ESSENTIAL REQUIREMENTS FOR DEVELOPING CORE GEOSPATIAL WEB SERVICES FOR NSDI

Jung-Hong Hong[1], Yi-Ling Huang[2] and Min-Lang Huang[3]
[1] Associate Professor, Dept. of Geomatics, National Cheng Kung University,
701, No.1, University Road, Tainan City, Taiwan; +886-6-2370876#851
E-mail:junghong@mail.ncku.edu.tw

[2] Master Student, Dept. of Geomatics, National Cheng Kung University,
701, No.1, University Road, Tainan City, Taiwan; +886-6-2370876#851
E-mail:a1001ritahi@yahoo.com.tw

[3] Ph.D Candidate, Dept. of Geomatics, National Cheng Kung University,
701, No.1, University Road, Tainan City, Taiwan; +886-6-2370876#851
E-mail: yoshi.huang@dprc.ncku.edu.tw

KEY WORDS: geospatial web service, NSDI

ABSTRACT: The importance of geospatial information has been widely recognized in current internet-based applications. The development of a National Spatial Data Infrastructure (NSDI) can take full advantage of the service-based technology to easily facilitate the sharing of geospatial data between different organizations. Although many geospatial web services have been successfully built, very few of them have taken the correct use of data in later applications into consideration. Consequently end users may naively misinterpret the data they acquired and make wrong decisions without noticing possible discrepancy. Especially for applications that depend on the real-time acquisition of geospatial data, the risk may tremendously increase with the increasing of the variety and volume of data. We argue that it is necessary to include the knowledge about data in the development of core geospatial web services of NSDI, such that the distributed data will be automatically possessed of built-in information to unambiguously identify itself and allow clients' applications to correctly analyze the correspondence or difference between heterogeneous data collected from various resources. Two strategies for developing such a service-based framework are proposed in this paper, the first is to analyze and identify the essential requirements for geospatial web services in NSDI, and the second is to design a standardized framework that ensures the necessary information can be correctly conveyed according to the types of services. The merits of NSDI come from the successful and transparent integrated use of geospatial data. However, only the "correct' use of data can really ensure the making of correct decision. Every participating organization therefore must understand its roles in NSDI and contribute necessary information for making the success of NSDI.

1. INTRODUCTION

With the booming developments of information technologies in recent years, acquiring various types of data meeting particular constraints becomes much easier than before. Internet-based search engines enable users to easily accomplish this task by only specifying keywords he or she has interests. Despite the convenience of obtaining data, the returned information is often not well organized and requires users to have in-depth knowledge to correctly interpret the context. Interoperability of heterogeneous data collected from different resources is therefore always regarded as a challenge to all levels of users. For the better use of geospatial information, the concept of Spatial Data Infrastructure (SDI) is introduced for building a standardized framework that can facilitate the use and distribution of geospatial data between different organizations. Many nations around the world use SDI as an integrated environment to connect the data producers, data providers, value-adders and data users together (Masser et al., 2008). The development of a SDI is of course a very complicated issue that not only demands consensus agreement among participants, but also innovated breakthroughs on geospatial technology. A SDI consists of five components: data, technology, standard, policy and institutional arrangement, and people (Mansourian et al., 2006; Nebert, 2009). In the view of Mansourian et al. (2006), both the data production and service development must follow standardized specifications to resolve the issues of interoperability. Such specifications are a collection of consensus agreement among different institutions and professionals. Policies are necessary in the development of SDI to enforce the development of the consensus framework. Access network serves as the channel for distributing data between organizations. Following Figure 1, the geospatial data is encoded and distributed following standards via internet enforced by the policy. All of these works are made possible by the participants (people) of SDI. The integration of these components is the key to the success of SDI.

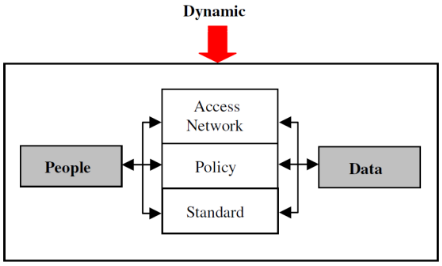


Figure 1 Relationship among the elements (Mansourian et al., 2006)

Web services are developed to bridge the communication between computer systems. According to OASIS's definition, "services are the mechanism by which needs and capabilities are bough together" (OASIS, 2006). A service-oriented architecture (SOA) can then be regarded as the working of web services for utilizing and organizing the distributed service resources. Within a SOA, the accesses to services are maintained with consistent constraints and policies, the users, however, don't really possess the knowledge about how the implementations are done or the essential information in evaluating for further applications if users do not send request (OASIS, 2006).

