



Bridging the gap between Hydrologic and Atmospheric communities through a standard based framework

E. Boldrini (1), F. Salas (2), D. R. Maidment (2), P. Mazzetti (3), M. Santoro (1), S. Nativi (3), and B. Domenico (4)

(1) National Research Council of Italy - Institute of Methodologies for Environmental Analysis (IMAA), Potenza, Italy enrico.boldrini@cnr.it, (2) Center for Research and Water Resources, University of Texas at Austin, Texas, USA maidment@mail.utexas.edu, (3) National Research Council of Italy - Institute of Atmospheric Pollution Research, Roma, Italy stefano.nativi@cnr.it, (4) Unidata, Boulder, Colorado, USA Ben@unidata.ucar.edu

Data interoperability in the study of Earth sciences is essential to performing interdisciplinary multi-scale multi-dimensional analyses (e.g. hydrologic impacts of global warming, regional urbanization, global population growth etc.).

This research aims to bridge the existing gap between hydrologic and atmospheric communities both at semantic and technological levels.

Within the context of hydrology, scientists are usually concerned with data organized as time series: a time series can be seen as a variable measured at a particular point in space over a period of time (e.g. the stream flow values as periodically measured by a buoy sensor in a river); atmospheric scientists instead usually organize their data as coverages: a coverage can be seen as a multidimensional data array (e.g. satellite images acquired through time).

These differences make non-trivial the set up of a common framework to perform data discovery and access.

A set of web services specifications and implementations is already in place in both the scientific communities to allow data discovery and access in the different domains.

The CUAHSI-Hydrologic Information System (HIS) service stack lists different services types and implementations:

- a metacatalog (implemented as a CSW) used to discover metadata services by distributing the query to a set of catalogs
- time series catalogs (implemented as CSW) used to discover datasets published by the feature services
- feature services (implemented as WFS) containing features with data access link
- sensor observation services (implemented as SOS) enabling access to the stream of acquisitions

Within the Unidata framework, there lies a similar service stack for atmospheric data:

- the broker service (implemented as a CSW) distributes a user query to a set of heterogeneous services (i.e. catalogs services, but also inventory and access services)
- the catalog service (implemented as a CSW) is able to harvest the available metadata offered by THREDDS services, and executes complex queries against the available metadata.
- inventory service (implemented as a THREDDS) being able to hierarchically organize and publish a local collection of multi-dimensional arrays (e.g. NetCDF, GRIB files), as well as publish auxiliary standard services to realize the actual data access and visualization (e.g. WCS, OPeNDAP, WMS).

The approach followed in this research is to build on top of the existing standards and implementations, by setting up a standard-aware interoperable framework, able to deal with the existing heterogeneity in an organic way.

As a methodology, interoperability tests against real services were performed; existing problems were thus highlighted and possibly solved.

The use of flexible tools, able to deal in a smart way with heterogeneity has proven to be successful, in particular experiments were carried on with both GI-cat broker and ESRI GeoPortal frameworks.

GI-cat discovery broker was proven successful at implementing the CSW interface, as well as federating

heterogeneous resources, such as THREDDS and WCS services published by Unidata, HydroServer, WFS and SOS services published by CUAHSI.

Experiments with ESRI GeoPortal were also successful: the GeoPortal was used to deploy a web interface able to distribute searches amongst catalog implementations from both the hydrologic and the atmospheric communities, including HydroServers and GI-cat, combining results from both the domains in a seamless way.