Spatial Web Portal for Building Spatial Data Infrastructure

Chaowei Phil Yang[1, 2], Ying Cao[1], John Evans[2], Menas Kafatos[1], Myra Bambacus[2]

[1]Earth Systems and GeoInformation Sciences and Center for Earth Observing and Space Research, School of Computational Sciences, George Mason University
[2]NASA GeoScience Interoperability Office, Code 604.0
E-mail: cyang3@gmu.edu

Abstract
The past decades have witnessed the rapid growth of heterogeneous geospatial information systems. An important way to share these valuable assets is a spatial data infrastructure (SDI). Recent developments in Web Services and distributed geospatial information services(Yang et al., 2005) provide a practical approach, Web Portals, to building a SDI. This article describes research, development, and challenges related to Web Portals for SDI.

Keywords
spatial Web portal, spatial data infrastructure, interoperability, Web services

I. SPATIAL DATA INFRASTRUCTURE

Geographical Information Systems (GIS) have seen widespread deployment in the public and private sectors, and have proven valuable for a variety of applications. GIS implementations range from commercial to open source software; their scope ranges from the global level to a specific community or person; and they are deployed on widely different hardware platforms. This heterogeneity often makes it difficult to find the right data or information in a usable form (Figure 1).

A Spatial Data Infrastructure (SDI) can facilitate harnessing and sharing heterogeneous GIS resources more effectively. SDI includes the technologies, standards, and policies needed to share geospatial data in the public, private, non-profit, and academic sectors. For over a decade, organizations such as U.S. Federal Geographic Data Committee (FGDC), the Open Geospatial Consortium (OGC), and the International Organization for Standardization (ISO) / Technical Committee 211 (TC211) have contributed to building such an infrastructure. ISO/TC211 defines information contents and component behavior; whereas OGC develops technical specifications for building geoprocessing software and services. FGDC works with U.S. Cabinet-level agencies to build the U.S. National SDI (NSDI) through guidance, seed funding, and coordination. As an international counterpart to FGDC, the Global SDI Association is working to build such an infrastructure globally.

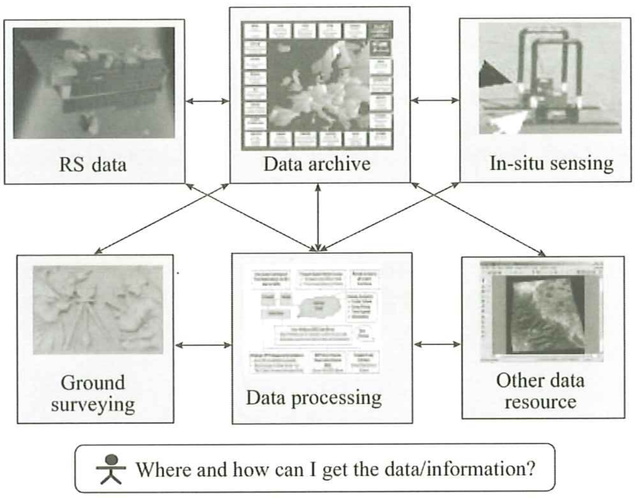

Figure 1. Different GIS datasets are collected through different techniques, kept in different forms and places, processed in different models and simulations. It's quite difficult to find the right data and information in a timely fashion when needed

Geospatial data also differ in their abstraction of real world phenomena, data-capture methods, origin and processing history, formats, and related applications. Bridging these differences is the goal of standards for data abstraction, data capture, computation, formats, and applications. For example, FGDC has identified seven fundamental data themes as "Framework" data (Cadastral, Digital Orthoimagery, Elevation, Geodetic Control, Governmental Units, Hydrography, and Transportation) and is defining Framework Content Standards in collaboration with American National Standards Institute (ANSI).

FGDC also spurs the development of networking, task forces, and other efforts to build different parts of the NSDI–so as to answer the question in Figure 1 within a heterogeneous environment of geospatial services and data. FGDC awards yearly seed grants to local, regional, national, and international SDI initiatives.