With the tremendous potential for SDI to easily facilitate the search and delivery of information from one organization to another, to access information that users have no prior knowledge would be very common in the future. Although the emergence of SDI indeed removes many obstacles of interoperability, easier availability of geospatial information does not guarantee the correct interpretation and use of the acquired data. We argue that more consideration must be added to enhance the current role of geospatial web services. Section 2 first discusses the core geospatial web services in SDI and proposes our categorization. Section 3 focuses on the analysis of the essential requirements according to the categorized core geospatial web services. Key factors for practicing are also included. In section 4, a discussion based on different practicing instances with a conceptual demonstration will be presented. Finally, in section 5 will be a short summary of this paper.

2. CORE GEOSPATIAL WEB SERVICES

Geospatial web services play an important role in the success of spatial data infrastructure, especially for improving the interoperability between different computer systems. A good practice for developing a national spatial data infrastructure (NSDI) is to deploy different types of core geospatial web services in an integrated fashion, such that users' various demands can be either met by a single service or the chaining of a number of services. This will be required to identify the categorization of geospatial web services. Table 1 summarizes the review of different viewpoints from the actual implementation of SDI and international standardization organizations. Although the terms used are not exactly the same, it appears that the data discovery service and data access service are common services for developing a SDI. This is not quite surprising as this is exactly how we search and access information in internet. As to the functionalities of both, they are regarded as fundamental service types within the working of SDI (Friis-Christensen, A. et al., 2007). The more functions a SDI is designed to offer, the more types of web services are needed. Each type of web service would need a consensus agreement on the standardized web interface to facilitate the communication between systems, so standards must be developed to meet such demands.

Table 1 Developments of SDI around the world

Nation	NSDI	Managed institution	Service types
United States of America*	US NSDI	US FGDC	1. Data discovery 2. Data visualization 3. Data access
United Kingdom**	UK NGDF	UK AGI	1. Planned to be compliant with INSPIRE
European Union***	INSPIRE ESDI	European Commission	1. Registry 2. Discovery 3. View 4. Download 5. Transformation 6. Invoke Spatial
International Organization for Standardization****			1. Human interaction service 2. Model/information management service 3. Workflow/task service 4. Processing service 5. Communication service 6. System management service

Nation	NSDI	Managed institution	Service types
Open Geospatial Consortium Inc.******			1. Application 2. Registry 3. Encodings 4. Data 5. Portrayal 6. Processing

Note：the content is the analyzed results of *FGDC(2010), ** Boulos(2004),*** Granell et al.(2010),**** Percivall(2002) and *****OGC

The various categorizations on the types of geospatial web services make the development much more complicated, as every type of service may has its own distinguished characteristic. Four types of geospatial web services are summarized from Table 1, namely, catalogue service, data service, portrayal/demonstration service and processing/functional service. Each type of service is respectively developed to denote one major function in the SDI.

Catalogue service
This type of service is developed for the registration of distributed geospatial resources, discovery of web service that may be useful to users and the supply of brief information or hyperlinks about the georesources found. It serves as a channel for users to reach all the geospatial web services available in the SDI.

Data service
This type of service allows authorized organizations to publish data via an open web service, such that remote users may easily acquire the data via internet. As the original data may be recorded in a variety of formats, the distributed data is often encoded in open and standardized format, e.g., GML and KML, to remove barriers caused by data heterogeneity.

Portrayal/demonstration service
This type of service is mainly used for the illustration of the content of geospatial data upon users' requests. Remote users are allowed to visually inspect the data located at remote servers.

Processing/functional service
This type of service is mainly used for modifying, transforming, processing and analyzing input data and responds with the desired content of geospatial data according to users' requirements. Ideally speaking, the way it works should be similar to how we use the functions offered by desktop GIS software. But as each individual service may have its own consideration and designed interface, users should be very cautious when dealing with a processing/functional service.

For developing an application in the SDI, all of the above four types of services have their distinguished roles. After the requirements of users' application are determined, catalogue service helps to locate geospatial web service that is helpful, the returned results may include all the other three types of services. Users can access required data from remote data services. In users' application environment, some data may be illustrated via portrayal service to simplify users' task and reduce the loading of data processing. If either users' data or data collected from data services does not exactly match the needs for operation, users can turn to processing/functional service to meet their application requirements. An ideal SDI must be able to meet the variety of demands from different domains of users. These users, however, may be also contributors to the SDI. Technically speaking, the standardization of these four types of geospatial web services has already established a transparent data sharing environment in the SDI. The correct use of and decision making on the data acquired from different georesources, however, requires more in-depth knowledge about the service used in the application. Users must be able to determine if the used service is a "trusted" service, or by the detailed extent of the service document (Granell, Dı´az et al., 2010). As this information can only be offered by its authorized organizations, we argue that more requirements must be added to ensure the correct use of data dynamically collected and processed by geospatial web services in the SDI.

