

Arizona Hydrologic Information System

Enhanced data accessibility using Web Services

Shankara Subramanya, Corinna Gries, Kathy Jacobs, Wolf-Dieter Otte, Ramon Vazquez, Joe Abraham, Raul Aguilar, James McGill, Huan Liu

Department of Computer Science and Engineering, Arizona State University; Global Institute of Sustainability, Arizona State University; Arizona Water Institute, Dept. of Computer Science, College of Engineering and Natural Sciences, Northern Arizona University; Sustainability of semi-Arid Hydrology and Riparian Areas (SAHRA), University of Arizona, P.O. Box 210158-B, Tucson, AZ 85721-0158

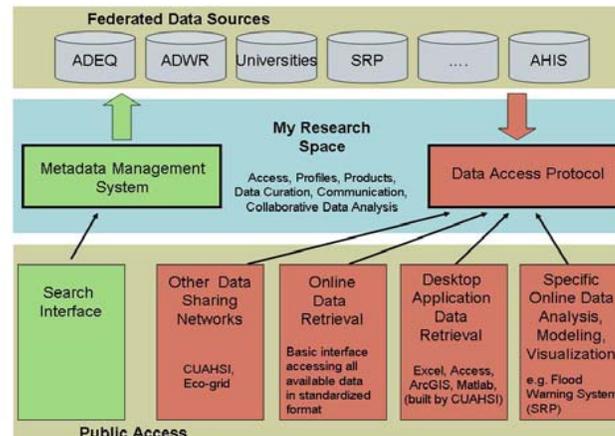
Introduction

In Arizona, diverse federal, state and local agencies and institutions independently gather, archive and present water related data. Although the Internet has improved access to these disparate data sources, many datasets are still generally unknown and inaccessible. Data that are accessible are provided on individual websites with proprietary data access protocols and data exporting formats. However, for efficient research and decision making in the realm of water management in the context of global change, population growth, and land management, rapid access to the existing wealth of data is imperative. This project addresses the issues currently preventing researchers, stakeholders, and decision makers from efficiently using and integrating existing data sources as a statewide, collaborative effort among the three state universities. The goal of AHIS is to develop the information infrastructure for data curation, including sustainable long-term archival and documentation, quality assessment and control, discovery, and access. In addition to the basic functionality, advanced data integration, analysis, modeling and visualization applications are being developed by AHIS and other entities.

The Arizona Water Institute (AWI) is a tri-university initiative (University of Arizona, Arizona State University, Northern Arizona University) partially funded by the Arizona legislature and the Board of Regents. AWI conducted a broad statewide needs assessment of 85 individuals and organizations in 2006, which indicated that an integrated data access system was the highest priority need both within the research community and across the communities, tribes and utilities of the state. The private sector also has a big stake in improved access to information, because the current approach to data collection is extremely labor intensive and may not identify all of the relevant data sources. Therefore, AHIS is a high priority effort of the Arizona Water Institute. It is the "cyberinfrastructure" that connects the three state Universities and state and federal agencies into a virtual clearinghouse for water information in Arizona.

Architecture

Built in a highly modular fashion (based on the principles of a 'Service Oriented Architecture' or SOA) and according to agreed-upon standards, AHIS will hide location and complexity of data from the user while allowing a variety of applications and tools to access the data. The design of the basic data access modules in AHIS will follow national standards set by the hydrological community in collaboration with CUAHSI (Consortium of Universities for the Advancement of Hydrologic Science) and SAHRA. Both have already developed access modules for national and regional databases and tools to use the data. Our efforts will therefore build on and complement a rich existing infrastructure.



Standards and Technology

Metadata

The Ecological Metadata Language (EML) will be used for encoding of the metadata (please see poster by Aguilar et al.)

Data Storage, Access and Exchange

The basic principle of AHIS is to accommodate the data storage format a data provider is currently using. Especially for large legacy databases change of the data model is not feasible. Through web services AHIS provides the data access mechanisms which expose the data to the system in a standard format. The web services implement four main methods which return data according to the WaterML standard for data exchange. Accordingly, custom webservice for each data source are responsible for mapping the existing data model to the exchange standard. The four main methods are:

- getValues: Given a site number, a variable, a start date, and an end date, this method returns a time series.
- getSite: Given a site number, this method returns the site metadata.
- getVariableInfo: Given a variable code, this method returns the variable metadata.
- getSites: Given an array of site numbers, this method returns the site metadata for each one.

WaterML is primarily designed for relaying fundamental hydrologic time series data and metadata between clients and servers, and to be generic across different data providers. Both the main methods and WaterML were developed by CUAHSI and are implemented here for compatibility.