SDI also provides a service-oriented environment for spatial data discovery, evaluation, and application for a broad set of users and providers(Nebert, 2004). This environment includes a coordinated "clearinghouse" of data catalogs with links to Geospatial One-Stop (GOS), the Global Change Master Database

1082-4006/06/1201-38$5.00

(GCMD) of the National Aeronautics and Space Administration (NASA), the National Climate Data Center (NCDC) of the National Oceanic and Atmospheric Administration (NOAA), and NASA's Earth-Sun System Gateway (ESG).

These systems have different sets of functions; require different inputs; and serve different user communities. But the World Wide Web and Web Services provide ways to enable these systems as different web portals and to link these infrastructures. Web Services are Web-based enterprise applications that rely on open, XML-based standards and transport protocols. Web portals use Web Services to publish available geospatial data and processing services, and help applications find them and invoke (bind to) services or retrieve data. OGC has developed several specifications for spatial Web Services, such as Web Mapping Service (WMS) and Web Coverage Service (WCS).These spatial Web Services provide an open, extensible foundation for sharing geospatial data and services, and a practical approach to building an SDI: spatial web portals.

This paper reports research and development conducted on the spatial web portals for building an SDI. Section II introduces an example, the MAGIC Portal. Section III discusses the interoperability requirements of a spatial web portal. Section IV is dedicated to the computing support of a spatial web portal. Section V discusses some challenges in building operational spatial web portal. Section VI provides some conclusion remarks.

II. MAGIC PORTAL: A SPATIAL WEB PORTAL EXAMPLE

A Web portal generally offers a broad array of resources and services, such as search engines and on-line shopping malls. A spatial web portal provides geospatial data and services. A typical portal has three components: resources and services, a catalog search service, and an application, usually customizable for different users.

The Virginia Access-Middle Atlantic Geospatial Information Consortium (VA-MAGIC), a NASA-funded project led by George Mason University (GMU), has collected data from different sources and archived them in a widely dispersed manner. To maximize the benefit of these resources, interoperable systems need to integrate these data within a distributed environment. With FGDC and NASA support, VA-MAGIC is building a spatial web portal to archive, search, discover, and access those resources—including a set of services built for transportation, using FGDC standards and OGC specifications. Figure 2 shows the architecture of the MAGIC portal.

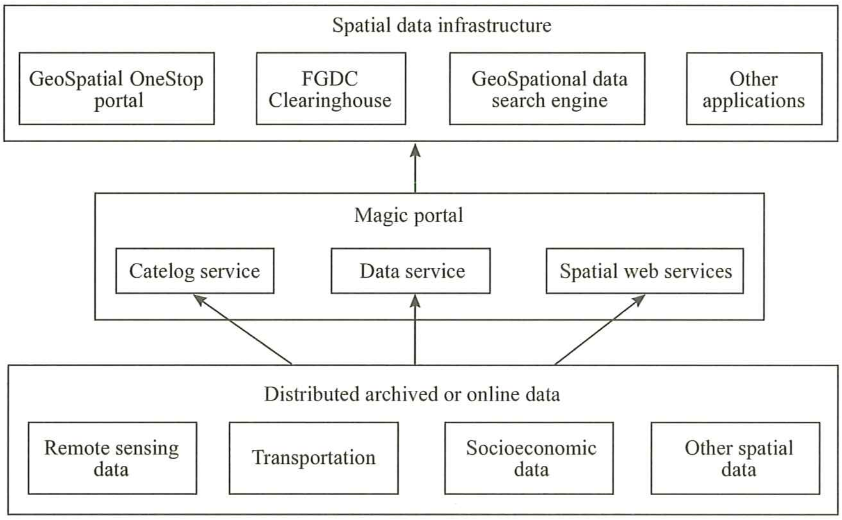

Figure 2. MAGIC portal 1) takes data services from different resources, such as satellite or airborne remote sensing data, transportation data, socioeconomic data, and other spatial data, 2) provides catalog services, and 3) other spatial web services to support different SDI applications, such as FGDC clearinghouse, GOS portal, and geospatial data search engine

