The Implementation of Geospatial Web Services and Workflows in GeoBrain

Liping Di, Peisheng Zhao, Weiguo Han, Xiaoyan Li, and Meixia Deng
Center for Spatial Information Science and Systems, George Mason University
6301 Ivy Lane, Suite 620, Greenbelt, MD 20770, USA
ldi@gmu.edu

GeoBrain is a standard-compliant, open, distributed geospatial web service system developed and operated by CSISS, GMU. The GeoBrain system has implemented ISO TC 211 standards for geographic metadata, and the Open Geospatial Consortium (OGC) and World-Wide Web Consortium (W3C) standards for the data and service interoperability. Similar to any other web service systems, the power and the flexibility of GeoBrain relies on the availability of a large number of standard-compliant chainable web services that can process the geospatial data and information.

The web services in GeoBrain come mainly from two sources, 1) conversion from traditional Geographic Information System/Digital Image Processing packages and 2) development from scratch. In GeoBrain, the majority of geospatial web processing services come from the first sources through the conversion of all processing functions in a public-domain GIS/image processing system, called Geographic Resources Analysis Support System (GRASS), into SOAP-based chainable geospatial web services. These web services are grouped into imagery, raster, and vector processing services respectively. The service interface structures are designed to facilitate application integration and to ensure interoperability. URL is used to pass geospatial data into and out of the services. During the conversion process, we encountered and solved issues related to the web services, such as modularity, WSDL description, and concurrent multiple connections of services. A small number of web processing services in GeoBrain were developed from scratch by developers in CSISS and other universities, mainly in response to the functional needs of OGC Web Service (OWS) initiatives. Examples of such services include reprojection service and feature cutting service. The interface for those services is compliant with OGC Web Processing Service (WPS) specification.

Earth System Science (ESS) research and applications often involve in collecting, analyzing and modelling with a huge amount of multi-source, multi-scale and multi-discipline geospatial data. Those data are processed step-by-step in geospatial analysis systems to extract information and knowledge products for applications and decision makings. Conceptually, the step-by-step processes from the raw data to a user-specific product form a geospatial processing model. The model represents the knowledge of geospatial domain experts on how to produce an application-specific product from available raw data sources. The model is expressed as a workflow in GeoBrain. In GeoBrain, two geospatial processing modelling approaches are implemented- expert creation and ontology-based automatic creation. For the expert creation of geoprocessing models, a graphic user interface is developed for assisting experts to create the processing models at abstract level. The abstract models are stored in GeoBrain as virtual products. The abstract model is converted automatically to an executable BPEL workflow when a user requests the virtual product. The workflow is executed by GeoBrain's BPELPower workflow engine to generate the product for the user on-the-fly. The automatic creation of geospatial processing models is driven by knowledge represented in geospatial and

application-specific ontologies, including 1) geospatial scientific ontology, which defines geospatial scientific concept and phenomena; 2) geospatial data ontology, which defines data product relevant to specific scientific concept and its metadata; and 3) geospatial processing ontology, which defines how to generate a data product and indicates the input, output, precondition and post-condition. To achieve the desired geospatial product, the ontology-driven modeling engine assembles individual geospatial processing services into a service workflow which represents the geospatial processing model. The workflow then is executed in BPELPower workflow engine to generate the product.

In GeoBrain, all distributed processing services, workflows, and data resources are accessible through the service portal called GeOnAS, a fully extensible online analysis system for discovering, retrieving, analyzing and visualizing geospatial and other network data. The most distinguished characteristics of GeOnAS are 1) online--users only need to have an Internet-connected PC with a Web browser; 2) OGC standard-based data access—the system provides a single point of entry to the geospatial data from any OGC compliant data services worldwide; 3) Web service-based analysis--all functions are provided through interoperable Web services so that users of the system can easily perform a specific analysis by integrating new services or chaining available services.

The GeoBrain system, although still in developing, has been deployed in an Apple G5 server cluster since 2005. The server cluster has been populated with more than 20 TB of typical EOS data for users to generate personalized, on-demand, value-added products. It also can access and analyze all on-line NASA EOS data catalogued in NASA ECHO. The GeoBrain URL is http://geobrain.laits.gmu.edu:81/OnAS/.