International Journal of Spatial Data Infrastructures Research, 2009, Vol. 4, 77-95.

Building an Online Learning and Research Environment to Enhance Use of Geospatial Data*

Meixia Deng, Liping Di
Center for Spatial Information Science and System, George Mason University
mdeng@gmu.edu, ldi@gmu.edu

Abstract

Geospatial data availability, interoperability, and integration remain a problem today. Current spatial data infrastructures (SDIs) are of limited use particularly to non-expert user communities. GeoBrain, a NASA funded project, has aimed to address those challenges, facilitate easy use of geospatial data and overcome some limitations of current SDIs through building a data-intensive, Service-oriented Architecture (SOA) based online learning and research environment. By adopting the latest developing Web services and knowledge management technologies, this online environment enables easy, open, seamless, and on-demand discovery, access, retrieval, visualization and analysis of distributed geospatial data, information, services, and models from any computer connected to the Internet. Such an online environment is able to serve the different needs of global Earth sciences research and higher education communities, bridge gaps between data user needs and provider capabilities, and greatly enhance use of geospatial data.

Keywords: geospatial data, interoperability, Web services, SDI, online learning and research environment

1. INTRODUCTION

As the amount of diverse geospatial data collected by private, public, and military sectors around the world grows, data availability, interoperability, and integration are becoming increasingly critical in almost all application domains. Significant efforts on and progress in increasing the availability and accessibility of geospatial data and facilitating the integration of multi-source diverse data have occurred in recent years. One of the most significant advances is the development of Spatial Data Infrastructures (SDIs) and associated technologies [Craglia et al, 2008]. Most recent large SDI developments in the U.S. have significantly enhanced the sharing of geospatial data and associated computing

DOI: 10.2902/1725-0463.2009.04.art4

resources over a large geographic region. Examples of such SDIs include the U.S. National Spatial Data Infrastructure (NSDI) and the System Grid project (ESG). The NSDI aims to provide "consistent means to share geographic data among all users could produce significant savings for data collection and use and enhance decision making" [NSDI Web site]. ESG uses the XML-based metadata representation and Grid computing technology to create a virtual environment, which links distributed centers, users, models and data as an integrated resource to support the next generation of climate modeling research [Bernholdt et al, 2005].

Despite the significant progress in interoperability and sharing of distributed geospatial resources, the current operational SDIs can mainly provide only data search, retrieval, and access. Thus, the accessed data from multiple sources will need to be further processed so that the data can be integrated and analyzed. Due to the notorious complexities of the geospatial data (e.g. multiple sources, heterogeneous data structure and types, different formats and projections, incompatible spatial resolutions, variable spatial and temporal coverage and scales), such processing needs expertise on the data, the processing algorithms, domain knowledge, and computer software and hardware resources. Such requirements prevent use of geospatial data by non-expert or resource-lacking users, as summarized by Craglia et al [2008]. In fact, even for expert users, the lack of semantics of data and mechanisms of transforming data to a common spatial or temporal reference system hampers integrating data from multiple sources and disciplines [Kuhn, 2003]. Consequently, Earth scientists are still crying, "Why is Earth science data so difficult to use?" For example, using multi-source geospatial data has been suffering from two key data usability problems, as identified in an extensive discussion during the 7th NASA Earth Science Data Systems Working Group Meeting (ESDSWG, 2008): 1) It is still difficult to locate specific data of interest, and 2) data are difficult to use once obtained. Current scientific studies are largely hampered by these data usability problems. Scientists spend considerable amount of time assembling the data and information into a form ready for analysis, even if the analysis is very simple. The process for a user from discovery to actually ingesting the data into the analysis systems usually takes weeks. Therefore, many studies requiring real or near-real time data cannot be conducted. Because users need to spend a lot of time and computing resources to process the data into the format, projection, geographic area or time period they want, they cannot perform many tasks, due to lack of computing resources or prior knowledge (e.g. familiarity with data formats and format conversions, reprojection tools, and subsetting/resampling packages).

Expert users are expecting to find and obtain the data and information needed in a ready-to-use form that has already been processed and integrated by a single tool or an integrated environment (not necessarily the same tool or environment for every person or task) so that they can spend more time and resources on conducting scientific study and less on obtaining, preprocessing, and integrating data. Non-expert users, such as decision makers and students, expect more geoprocessing and data analysis services, information extraction, and knowledge generation tools for problem-oriented help and guidance. In general, all user communities ultimately desire easy-to-use, dynamic and on-demand data services with coupled intelligence and automation. To realize this vision, we need to solve a key problem -- how to let users easily obtain data and information from separate data archives in the exact forms they want, regardless where and how the data and information are stored and managed in the archives. Solving this problem will also provide the solution to many challenging issues of geospatial data availability, interoperability, and integration. This paper describes the well-integrated, geospatial Web-services based, data-intensive online learning and research environment established by the GeoBrain project, which provides a solution to this problem.