The MAGIC portal is a joint effort of the Center for Earth Observing and Space Research (CEOSR) at GMU, Virginia Department of Transportation (VDOT), Bureau of Transportation Statistics at the Department of Transportation (BTS/DOT), and Intergraph Corporation. It demonstrates Transportation Framework Data Services using OGC's Web Feature Service (WFS), Web Mapping Service (WMS), and the American National Standards Institute/National Information Standards Organization ANSI/NISO Z39.50 Standard to share transportation data between BTS/DOT (federal government), VDOT (state government), and Intergraph (industry). CEOSR developed the solution using Intergraph and ESRI software and internal distributed geospatial computing research(Yang et al., 2005). The Web services developed within the MAGIC portal are integrated with operational geospatial systems at CEOSR.

VA-MAGIC data include the National Transportation Atlas

Database (NTAD); VDOT's public framework transportation data (integrated into CEOSR's operational system and also shared through WFS), and near real-time rainfall data from Unidata, (integrated to detect correlations of weather and traffic patterns). All feature data are consistent (where applicable) with the draft INCITS L1 Framework Data Standard, and available via WMS and WFS using Intergraph's GeoMedia WebMap 6.0 WFS toolkit, ESRI's ArcIMS 9.2 interoperable toolkit, and custom software.

Several different SDI clients have been developed and/or linked to the VA-MAGIC Web Portal. Figure 3 illustrates the Geospatial Data Searching Engine (GDSE), which uses Web Services and Z39.50 to search the portal. GDSE provides different types of searching methods, such as key words, geographic region, and free text. Search results include brief descriptions, links to the full FGDC metadata (FGDC Content standard for digital geospatial metadata 1998), and WMS example queries; a Web Map viewer provides pan, zoom, center, and identify functions for exploring the found data. Users may download a dataset if they determine that it supports their needs.

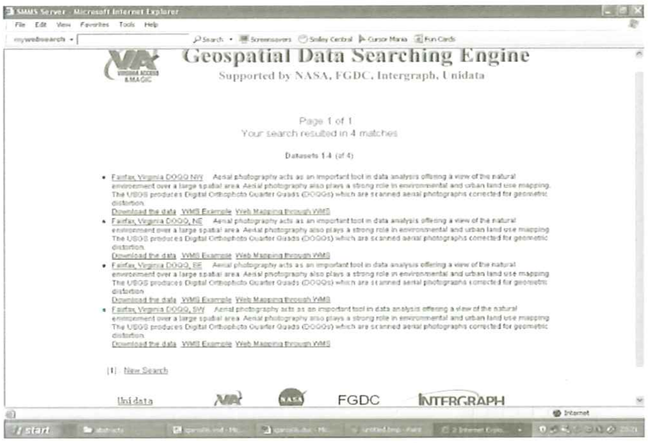

Figure 3. Geospatial data search engine interfacethe geospatial data search engine provide a place to find, review, and mapping relevant data through interoperable services such as FGDC metadata, Z39.50, and WMS

Figure 4 illustrates the VA-MAGIC portal's support to the GOS portal through Z39.50 for Geospatial Metadata and the OpenGIS Catalog Service for the Web (CS-W(Schutzberg et al., 2005)): a search through the GOS portal finds data kept in the VA-MAGIC portal and originated by the Department of Transportation through its Bureau of Transportation Statistics.

The magic portal also provides a web-based applet client prototype developed at CEOSR. The client can integrate and view the output of different WMS, WFS services in an integrated mapping interface; users can customize the map through different layer selections.