3. ESSENTIAL REQUIREMENT FOR GEOSPATIAL WEB SERVICE
As each type of geospatial web services plays a particular role in the SDI, we argue that the essential requirements for fulfilling such distinguished characteristics must be considered beforehand and enforced by the respective components during the development of SDI. The following discusses the essential requirements for each type of geospatial web services in section 2:

Catalogue service

The major purpose of catalogue service is to bridge the communication between data providers and users, so the most essential requirement will be to ensure “all” the available georesources are registered in a standardized way. This would require the adoption of CSW and rigorous policies from the authorized organization of SDI. All participants must submit their registered information to the SDI that precisely describes the purpose, interface, standards and even the data of the geospatial web services. After the standardized description is effectively managed by the SDI, users can issue queries to the SDI to search qualified geospatial web services and determine their fitness of use according to the descriptive information. Despite the advantages of standardization, most of the current systems only offer limited search capabilities like “where”, “what” and “when”. To meet the searching from the continuously growing volume of data, more search constraints like data quality should be taken into consideration. Furthermore, the ranking of returned results is also helpful for users to select the best available georesources, which would require more insights about the knowledge of ranking behaviors.

Data service

To improve the data interoperability at clients’ application, both the data content and the service interface must be provided in an open way. Most of the current SDI implementations are based on developing a framework of data content standards and using OGC standardized web services like WFS. Data content standards allow a specific domain of geospatial data to be distributed and parsed in a standardized and consistent way. The use of standardized web interface simplifies the heterogeneous interfaces provided by the participants. As different themes of geospatial data are administered by different organizations, an ideal SDI must identify common data services widely demanded by participants in the SDI and ask their authorized organizations to publish the data in an OpenGIS way. The application schema design of the distributed data must include consideration of identification, spatial, temporal, application-dependent attributes, and most importantly, the data quality about the data. Since the handling of unfamiliar datasets acquired from diverse resources is very common in SDI, data quality information, e.g., completeness, logical consistency, spatial accuracy, temporal accuracy and thematic accuracy, help users to distinguish the differences between selected datasets and avoid wrong decision making. All of the distributed information must be encoded in a standardized way for the clients’ program to unambiguously and consistently parse the required information.

Portrayal/demonstration service

Being developed to meet the visual demands, the essential requirement of a portrayal/demonstration service is to vividly describe the original data behind the service and how the selected data is portrayed. The use of open portrayal service reduces the loading of clients, but it on the other hand may also hide the differences and discrepancies of the selected datasets. As more than one dataset may be illustrated simultaneously via portrayal service, the description must also be encoded in a standardized format to allow clients’ program to correctly parse the required essential information. Furthermore, the portrayal specification the service follows determines the outcomes of the illustrated results, so the semantic meaning of map symbols must be also provided to clients for action reference. An ideal clients’ application program must be able to correctly identify the differences of the selected datasets in a standardized way.

Processing/functional service

The most distinguished characteristic of a processing service is the status of the original data is changed to a new status, so the most essential requirement is to provide useful reference regarding how data changes. Unlike WFS and WMS, there can be varieties of processing services following the guidelines of the WPS. These services may have their own lists of operations and the differences to the outcomes for each service may be totally different. Although the comparison result appears to be domain dependent, the data quality of the output data can still be recorded in a standardized way. This implies that whenever we develop a processing service, we need to emphasize more about the difference the service will bring. For example, a Web Coordinate Transformation Service can change the coordinates of a selected dataset from one reference system to another, but the positional accuracy of the features after transformation may be totally different from their original positional accuracy simply because the service only offers an approximated transformation. As users may naively ignore how a processing service is designed, the lack of reference information certainly endangers how geospatial data in SDI is integrated and used.