2. FUNCTIONAL REQUIREMENTS AND APPROACHES

The goal of the GeoBrain project is to provide solutions addressing the needs of the Earth Sciences communities for timely and ready access to and easy use of Earth and environmental (geospatial) data. With the current data infrastructure and information technologies, there are many challenges or difficulties in addressing those needs. A key challenge, as just introduced, is how to let users easily obtain data and information from different data centers in the exact forms they want. Meeting this key challenge requires removing major obstacles that prevent current data and information systems from being interoperable at the data, functional services, and system levels. In other words, three types of interoperability need to be addressed: catalogue interoperability (consistent metadata description), data interoperability (consistent data models and access protocols), and service interoperability (interoperable services). The GeoBrain project addresses the challenges and interoperability issues by developing and operating a standards-compliant, Web service based data, information, and knowledge building system—GeoBrain [Di, 2004; Deng and Di, 2006]. A Web service is a software system designed to support interoperable machine-to-machine interaction over a network. By connecting interoperable Web services, it is easy to construct a geoprocessing workflow and create a high-level knowledge model.

The GeoBrain system aims to provide innovative capabilities for discovering, accessing, visualizing, processing, integrating and analyzing geospatial data and to enable geospatial processing, modeling and knowledge building by establishing a well-integrated, data-intensive, and collaborative online learning and research environment. This online environment is an interoperable geospatial Web service based portal system whose objectives are to address the issues that make use of Earth science data so difficult, among them:

1) The distribution of geospatial data over different data centers with diverse data access protocols.
2) The significant time required by current data obtaining procedures.
3) The format, projection, and coverage complexities and data integration difficulties of geospatial data.
4) Lack by most users of enough computing resources to do processing, analysis or modeling of data.

The following outlines the functional requirements of building the GeoBrain online environment and the approaches to implement these requirements.

A. Making Data and Services Resources Easily Online Available

   Making distributed geospatial data and services resources easily available to users online, with a single point of entry, will free users from worrying about lack of data and computing resources, save users significant time and effort in discovering and acquiring data, and simplify some complicated scientific tasks. The Geospatial Catalogue Federation Service provides a method for distributed and integrated metadata discovery [Bai et al., 2007]. The major challenge in building a catalog federation service is to solve the interoperability of different catalog protocols and information models used by different data sources. With a Catalogue Federation Service, data and services in distributed systems can be accessed through machine-to-machine interfaces from a single access point and provided to users online.

B. Customizing Data At User Request

   Most current data systems still adopt the one-size-fits-all approach, without recognizing different users' requirements for data. The complexities of geospatial data cause big gaps between data users' needs and data providers' capabilities. Providing customized data at user request effectively bridges the gaps and eliminates the difficulties in using geospatial data. By developing interoperable, personalized, and on-demand data access and services (IPODAS), the online environment recognizes users' different requirements for data. "Personalized" will ensure that data services tailor data according to user's needs. "On-demand" means that data products must be created on-demand because the exact needs of users are not known. After users have found data of interest through a federated search from multiple (distributed) resources, IPODAS then processes them into the user-defined form with on-the-fly subsetting, reformatting, reprojecting, and resampling services (automated data preprocessing). IPODAS ensures that users obtain data in the form that exactly matches user's needs in term of format, spatial and temporal coverage, projection, and resolution.

C. Towards Automated Multi-source Geospatial Data Integration

Integration of geospatial data from multiple sources requires adequate data preprocessing skills (e.g. subsetting, reformatting, reprojecting, resampling) and thorough understanding of the data syntax and semantics. Full automation of multi-source data integration remains a very challenging research issue. The online environment enables automated or semi-automated data integration through a workflow (service chain) based approach: first developing both data and service ontologies to semantically describe the data and services, and then developing service chaining techniques (e.g., the path planning algorithm, and workflow management). A service chain is defined as a sequence of services where, for each adjacent pair of services, occurrence of the first action is necessary for the occurrence of the second action. Interoperable geospatial Web services can be chained together (as a workflow) to solve complex tasks [Deng et al., 2004]. A data integration workflow can be automatically (or semi-automatically) constructed through service chaining methods (e.g. ontological reasoning) and executed by the workflow engine to generate the integrated data.

D. Providing Online On-demand Data Mining, Visualization and Analysis

Geospatial data, such as remote sensing data, are mostly multi-spectral or even hyper-spectral images. They have large dimensions, with multiple spectral bands (sometimes more than two hundred), far from the "readability" of other simple data types, such as text and numbers. Visual interpretation, mining (extracting useful thematic information from spectral signatures of the image) and analysis of remote sensing data rely largely on proper computer-based techniques. Enabling online on-demand geospatial data mining, visualization and analysis will dramatically facilitate scientific studies that need to use geospatial data and will ultimately free users from worrying about the lack of computing resources or skills. Geospatial Web services provide great possibilities and flexibility for geospatial processing or for computing tasks in distributed environments [Deng et al., 2003; 2004]. Online on-demand data analysis and visualization services can be provided by extensible geospatial Web service based portal system with the following characteristics:

1) Accessible online -- users only need to have an Internet connected PC with a Web browser to conduct complicated tasks needing huge computing resources.
2) Standards-based data access-- the system provides a single point of entry to the geospatial data using standards compliant data services.
3) Geospatial Web service based visualization and analysis -- all functions are provided through interoperable geospatial Web services, allowing users of the system to easily perform on-demand analysis by integrating new services or chaining built-in services.

E. Geospatial Processing, Modeling, and Knowledge Building and Sharing

Most geospatial scientific products are not obtained directly from measurements but derived from other data by building and executing a geospatial model. A geospatial model always represents behavior-based high-level geospatial knowledge and involves a computational procedure. A Web interface with geospatial modeling toolkits can be used as the approach to promote geoprocessing, geospatial modeling, and knowledge building. More specifically, an Abstract Model Designer based on Java Applets can be developed to enable domain experts to use data types, service types, and existing abstract models as basic components of new abstract models constructed online and registered in the system.