Figure 5 shows a sequence diagram depicting interactions of three different types of users with five components of the VA-

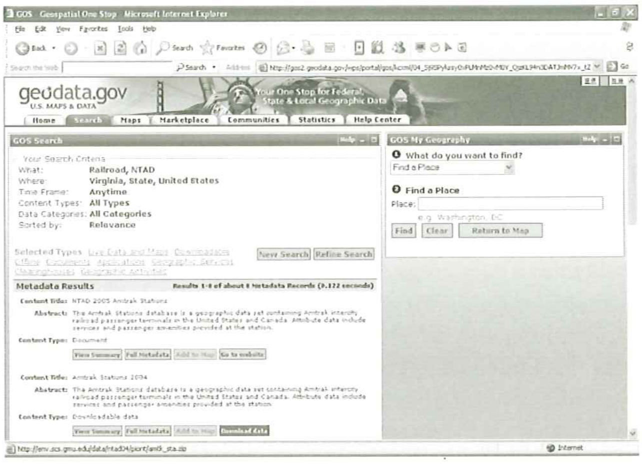

Figure 4. The Web portal supports the geospatial one stop by providing metadata which is searchable through Geospatial One-Stop portal, the figure shows finding the datasets hosted at MAGIC portal through GOS portal

MAGIC portal: (1) *clients* handle interactions with users; (2) a *Chaining Service* provides basic work flow support through predefined application logic; (3) *Catalog Services* use the FGDC metadata standards and relational database support to record meta information about data and services; (4) *Editing* allows a system developer to input service metadata into the catalog service, to configure different data, mapping, coverage, and feature services, and to configure specific work-flow based chaining services; (5) *Data and information services* provide the core system support, via interoperable interfaces to distributed services.

An advanced user can interact with the client to configure a service chain for a specific application—for example, a real time traffic-flow map built from sensor data, reprojected with roads lines, using services for map integration and visualization. Public users may invoke a predefined work flow — for example, a school teacher might display a map of world ecoregions through a predefined Web Mapping chain. Figure 6 also illustrates the sequence of function calls chained by a specific user. For example, figure 3 and 4 illustrated public user scenarios. Figure 5 illustrated an advanced user scenario. System developer operations normally occur offline because of security and system performance considerations.

The VA-MAGIC portal illustrates the use of Web Services to support SDI. Other organizations have also built portals to support SDI, such as the NOAA NCDC portal, NASA ESG, GOS portal, etc. Interoperable solutions based on standard or specifications facilitate sharing and reusing these assets so that a larger community of users can benefit.