4. INTEGRATED SHARING AND APPLICATION ENVIRONMENT BASED ON SERVICE TECHNOLOGY

Two test cases based on geospatial web services are discussed in this section. The first case is a simulated data service for river water quality monitoring data. The original data is acquired from the Environmental Protection Administration (EPA), we published the data via Web Feature Service (WFS) using SnowFlake’s “Go Publisher WFS” software. Figure 2 shows the data services connection dialog we developed in ESRI ArcMap 9.3.1. Users input URL of WFS and click “Get Data” button to request the data. In this use case, the service returns the monitoring data of water quality sites within the specified bounding box. The clients’ function automatically parses the essential information of completeness, temporal status, coordinate system and positional accuracy, and displayed the results in the map interface. After the connection to a data service is established, it works like a data layer in the local machine. Some of the functions, e.g., coordinate conversion or transformation, are even automatically executed to reduce the complexity of users' application environment. Such convenience, however, does not imply the differences of the data contents is automatically handled and correctly shown to the end users. The essential requirement we proposed for the data service thus provides basic but important information of returned data set from data services. Only with standardized framework can the application program automatically parse the essential information.

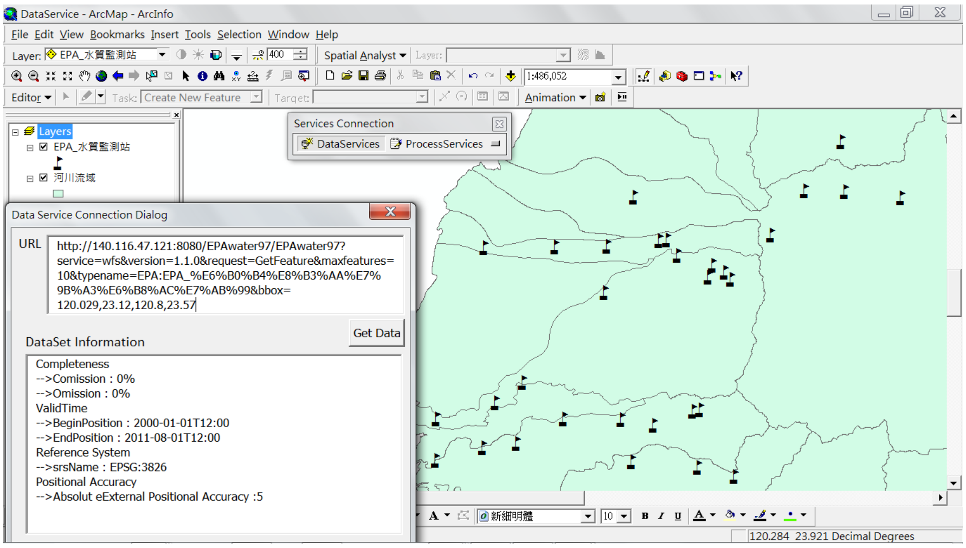


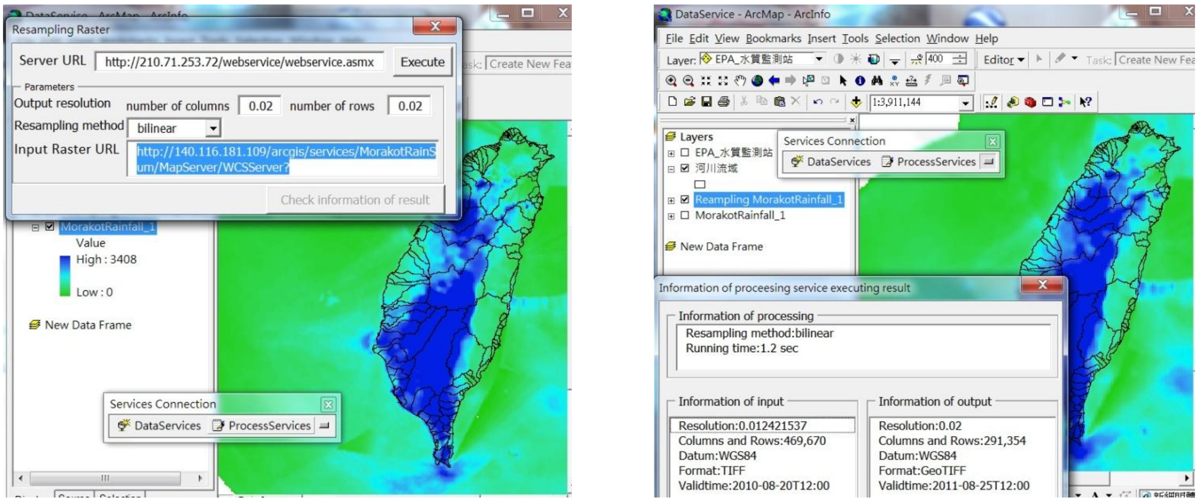
Figure 2 The demonstration of service document of data service type