3. RELATED WORK

The GeoBrain online learning and research environment is built on numerous previous or on-going research efforts. The following projects are relevant to the online environment.

The Data and Information Access Link (DIAL) [Di et al, 1999; Suresh et al, 2001] was a project funded by the NASA Earth Science Data and Information System (ESDIS) Program. The project developed the first Web-based satellite remote sensing data distribution system. The system enables users to interactively discover archived satellite remote sensing data and retrieve the data in user requested format, spatial and temporal coverage, and bands/parameters.

The Joint Geoinformatics Research project, funded by NASA ESDIS, develops interoperability standards and technology for sharing geospatial data, information, knowledge, systems, and algorithms at distributed environments. The project has contributed to the development of Federal, national, and international standards and specifications on geospatial data and system interoperability [Di, 2003], and led to the development of the Federal Geographic Data Committee (FGDC) Content Standard for Geospatial Metadata, Extensions for Remote Sensing Metadata [Di et al, 2000; FGDC, 2002], ISO 19130 Geographic Information-Sensor and Data Models for Imagery and Gridded Data [Di et al., 2002a; ISO TC 211, 2004], and the Open Geospatial Consortium (OGC) Web Coverage Service (WCS) Implementation Specification [Evans et al., 2003]. The project has also developed the NASA HDF-EOS Web GIS Software Suite (NWGISS) based on OGC specifications [Di et al., 2002b; Di, 2006 a, b]. NWGISS provides experiences to develop IPODAS for the online environment addressed in this paper.

The Thematic Real-time Environmental Distributed Data (THREDDS) [Domenico et al., 2002; Fulker et al., 1997] project is part of National Science Digital Library (NSDL). THREDDS is a collaborative initiative to build a software infrastructure to provide students, educators, and researchers with Internet access to large collections of real-time and archived environmental datasets from distributed servers. In the initial phase, THREDDS established a solid, working prototype of services and tools enabling data providers to create inventory catalogues of the data holdings at their sites as well as enabling educational model builders to author compound documents with embedded pointers to environmental datasets and analysis tools. These catalogues and data-interactive documents could then be harvested into digital libraries using standard protocols.

The Earth Science Markup Language (ESML) [Ramachandran et al, 2003 a,b and 2004] project is funded by the NASA Earth Science Technology Office to study, design, and develop ESML as an *interchange technology*. An interchange technology solves the data-to-application interoperability problem by enabling applications to work seamlessly with datasets of heterogeneous format, using external metadata. ESML specifications and libraries are publicly available.

The Algorithm Development and Mining (ADaM) Toolkit**,** funded by NASA, consists of over 75 interoperable data mining and image processing components for Earth System Science data that can be arbitrarily chained to solve a range of problems. These components can carry out distributed mining and image processing services in Web and GRID applications. Individual ADaM components can execute in a standalone mode, allowing their use in distributed and parallel processing systems. Integrating the components provides pattern recognition, image processing, optimization, and association rule-mining capabilities, packaged as C/C++ executables and Python modules.

The Semantic Web for Earth and Environmental Terminology (SWEET) [Raskin, 2003] is a NASA-funded prototype for improving semantic interoperability. This project has developed a set of ontologies for Earth science that serves as a scalable, common semantic framework. The SWEET ontologies are written in the

XML-based Ontology Language for the Web (OWL), a standard adopted by the W3C. The project also has developed an ontology-assisted search tool to improve search performance.

The Geosciences Network (GEON) is a multi-institution coalition of IT and Earth science researchers building a cyberinfrastructure for the geological sciences. It develops technologies to enable geoscientists to integrate, analyze, model, and visualize enormous, complex multidisciplinary 4-D solid Earth data sets. By providing leading-edge data integration and Grid computing services to support geosciences research and collaboration on unprecedented scales, GEON enables new insights into the complex dynamics of the solid Earth system. Ontology is integrated into GEON-Grid visualization and data modeling environment.

4. ARCHITECTURE DESIGN

A Service-oriented Architecture (SOA) has been adopted to establish an online learning and research environment with the above functional objectives. The SOA uses loosely coupled and interoperable Web services to implement system requirements. All the services within the SOA are independent so that they can be accessed in a standard way without knowledge of how the service actually performs its tasks. Moreover, the SOA can support integration and orchestration of individual services into composite services. The SOA design enables the system to link distributed computational resources to support geospatial analysis and make it fully extensible and self-evolvable. The more users are involved, the more powerful the system becomes.

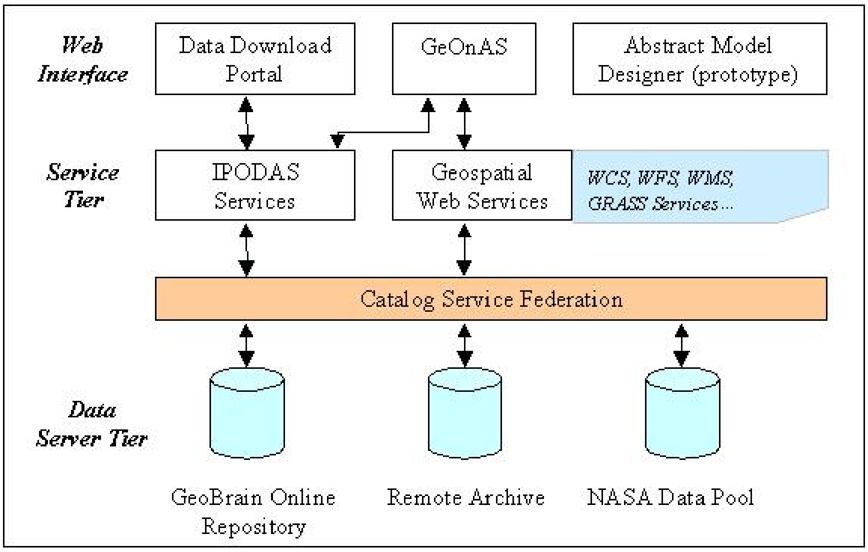
Based on SOA, the GeoBrain online learning and research environment is an open, three-tier, multi-component system. Figure 1 is the system architecture schematic showing major components in the three tiers:

The Data Server Tier is at the back-end. It includes many petabytes of Earth Observing System (EOS) data at NASA online data pools running at DAACs, remote data archives of NOAA and USGS, and the GeoBrain online data repository. These distributed geospatial data resources, which are important for Earth and environment sciences studies, can be discovered and accessed transparently and simultaneously by the system through the Catalog Service Federation.

The Service Tier is a geospatial service and knowledge management middleware tier**.** It consists of multiple components and numerous service modules that perform geospatial data processing, information extraction, and knowledge management. The IPODAS, the on-demand data visualization, mining, analysis services, the workflow construction and execution services, and other value-added geospatial Web services are enabled through this tier.

The Web Interface Tier is at the front-end. It currently includes the Web portals for download of geospatial data and products, the GeoBrain Online Analysis System (GeoOnAS), and the Abstract Model Designer. This tier interacts with the middle Service Tier and provides end users easy online access to all real data/information available through the system and virtual products generated by the system, the geospatial modeling and knowledge sharing functionalities, and the collaborative development capabilities enabled by the system.

Figure 1 System Architecture Schematic Showing Major Components



5. IMPLEMENTATIONS

To implement the online learning and research environment, OGC standards and technology specifications for geospatial data, service and system interoperability are adopted, developed, or extended for promoting data, functional services, and system level interoperability. OGC is the only international organization dedicated to developing geospatial implementation standards based on ISO, FGDC, INCITS, and other organizations' abstract or content standards. OGC standards have been fully tested both in NASA's and other data providers' environments. OGC specifications are widely used by geospatial communities for sharing data and resources. Besides OGC standards, many widely used international standards and specifications such as W3C and ISO standards are complied.

Many standards-compliant technologies have been developed for establishing the GeoBrain online environment, including but not limited to:

1) The OGC standards-based Web Coverage Service (WCS), Web Map Service (WMS), Web Feature Service (WFS), and Catalog Service-Web Profile (CSW) servers and clients.
2) The OGC Geoprocessing services, e.g., Web Coordinate Transformation Service (WCTS), Web Image Classification Service (WICS), Feature Cutting Service, Reformatting Service.
3) Many geospatial Web services, either converted from Geographic Resources Analysis Support System (GRASS) or built on with Geospatial Data Abstraction Library (GDAL), and functional as either OGC WPS services or SOAP services.
4) The Business Process Execution Language (BPEL) Workflow Engine extended for Geospatial Web Services, called BPELPOWER.
5) Web service chaining and modeling tools.

These technology implementations provide the foundation for meeting the functional requirements discussed in Section 2. For example, IPODAS is based on the OGC Web data services specifications WCS, WMS, WFS, and CSW, since these technology specifications allow seamless discovery of and access to geospatial data in a distributed environment, regardless of the format, projection, resolution, and archive location. Similarly, standards and specifications from OGC, W3C, ISO, and other international standards and technology groups provide the technology foundation for developing interoperable Web services and geospatial Web services, knowledge representation and management, and workflow management.

5.1 GeoBrain Online Data Resources

Remote sensing data, particularly those produced by the NASA Earth Observing System (EOS) program, are the most common data used in Earth, climate and environment sciences research and education because of the variety of data sources, multi-disciplinary coverage, science-oriented data collection, and no or low cost of the data. GeoBrain makes large amounts of NASA EOS data and other typical geospatial data sets, useful for global change study, accessible to users easily online in two ways. An online data repository is embedded in GeoBrain local storage to hold key remote sensing data for fastest and easiest access, and machine-to-machine interfaces are established between the GeoBrain and NASA EOS online data pools.

The GeoBrain Online Data Repository is now populated with about 20 TB of NASA EOS data (e.g. global coverage LANDSAT data and typical ASTER data samples) and datasets from sources other than or derived from the NASA EOS program but important for global change research activities (e.g. WindSat, Blue Marble and DMSP city lights data). About 600 TB of EOS data in four NASA EOS online data pools, running at NASA GSFC, USGS EDC, NSIDC, and NASA LaRC, are available, using machine-to-machine interfaces through the GeoBrain system.

Users get transparent access to all GeoBrain data resources through its catalog federation service [Bai, et al, 2007], which provides a single access point to two distributed catalogues: a local CSW catalog and an external catalog system -- NASA Earth Observing System Clearinghouse (ECHO). The local CSW catalog maintains metadata for all data and services in the GeoBrain system and any data resources in other distributed online systems, which are shared by registering their metadata. NASA's ECHO is a metadata clearinghouse and order broker being built by NASA's Earth Science Data and Information System (ESDIS) to support efficient discovery and access to Earth Science data, especially those at NASA Distributed Active Archive Centers (DAACs).