III. INTEROPERABILITY: REUSING DATA, INFORMATION, AND OTHER SERVICES

An interoperable approach will ensure a more usable portal,

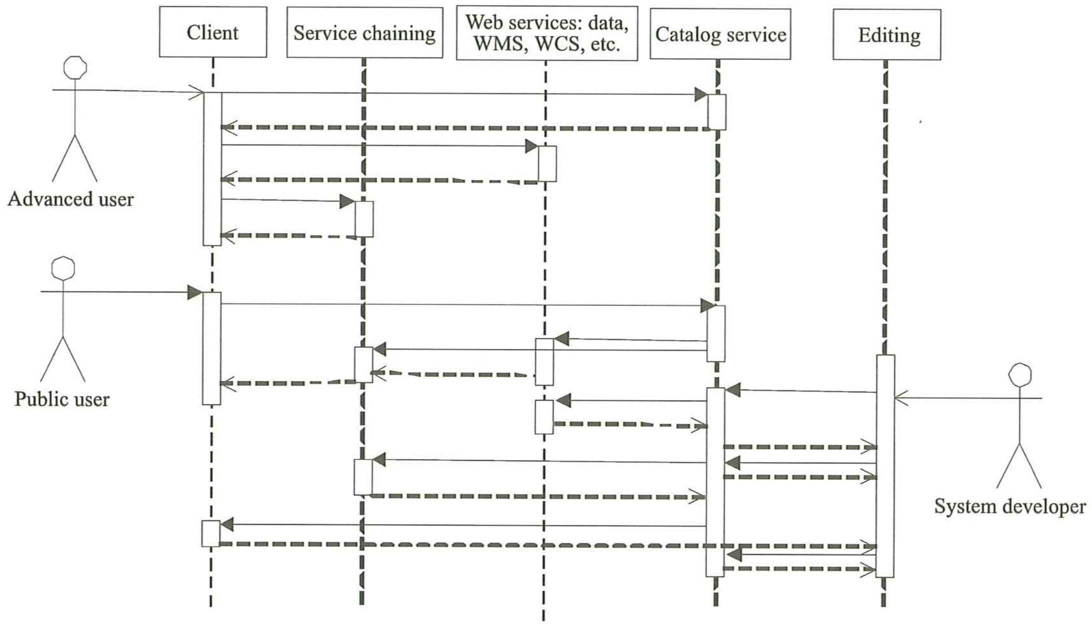

Figure 5. Different users using different approach to accomplish different tasks within a portal supporting the spatial data infrastructure: system developer can perform system configuration and application development offline, advanced user can access the SDI interactively, general public user transparent can access predefined applications transparently

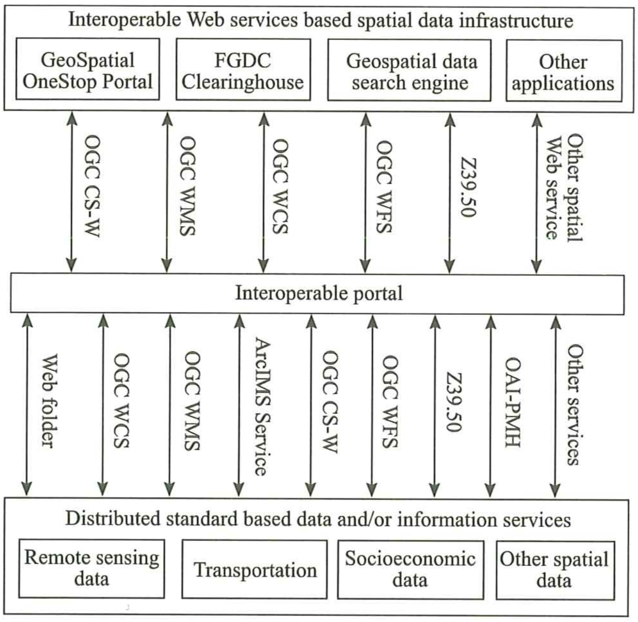

Figure 6. An interoperable architecture of a portal access different data/information services to provide interoperable services to different SDI clients that implementing the interoperable access to different interoperable spatial Web services

reusable portal components, and reduce costs in the long run for a spatial Web Portal(Bambacus et al., 2006). The interoperable solution is based on implementing standards-based component interfaces. The VA-MAGIC portal shows that a web portal supporting SDI includes 3 principal tiers: the core data, information, and application services; a set of interoperable Web services supporting the "publish, find, and bind" pattern; and a client to interact with users.

The interoperability of such a web portal relies on standards or specifications at and between those three tiers. Figure 6 shows how such a portal embeds interoperable solutions to facilitate access to heterogeneous distributed geospatial information services.

The bottom tier implements data and information services in standard-based formats or models; these services are understandable to the client tier, which takes the service input and handles interaction with users.

The interface between the middle and bottom tiers will use several different Web Services: (1) a Web folder is a folder accessible through HTTP based file downloading; (2) OGC WCS supplies images or other coverage; (3) OGC WMS provides maps and simple visualization; (4) ArcIMS service provides a "*de facto* standard" web mapping service; (5) OGC CS-W provides web catalog searching; (6) OGC WFS provides discrete features; (7) Z39.50 provides catalog querying access, and (8) Open Archives Initiative-Protocol for Metadata Harvesting (OAI-PMH) provides open access to metadata archives; (9) Other web services may also be available.

In the middle tier, the portal will provide facilitating SOA services and some transforming services— for example, a portal could combine services with different projections by reprojecting them into one projection, and integrating them as one response to the client.