The second test case is about the use of processing/functional service for resampling a raster image. The essential requirement of such a service is the ability to vividly explain the difference before and after the operation. The raster resampling web service is developed with GDAL (Geospatial Data Abstraction Library) and Visual Studio .NET 2008. GDAL is a translator library for raster-based geospatial data released under an X/MIT style Open Source license by the Open Source Geospatial Foundation. In this service we use "gdalwarp" utility of GDAL to request the resampling operation of raster data. Figure 3 shows the input dialog of raster resampling service we developed. The input raster data is directly accessed from OGC web coverage service by URL and the output data after resampling is also returned in URL of WCS. For the completion of this service, other parameters include the width (number of columns) and the height (number of rows) of the raster data. The essential information about the proposed WPS is recorded in a XML file and sent to users together with the resampling results.

Figure 3 The demonstration of raster resampling service

The client application is developed with ESRI ArcMap 9.3.1. In addition to the resampling request, the resolution differences between the input and output data is also provided after the resampling operation is accomplished. Figure 4a shows the parameter setting dialog for resampling raster function. Users can set the URL of the processing services, resolution of output, resampling method and the URL of the input raster data. The accumulated rainfall during the typhoon MORAKOT in 2009 is chosen as our input raster data. The rainfall data was published via WCS by Disaster Prevention and Research Center (DPRC), the National Cheng Kung University (NCKU) in Taiwan. After all parameters are specified, users press "Execute" button on the dialog to issue the processing requests to the services. After data is returned, users can click the "Check information of result" button to open "information checking dialog" (Figure 4b) to inspect the information about the processing, input data and output data after resampling process.

Figure 4 shows the process of using the bilinear resampling method and it takes 1.2 second to complete the resampling process. Compared with the original resolution of 0.01241537 degree, the resolution of the output raster data is changed to 0.02 degree.



(a) Parameter setting dialog (b) Information checking dialog

Figure 4 The client application of raster resampling service

5. CONCLUSION

The development of spatial data infrastructure can facilitate easier data sharing between organizations and expand the possibility of innovated applications via internet. While current SDI technology mainly focuses on the successful search, distribution and parsing of geospatial datasets in a standardized way, we argue that the development of SDI must not only consider the choices of core geospatial web services, but also the distinguished characteristics according to the roles of geospatial web services play in the SDI. Based on the proposed categorization of geospatial web services, the essential requirements for each type of geospatial web services are discussed to improve the interoperability of distributed datasets and ensure the through control of available georesources. The results indicate that all five components of SDI must be introduced to enhance the current service-based working environment. The success of a SDI will be based upon the seamless integration of geospatial web services and only geospatial web services that meet the essential requirements can ensure the correct use and dynamic integration of distributed geospatial data.

6. REFERENCE

Boulos, M. N. K., 2004. Towards evidence-based, GIS-driven national spatial health information infrastructure and surveillance services in the United Kingdom. International Journal of Health Geographics, 3(1)

FGDC, 2010. Introduction to Geospatial Web Services, NSDI training porgram.[onlie resources] http://www.fgdc.gov/training/nsdi-training-program/materials/GeoWebservices_Intro_20100604.ppt

Friis-Christensen, A., N. Ostländer, et al., 2007. Designing Service Architectures for Distributed Geoprocessing: Challenges and Future Directions. Transactions in GIS, 11(6), pp.799-818

Granell, C., L. Dı´az, et al., 2010. Service-oriented applications for environmental models: Reusable geospatial services. Environmental Modelling & Software, 25(2), pp.182-198

ISO, 2003. ISO 19112 : Geographic information — Spatial referencing by geographic identifiers. First edition

Mansourian, A., A. Rajabifard A., et al., 2006. Using SDI and web-based system to facilitate disaster management. Computers & Geosciences, 32(3), pp.303-315

Masser, I., A. Rajabifard, et al., 2008. Spatially enabling governments through SDI implementation, International Journal of Geographical Information Science 22(1), pp. 5-20.

Nebert, D. D. et al., 2009, Spatial Data Infrastructure Cookbook 2009

OASIS, 2006. Reference Model for Service Oriented Architecture 1.0

OGC. Open GeoSpatial Consortium Standards [on-line resources], Retrieved 2011/08/05, from http://live.osgeo.org/en/standards/standards.html

Percivall, G., 2002. ISO 19119 and OGC Service architecture. FIG XXII International Congress. Washington DC, USA

Open Source Geospatial Foundation, 2011, GDAL [on-line resources], http://www.gdal.org/index.html