More data resources located in distributed systems are expected to be available through GeoBrain by either registering their metadata in the GeoBrain local CSW or federating their catalogues similarly as ECHO.

5.2 Data Products Download Portal with IPODAS

The Data Products Download Portal is designed and implemented to provide users advantages in discovering, accessing, and downloading data with IPODAS from distributed GeoBrain data resources through regular Web browsers such as Internet Explorer and Firefox. With the Portal, users can get ready-to-use data for their further study instead of wasting significant effort and computing resources in preprocessing data.

The Portal provides an integrated easy-to-use Web interface for users to search for, based on different searching criteria, data from distributed locations (e.g. in GeoBrain Online Data Repository or in different NASA online data pools), to visualize data, and to obtain data in the exact form desired (e.g., format, coverage, resolution, projection). Figures 2, 3 and 4 are snapshots of use of the GeoBrain Data Download Web Portal to search for and retrieve data. Data search criteria can cater to users' specific interests e.g. spatial coverage, temporal coverage, topics, sensor platforms, and parameters (Figure 2). The Portal returns only those data sets matching users' requests. Users then select data sets for downloading (Figure 3). Users can customize this downloading by specifying spatial coverage, resolution. projections, and formats so that they can obtain the data sets in the exact form they want through the embedded subsetting, resampling, reprojection and reformatting services (Figure 4).

With the portal, multi-source geospatial data acquisition and preprocessing typically needing hours, days or even weeks can be shortened to just *minutes or seconds*.

Figure 2: Defining Data Search Criteria

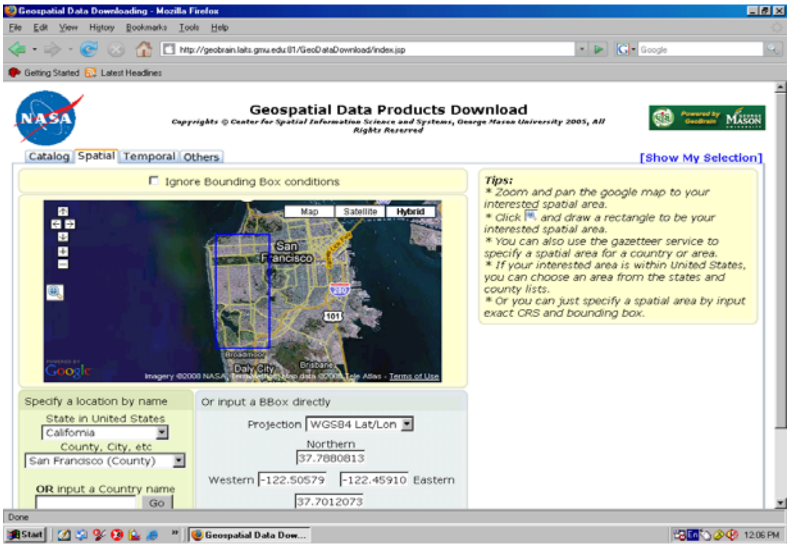


Figure 3: Data Found and Selection for Downloading

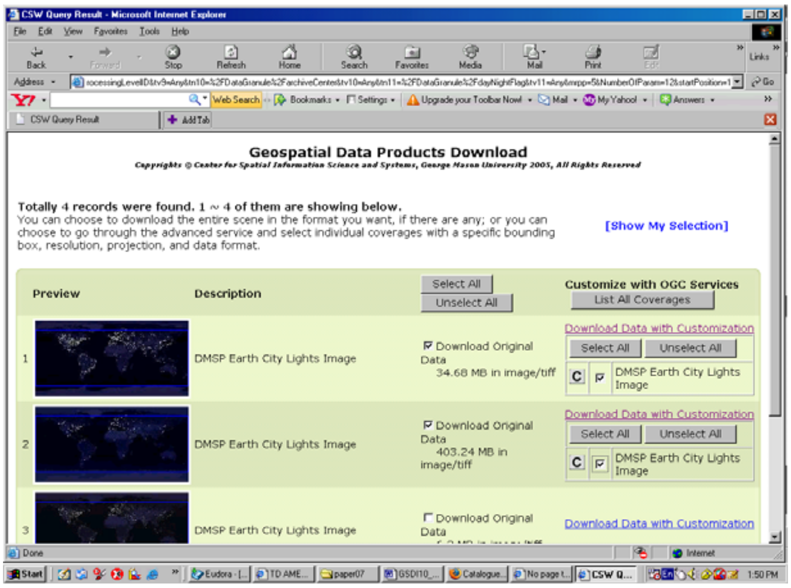
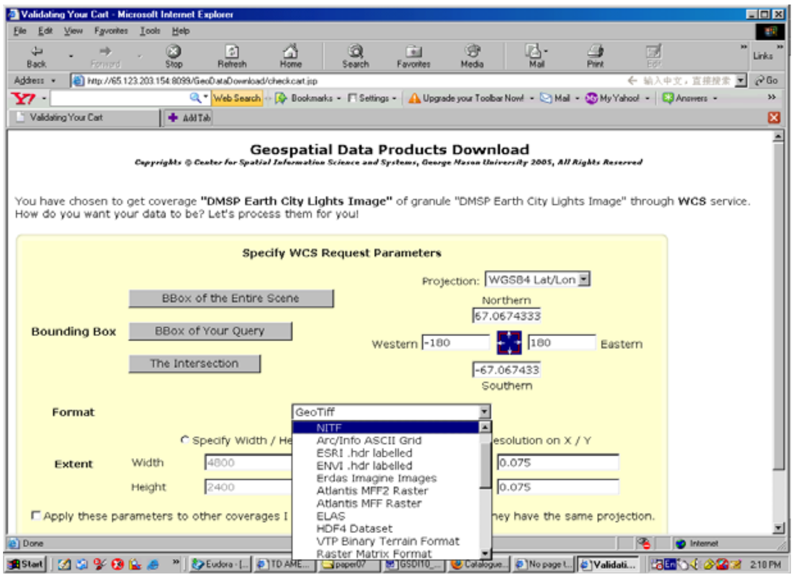