Between the middle and top tiers, web services are supported through either the interoperable service provided by the bottom tier or some transformation services provided by the middle tier. These services could be WCS, WFS, WMS, CS/W, and other widely-supported web services.

The top tier provides interoperable access to the services brokered or provided through the middle tier.

This interoperable architecture is one practical approach for building interoperable spatial Web Portals. OGC also works on a Geospatial Portal Reference Architecture(Rose, 2004).

IV. SYSTEM SUPPORT: COMPUTING BASED ON THE INTERNET

Implementation of a Web portal and support for SDI is inseparable from the system support provided by computing technologies, and especially the Internet. System support from Internet computing depends on Computing and Networking, System Performance, Application logic and Service Chaining, and the Graphical User Interface (GUI).

A typical GIS process involves many large datasets and processing steps, for example, to make a map, data needs to be found, subsetted, reprojected, fused/mined, and visualized according to a specific schema. The processes can be distributed across the Internet on servers. And the Internet is responsible for the communication and transmission of information exchange between different steps in the process.

Computing and networking involves both computer's computing capacity and network's transmission capacity.

Fortunately, evolution of computing technologies provides support to this process on both computing and networking, for example, today's desktop processor is more powerful than mainframes of the 1960s and supports each step of the process. Internet bandwidth has also increased significantly in the past several years: for example, the Lambda Rail has increased backbones to a magnitude of 10 Gigabits per second (Gbps). An organization's internal connection can easily have a 1 Gbps connection; and 100Mbps is the norm for desktop and laptop connectivity.

Other developments, such as web sensors and ubiquitous computing, have enabled ubiquitous, on-demand implementation and integration of different components, especially the data capturing and client functions, of an SDI.

System performance is another concern of a widely adopted SDI because of the networking environment and the complexity of process within the SDI. A number of researchers (Tu et al., 2001) have addressed the performance issues within a network GIS from a specific or systematic point of view, and provide a set of valuable results on improving performance. According to the structure of a SDI and the integration of different service, a normal approach for improving performance is to develop and integrate an approach with multi-threading, dynamic caching, and binary streaming to achieve a better performance(Yang et al.).

Commercial products, such as Google Earth and TerraServer, provide good examples of some of these performance techniques on rendering images and transforming predefined information.

Application logic and service chaining is another important issue for system support. Specific application logic may require a specific set of geospatial information services and a specific sequence to integrate these services through service chaining process(Alameh, 2003). Chaining different services to form an application allows information flow from the data acquisition to a specific application. As illustrated in Figure 5, a user who understands the application and the service chaining process may build this service chain interactively–and save it for future access by more casual users, or as a predefined application for a specific group of users.

The **Graphical User Interface** (GUI) is another system support needed for Web Portals and SDI. As seen in Figure 3, there are clients for data discovery, data publishing, data viewing, and data integration. How many functions can be integrated into the GUI is a balance on both functionality and complexity. Component-based software development provides a convenient way to develop a flexible client GUI, or to integrate an existing GUI.

V. CHALLENGES: THE FUTURE EVOLUTION OF SPATIAL WEB PORTAL AND SDI

System performance within Web Services environment may be a significant issue(Tu et al., 2004) when using standards-based processing, interfacing, service transformations: for example, encoding information as text or GML may significantly increase system response times. Thus, interoperable services such as WMS, WCS, and WFS may need special consideration on their performance implications for data access and processing. For example, a pyramid index method can be used for images to serve WMS; WFS can use a binary encoding of GML. Also needed is systematic research for integrating an application through the workflow chaining service. Performance under load is also a concern requiring collaborative research between data /information specialists and computer scientists.

Knowledge-based automatic service chaining: Service chaining provides a flexible process for building applications. However, system designers cannot predefine all applications before users access them from the GUI. Thus, chaining can be problematic without an expert to specify the chaining process. If the chaining can be automatically performed based on previous knowledge gained, the process may be more flexible and more fully interoperable. Ideally, a web portal supported SDI could assemble a service chain dynamically to solve any problem submitted by users; but this will require extensive work in knowledge management and discovery.

