Use Cases

In the Concept Development Study, we have so far identified four use cases:

1. **Publication** – a data Provider publishes data and their associated metadata using web services.
2. **Cataloging** – the data Provider registers its web services with a centralized catalog, which harvests the metadata to support comprehensive search.
3. **Discovery** – a data Consumer searches a catalog to identify data series or datasets of interest.
4. **Access** – the Consumer acquires the selected data from the data Provider.

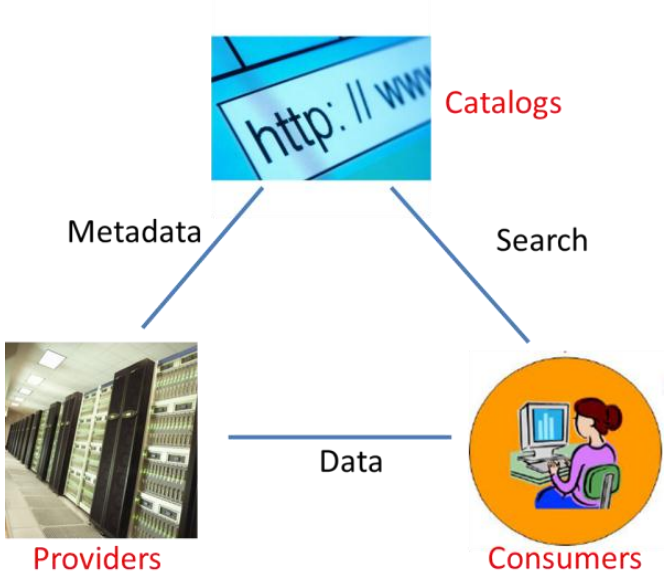

Figure 2. The interaction of water data providers, catalogs and consumers

Hydrologic Information System

This web-integrated Hydrologic Information System supports any number of data providers, catalogs and consumers. Each data Provider publishes its own data, metadata and catalog connection. The catalog connection can be registered in any number of portals, thus proving access to the information services to specialized outlets like NIDIS as well as generalized ones like data.gov. Researchers associated with the CUAHSI Hydrologic Information Systems team have built a prototype of the architecture envisaged, with a national water data catalog at the San Diego Supercomputer Center called HIS Central that indexes information from the USGS, EPA, NCDC, USDA, USACE and regional water sources, including NSF-supported observatories. In Texas, the Texas Water Development Board has assumed the responsibility of being the central data node for state water agency data, and the web application shown in Figure 3 searches across both federal and state catalogs to get dissolved oxygen observations for Texas and the surrounding area.

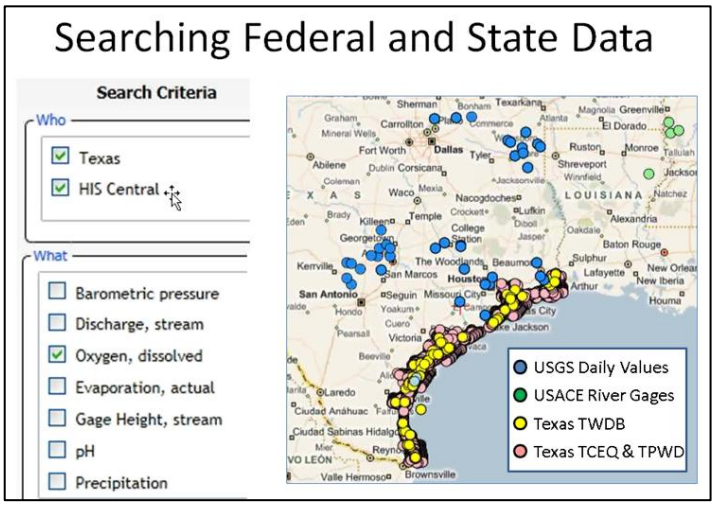

Figure 3. A web application that searches across federal and state water data catalogs

Although the principal focus of the CUAHSI Hydrologic Information System project has been on water observations data, we have desired from the beginning of our project to provide access to an integrated array of information including static spatial water data in GIS, hydrologic modeling information, weather and climate model outputs, and remote sensing products, as illustrated in Figure 4. The information architecture we are envisaging with the Concept Development Study is focused on water observations time series but is sufficiently general that it can also be applied to acquire the other information types as well, which form components of a Hydrologic Information System.

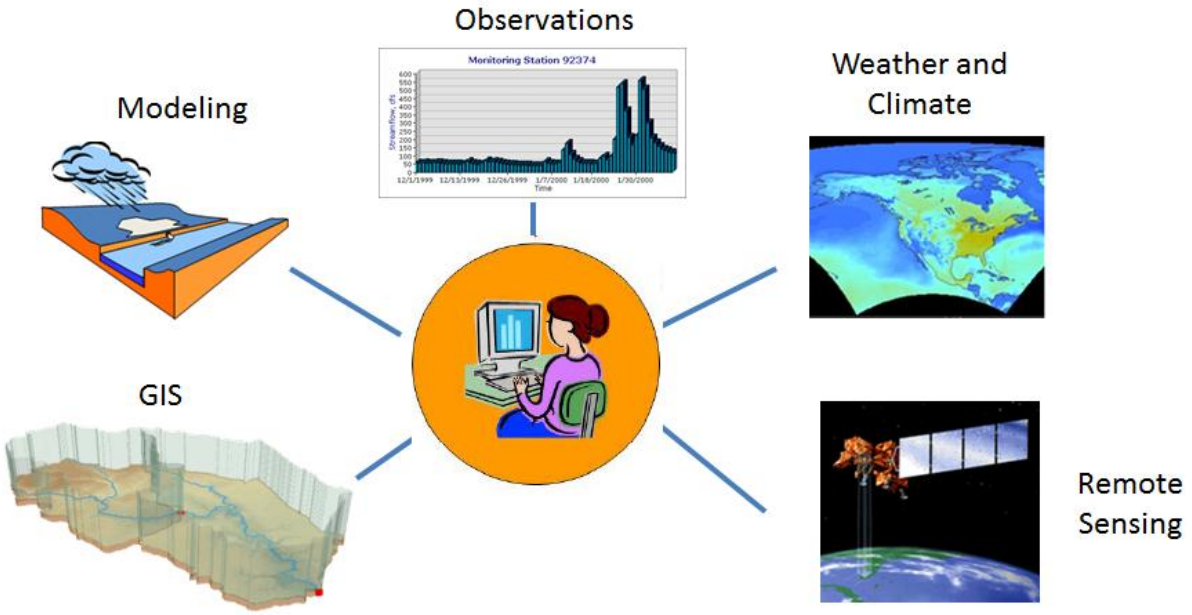

Figure 4. Components of a Hydrologic Information System

Conclusion

The purpose of this short paper is to provide the SWAQ with assurance that the information technology is available that can support a web-integrated hydrologic information system for the nation. What is needed is the articulation of a federal science and technology strategy that articulates from federal viewpoint what the Community Objectives for such a system should be, and why during a restricted budget period it is worth investing resources into the creation of such a system.

Reference

OGC (2008), Open Geospatial Consortium Reference Model, document ref: OGC 08-062r4 http://www.opengeospatial.org/standards/orm

Maidment, D.R., (2011) Water Data Sharing in the United States, Memorandum prepared for the Subcommittee on Water Availability and Quality, Executive Office of the President, 19 Jan 2011 http://www.caee.utexas.edu/prof/maidment/papers/SWAQ/WaterDataSharing.pdf