Figure 4: Downloading Data with Customization



5.3 GeoBrain Online Analysis System (GeoOnAS)

The GeoBrain Online Analysis System (GeoOnAS) is designed for online on-demand visualization, mining, and analysis of all data searchable through GeoBrain in order to provide users derived information products, in addition to raw data or customized data products. With GeoOnAS, users can be largely relieved from worrying about their lack of computing resources, data processing and analysis skills and knowledge so that they can focus on doing scientific explorations or fulfilling complicated research tasks.

GeoOnAS is a fully extensible portal system based on Service-Oriented Architecture (SOA) for multi-source discovery, heterogeneous retrieval, simultaneous visualization, and dynamic computation and analysis of geospatial and other network data [Di, et al, 2007]. The SOA uses loosely coupled and interoperable Web services to implement system requirements. All the services within the SOA are independent so that they can be accessed in a standard way without knowledge of how the service actually performs its tasks. Moreover, the SOA can support integration and orchestration of finely grained services into coarsely grained composite services. The functionalities of GeoOnAS are powered by numerous interoperable geospatial Web services for manipulating and analyzing vector and raster geospatial data and by its mechanism for integrating any standard based Web services into the system. GeoOnAS allows users to dynamically preprocess, integrate, and analyze any part of the petabytes of the data searchable through GeoBrain and get back derived information products as well as original and customized data. GeoOnAS has all the functionalities that the Data Product Download Portal has for data discovery, access, and retrieval and provides significant values by improving the overall efficiency and accuracy of processing, integrating, and analyzing distributed heterogeneous geospatial data over the Web. With GeoOnAS, users only need to use regular Web browsers to perform many complicated geospatial projects. Figures 5, 6 and 7 are screenshots showing how GeoOnAS is used.

Figure 5: Launching GeoOnAS through Web Interface

Figure 6: Defining Study Area for New Project

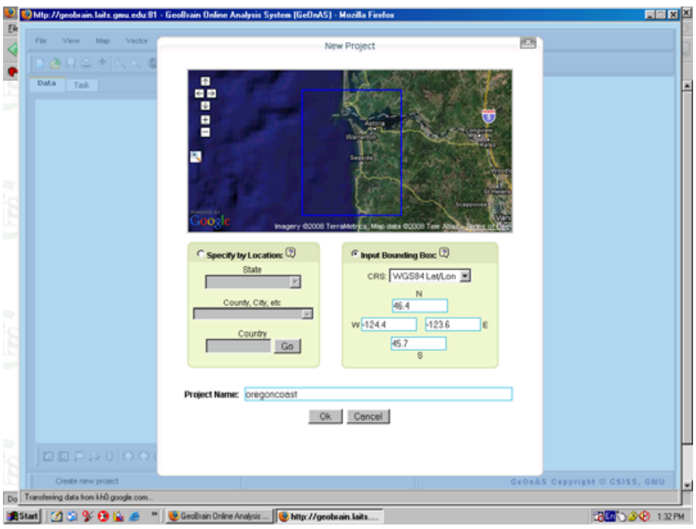
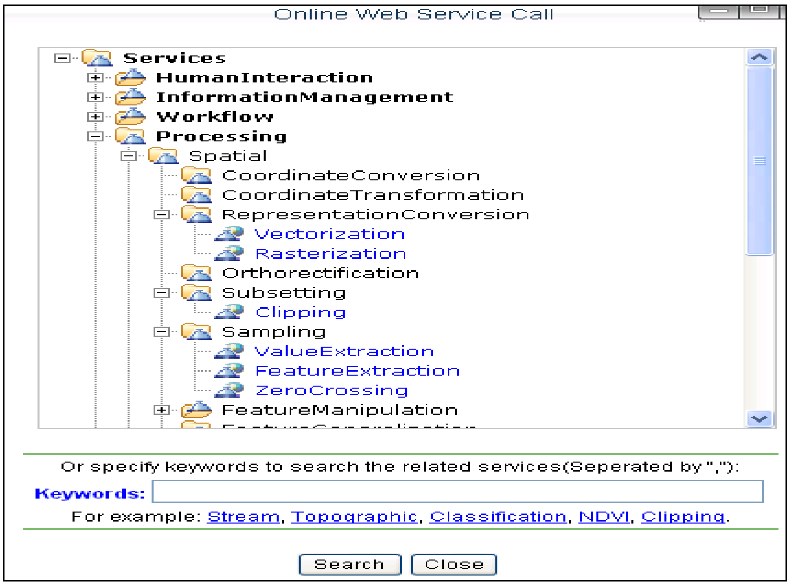


