



Federating Metadata Catalogs

C. Baru and K Lin

University of California San Diego, San Diego Supercomputer Center, La Jolla, United States (baru@sdsc.edu)

The Geosciences Network project (www.geonetwork.org) has been developing cyberinfrastructure for data sharing in the Earth Science community based on a service-oriented architecture. The project defines a standard “software stack”, which includes a standardized set of software modules and corresponding service interfaces. The system employs Grid certificates for distributed user authentication. The GEON Portal provides online access to these services via a set of portlets. This service-oriented approach has enabled the GEON network to easily expand to new sites and deploy the same infrastructure in new projects.

To facilitate interoperability with other distributed geoinformatics environments, service standards are being defined and implemented for catalog services and federated search across distributed catalogs. The need arises because there may be multiple metadata catalogs in a distributed system, for example, for each institution, agency, geographic region, and/or country. Ideally, a geoinformatics user should be able to search across all such catalogs by making a single search request.

In this paper, we describe our implementation for such a search capability across federated metadata catalogs in the GEON service-oriented architecture. The GEON catalog can be searched using spatial, temporal, and other metadata-based search criteria. The search can be invoked as a Web service and, thus, can be imbedded in any software application. The need for federated catalogs in GEON arises because, (i) GEON collaborators at the University of Hyderabad, India have deployed their own catalog, as part of the iGEON-India effort, to register information about local resources for broader access across the network, (ii) GEON collaborators in the GEO Grid (Global Earth Observations Grid) project at AIST, Japan have implemented a catalog for their ASTER data products, and (iii) we have recently deployed a search service to access all data products from the EarthScope project in the US (<http://es-portal.geonetwork.org>), which are distributed across data archives at IRIS in Seattle, Washington, UNAVCO in Boulder, Colorado, and at the ICDP archives in GFZ, Potsdam, Germany. This service implements a “virtual” catalog—the actual/“physical” catalogs and data are stored at each of the remote locations.

A federated search across all these catalogs would enable GEON users to discover data across all of these environments with a single search request. Our objective is to implement this search service via the OGC Catalog Services for the Web (CS-W) standard by providing appropriate CSW “wrappers” for each metadata catalog, as necessary. This paper will discuss technical issues in designing and deploying such a multi-catalog search service in GEON and describe an initial prototype of the federated search capability.