WaterML, an Information Standard for the Exchange of in-situ hydrological observations

D. Valentine (1), P. Taylor (2), and I. Zaslavsky (3)

(1) University of California, San Diego, United States (valentin@sdsc.edu), (2) CSIRO Information, Communication and Technology Centre, Hobart, Australia (Peter.Taylor@csiro.au), (3) University of California, San Diego, United States (zaslavsk@sdsc.edu)

The WaterML 2.0 Standards Working Group (SWG), working within the Open Geospatial Consortium (OGC) and in cooperation with the joint OGC-World Meteorological Organization (WMO) Hydrology Domain Working Group (HDWG), has developed an open standard for the exchange of water observation data; WaterML 2.0. The focus of the standard is time-series data, commonly generated from in-situ style monitoring. This is high value data for hydrological applications such as flood forecasting, environmental reporting and supporting hydrological infrastructure (e.g. dams, supply systems), which is commonly exchanged, but a lack of standards inhibits efficient reuse and automation.

The process of developing WaterML required doing a harmonization analysis of existing standards to identify overlapping concepts and come to agreement on a harmonized definition. Generally the formats captured similar requirements, all with subtle differences, such as how time-series point metadata was handled. The in-progress standard WaterML 2.0 incorporates the semantics of the hydrologic information: location, procedure, and observations, and is implemented as an application schema of the Geography Markup Language version 3.2.1, making use of the OGC Observations & Measurements standards. WaterML2.0 is designed as an extensible schema to allow encoding of data to be used in a variety of exchange scenarios. Example areas of usage are: exchange of data for operational hydrological monitoring programs; supporting operation of infrastructure (e.g. dams, supply systems); cross-border exchange of observational data; release of data for public dissemination; enhancing disaster management through data exchange; and exchange in support of national reporting

The first phase of WaterML2.0 focused on structural definitions allowing for the transfer of time-series, with less work on harmonization of vocabulary items such as quality codes. Vocabularies from various organizations tend to be specific and take time to come to agreement on. This will be continued in future work for the HDWG, along with extending the information model to cover additional types of hydrologic information: rating and gauging information, and water quality. Rating curves, gaugings and river cross sections are commonly exchanged in addition to standard time-series data to allow information relating to conversions such as river level to discharge. Members of the HDWG plan to initiate this work in early 2012. Water quality data is varied in the way it is processed and in the number of phenomena it measures. It will require specific components of extension to the WaterML2.0 model, most likely making use of the specimen types within O&M and extensive use of controlled vocabularies.

Other future work involves different target encodings for the WaterML2.0 conceptual model, such as JSON, netCDF, CSV etc. are optimized for particular needs, such as efficiency in size of the encoding and parsing of structure, but may not be capable of representing the full extent of the WaterML2.0 information model. Certain encodings are best matched for particular needs; the community has begun investigation into when and how best to implement these.