Figure 7: Finding Services in GeoOnAS for Project



6. CONCLUSION

The GeoBrain online learning and research environment has been successfully established. It makes vast data, services, and computing resources more available and easily accessible. It addresses key challenges of geospatial data usability and the interoperability at the data, functional services, and system levels. It takes into account the current limitations of data-intensive and distributed computing and improving the computational capabilities. Users are provided with great functionality, flexibility and ease in obtaining, accessing, and using geospatial data and information by two operational components implemented in the environment: the Data Download Portal with IPODAS and GeoOnAS. The online environment is a standards-compliant, data-enhanced, interoperable geospatial Web service based, fully extensible geospatial Web portal system. It has significantly enhanced the use of geospatial data by enabling data and information providers, scientists, educators, and students to more easily publish, discover, access, process, integrate, and analyze Earth science data and information from distributed Web sources. Such an environment will also have positive impact on how and by whom the research is conducted (precious resources that used to be only available to a few scientists or professionals in a limited way are now available on-demand 24/7 to everyone), and to the students' learning behavior.

ACKNOWLEDGEMENT
This research was supported by a grant from NASA REASoN program (NNG04GE61A, PI: Dr. Liping Di).

WEBSITES

ADaM, http://datamining.itsc.uah.edu/
ECHO, http://www.echo.nasa.gov/
ESG, http://www.earthsystemgrid.org/
GeoBrain, http://geobrain.laits.gmu.edu
GeoBrain Data Download Portal,
http://geobrain.laits.gmu.edu:81/GeoDataDownload/
GeoOnAS, http://geobrain.laits.gmu.edu:81/OnAS/
GEON, http://www.geongrid.org/
GRASS, http://grass.osgeo.org/intro/general.php
GSDI, http://www.gsdi.org/
ISO, http://www.iso.ch/iso/
NVO, http://www.us-vo.org/
NSDI, http:// http://www.fgdc.gov/nsdi/nsdi.html
OGC, http://www.opengeospatial.org/
SWEET, http://sweet.jpl.nasa.gov
THREDDS, http://www.unidata.ucar.edu/projects/THREDDS/
W3C, http://www.w3.org

REFERENCES

Bai, Y., L. Di, A. Chen, Y. Liu, and Y. Wei, 2007. “Towards a Geospatial Catalogue Federation Service”, *Photogrammetric Engineering & Remote Sensing*. Vol. 73, No. 6, June 2007, pp. 699–708.

Bernholdt, D., S. Bharathi, D. Brown, K. Chanchio, M. Chen, A. Chervenak, L. Cinquini, B. Drach, I. Foster, P. Fox, J. Garcia, C. Kesselman, D. Middleton, V. Nefedova, L. Pouchard, A. Shoshani, A. Sim, G. Strand, and D. Williams, 2005. “The Earth System Grid: Supporting the Next Generation of Climate Modeling Research”, *Proceedings of IEEE*, Vol. 93, No. 3, March 2005.

Craglia, M., Goodchild. M., Annoni, A., Camara, G., Gould, M., Kuhn, W., Mark, D., and Masser, I., 2008. “Next-Generation Digital Earth”, International *Journal of Spatial Data Infrastructure Research*, 2008, Vol. 3, pp. 146-165.

Deng, M., Liu, Y., and Di., L., 2003. “An Interoperable Web-based Image Classification Service for Remote Sensing Data”, *Proceedings of 2003 Asia GIS Conference*, *Asia GIS Society,* Oct 16-18, Wuhan, China, 11p. (CD-ROM publication).

Deng, M., Zhao, P., Liu, Y., Chen, A. and Di L., 2004, “The Development of A Prototype Geospatial Web Service System for Remote Sensing Data”, *The*

International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences, http://www.isprs.org/congresses/istanbul2004/comm2/papers/126.pdf

Deng, M. and Di, L., 2006. “Utilization of Latest Geospatial Web Service Technologies for Remote Sensing Education Through GeoBrain System”, *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2006)*, Colorado, USA.

Di, L., R. Suresh, K. Doan, D. Ilg, and K. McDonald, 1999. "DIAL-an Interoperable Web-based Scientific Data Server". In M. Goodchild, M. Egenhofer, R. Fegeas, and C. Kottman edit, Interoperating Geographic Information Systems, Section 4. System Experiences. Kluwer Academic Publishers, Norwell, MA. p 301-312.

Di, L., and K. McDonald, 1999. "Next Generation Data and Information Systems for Earth Sciences Research", proceedings of the First International Symposium on Digital Earth, Volume I. Science Press, Beijing, China. p 92-101.

Di, L., B. Schlesinger, and B. Kobler, 2000. " U.S. FGDC Content Standard for Digital Geospatial Metadata: Extensions for Remote Sensing Metadata", *International Archives of Photogrammetry and Remote Sensing (IAPRS)*, Vol. XXXIII, http://www.isprs.org/publications/archives.aspx

Di, L., W. Kresse, and B. Kobler, 2002a. “The Progress and Current Status of ISO Standard Development Project 19130-Sensor and Data Models for Imagery and Gridded Data.” In Chen and Jiang, eds., *The International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*, Vol. XXXIV, Part 2, Commission II, pp 107-110.

Di, L., W. Yang, M. Deng, D. Deng, and K. McDonald, 2002b. “Interoperable Access of Remote Sensing Data Through NWGISS”, *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2002)*, Toronto, Canada, p. 255-257.