There are several challenges: (1) semantic or ontology processing to understand or translate key terms and concepts; (2) a "credit alike" for rating different services(akin to recommendations on shopping sites like Amazon.com); (3) a learning system to detect and mimic patterns in user-defined service chains.

Quality control/relevance grading: The quality of service (QoS) (performance, accuracy, reliability, etc.) of SDI and spatial web portals is also a crucial success factor. It may arise from user evaluations (such as book ratings and recommendation at amazon.com). Another mechanism is to provide rules based on the real service content to rate the QoS. This mechanism requires research on spatial data patterns, data mining, automatic chaining, or information science within each specific application domain.

Security, economic, and other challenges also exist as we target a successful SDI through Web Portals. Security problems frequently arise; for example, the need to protect sensitive data from unauthorized users at several levels of granularity—maps, map layers, features, or even certain coordinates or feature attributes.

Security concerns may also arise for a specific service; for example, a pattern recognition service may need protection from unauthorized use. From a systematic point of view, the open, interoperable solution must not circumvent the security of a computer system or communications network.

Such a spatial web portal supported SDI may also change business practices in the geospatial world. For example, instead of purchasing full-fledge software product suites, we can make a usable image via the Internet by integrating data services from USGS, computing services from Intel, integration services from Intergraph, and a visualization service to construct an image. One concern that arises in this process is how to manage the paying and credit system.

Other issues include social concerns from different communities, for example, some country may have restrictions on data that is open to the public in another country.

VI. CONCLUDING REMARKS

This paper reports ideas and practices in building spatial data infrastructure at George Mason University and NASA through spatial web portals. We expect that the reported ideas and experience may help in (1) building different applications and community web portals to support SDI using interoperable solutions; (2) guiding future research on building and operating SDI; (3) serving as a reference for developing new spatial web portals, such as ESIP Federation Portal and Transportation geospatial web portal; (4) adding to the knowledge base of SDI activities.

ACKNOWLEDGMENTS

Ideas, research, and development reported are sponsored by FGDC CAP project 05HQAG0115, NASA VA-MAGIC project NNS04AB26G, and Intergraph (through 2002-2006 Registered Research Laboratory).

REFERENCES

[1] Alameh N., 2003, Chaining Geographic Information Web Services, IEEE Internet Computing, 6(18): 22–29.

[2] Bambacus, Reichardt, 2006, Invest on interoperability, Geospatial, Februrary, pp. 26–30.

[3] Nebert D., 2004, Global Spatial Data Infrastructure Cookbook, Version 2, pp.1–3.

[4] Rose L., 2004, OGC Geospatial Portal Reference Architecture, https://portal.opengeospatial.org/files/?artifact_id=666

[5] Schutzberg A., Francica J., 2005, The Technology Behind the New Geodata.gov and the Non-Technology Challenges Ahead, http://www.directionsmag.com/article.php?article_id=784&trv=1

[6] Tu S., Flanagin M., Wu Y., Abdelguerfi M., Normand E., Mahadevan V., Ratcliff J., Shaw K., 2004, Design Strategies to Improve Performance of GIS Web Services. ITCC (2): 444–450.

[7] Tu S., He X., Li X., Ratcliff J., 2001, A Systematic Approach to Reduction of User-Perceived Response Time for GIS Web Services. ACM-GIS, pp. 47–52.

[8] Yang C., Tao V., 2005, Chap. 5: Distributed Geospatial Information Service, *Frontiers of Geographic Information Technology*, Edited by Sanjay Rana and Jayant Sharma, London: Springer, pp. 113–130.

[9] Yang C., Wong D., Li B., 2005, Introduction to computing and computational issues of distributed GIS, *Geographic Information Sciences,* 11(1):1–3.

[10] Yang C., Wong D., Yang R., Kafatos M., Li Q., WebGIS performance improving techniques, IJGIS (19): 319–342.