Di, L., 2003. “The Development of Remote-Sensing Related Standards at FGDC, OGC, and ISO TC 211.” *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2003)*, July 21-25, Toulouse, France.

Di, L., A. Chen, W. Yang, and P. Zhao, 2003. “The Integration of Grid Technology with OGC Web Services (OWS) in NWGISS for NASA EOS Data,” *Proceedings of the Eighth Global Grid Forum (GGF8)*, 24-27 June 2003, Seattle, WA, USA.

Di, L., 2004. “GeoBrain-A Web Services based Geospatial Knowledge Building System.”, *Proceedings of NASA Earth Science Technology Conference 2004. June 22-24, 2004, Palo Alto, CA, USA. (CD-ROM).*

Di, L., 2006a. "The Open GIS Web Service Specifications for Interoperable Access and Services of NASA EOS Data", in Qu, J. et al. (Eds.). *Earth Science Satellite Remote Sensing*, Springer-Verlag. pp. 254-268.

Di, L., 2006b. "The NASA HDF-EOS Web GIS Software Suite (NWGISS)." In Qu, J. et al. (Eds.), *Earth Science Satellite Remote Sensing*. Vol. 2, Chapter 18, Springer-Verlag.

Di L., Zhao P., Han W., Wei Y. and Li X., 2007. "Web Service-based GeoBrain Online Analysis System (GeoOnAS)", *Proceedings of NASA Earth Science Technology Conference 2007. June 19-21, 2007. College Park, MD, USA. (CD-ROM).*

Domenico, B., J. Caron, E. Davis, R. Kambic and S. Nativi. 2002. "Thematic Real-time Environmental Distributed Data Services (THREDDS): Incorporating Interactive Analysis Tools into NSDL," *Journal of Digital Information*, 2 (4), http://journals.tdl.org/jodi/article/view/51/54

Evans, J. Bernhard, B., Di. L., etc., 2003. Web Coverage Service (WCS), Version 1.0.0, OGC 03-065r6. OpenGIS© Implementation Specification. Open GIS Consortium. http://ww.opengis.org/docs/03-065r6.pdf (101p).

FGDC, 2002, "The FGDC Content Standard for Digital Geospatial Metadata: Extensions for Remote Sensing Metadata", FGDC-STD-012-2002. U.S. Federal Geographic Data Committee. Reston, VA (144 p).

Fulker, D.W., S. Bates, and C. Jacobs, 1997. "Unidata: A Virtual Community Sharing Resources via Technological Infrastructure", Bulletin of the American Meteorological Society, Vol 78, March, 457-468.

ISO TC 211, 2004. "ISO 19130 Geographic Information-Sensor and Data Models for Imagery and Gridded Data". International Organization for Standardization (ISO), Technical Committee 211. (Committee Draft, Project Chair : L. Di).

Kuhn, W., 2003. "Semantic Reference Systems", *International Journal of Geographic Information Science*, 17 (5), 405-409.

Pavlovic, D., and Smith, D. R., 2001. "Composition and refinement of behavioral specifications", *Proceedings of Automated Software Engineering Conference* (2001), IEEE Computer Society Press, pp. 157-165. http://citeseer.nj.nec.com/pavlovic01composition.html

Ramachandran, R., J. Rushing, H OGC, 2003a. Geography Markup Language 3.0. OpenGIS Consortium. (Adopted Standard, Ron Lake, et. al, Editors).

Ramachandran, R., H. Conover, S. Graves, S. Christopher, 2003b, "Earth Science Markup Language: A Solution to the Earth Science Data Format Heterogeneity Problem", American Meteorological Society's (AMS) 19th International Conference on Interactive Information Processing Systems

(IIPS) for Meteorology, Oceanography, and Hydrology, Long Beach, CA, Feb. 9 – 13.

Ramachandran, R., S. Graves, H. Conover, and K. Moe, 2004, Earth Science Markup Language (ESML): a solution for scientific data-application interoperability problem, *Computers & Geosciences*, 30 (1),117-124

Raskin, R., 2003. “A semantic web for Earth and Environmental Terminology (SWEET)”, International Semantic Web Conference Workshop on Technologies for Searching and Retrieving Scientific Data, Sanibel Island, FL, Oct. 2003.

Reddy, M., LeClerc, Y. G., Iverson, L., Bletter, N., 1999. TerraVision II: Visualizing Massive Terrain Databases in VRML, IEEE Computer Graphics and Applications (Special Issue on VRML), vol. 19, no. 2, pp. 30-38, 1999. http://www.ai.sri.com/pubs/files/412.pdf

Stickel, M., Waldinger, R., Lowry, M., Pressburger, T., and Underwood, I, 1994. The deductive composition of astronomical software from subroutine libraries, in Twelfth International Conference on Automated Deduction (CADE), Nancy, France, pp. 341-355, Jun 1994. http://www.ai.sri.com/pubs/full.php?id=213

Stickel, Mark E. and Waldinger, Richard J. and Chaudhri, Vinay K., 2000. A Guide to SNARK, Technical Report. AI Center, SRI International, May 2000. http://www.ai.sri.com/pubs/full.php?id=784

Suresh, R. L. Di, L. Kumar, K. McDonald, 2001. “ DIAL - Enhanced Interoperability of Geo-spatial Data A Distributed Catalog and Data Services System for Remote Sensing Data”, *Geoinformatics*, p 36-39. Jan/Feb, 2001.