Development of Spatial Data Infrastructure in the Mongolian Academy of Sciences

Igor V. Bychkov, Alexey O. Shigarov, Viacheslav V. Paramonov, Gennadiy M. Ruzhnikov, Alexey E. Hmelnov, Roman K. Fedorov, Alexey S. Burlakov

Institute for System Dynamics and Control Theory of SB RAS, 134 Lermontov st., 664033, Irkutsk, Russia

Abstract. The paper describes the background and prospects of the development of the spatial data infrastructure (SDI) in the scientific institutions of the Mongolian Academy of Sciences (MAS). This paper also discusses the integration of information resources for the development of SDI. Some original technological solutions are proposed for building SDI in the MAS based on international standards.

Introduction

The spatial data infrastructure (SDI) has been actively formed worldwide during the last two decades. It consists of a system of basic spatial data, metadata, organisational structures, legal regulation mechanisms, methodical framework, methods and means. It also provides users with easy access to an effective use of spatial data. Such concepts and terms are discussed in detail in the works of leading specialists [1-5] in the field of SDI.

Today, the complex use of spatial data in the Mongolian Academy of Sciences (MAS) is limited by a number of complex issues. These issues include the low quality of basic spatial data, the absence of searching and catalogization mechanisms, and the inconsistency of format and representation standards. The forming of the SDI may be considered as a challenge solution guide that addresses the effective use of spatial data in the scientific research of the MAS. Mongolian national SDI development issues attract the interest of governments. The SDI of the MAS would probably become a part of the Mongolian national SDI. This will allow for the creation of an integrated in-formational environment for spatial data storage and processing.

The background technologies that form the SDI were created by the collaborative work of the Siberian Branch of the Russian Academy of Sciences (SB RAS) and the MAS. This collaboration is directed towards the development of informational resources for the MAS scientific institutes. Particularly, the spatial data of the MAS were analysed to discover the capability and actuality of the SDI formation. The MAS scientific institutes' informational resources integration system was created by the use of open source software, such as GeoNetwork (the system to manage metadata and to catalogue the metadata, grouping them thematically and spatially) and GeoServer (a server for spatial data storage and processing). This system uses up-to-date international representations and metadata exchange standards. Moreover, a set of unique technologies was developed in the Institute for System Dynamics and Control Theory of the SB RAS to create the Geoportal and geoprocessing services of the MAS. Thus, essential elements of SDI are already prepared to facilitate scientific works in the MAS institutes and correspond to up-to-date standards. They provide easy worldwide access to the scientific spatial information resources of the MAS.

The SDI is being developed in correspondence with the MAS information resources author's project. This project includes the development of the database system, visualization and geoprocessing services, development of Geoportal, formation of basic spatial data and metadata for the scientific institutes of the MAS.

Currently, there are unique problem-oriented and domain-oriented information resources formed in the MAS scientific institutes and Mongolian universities. These resources are related to different

scientific, educational and production fields. Many of them have spatial features and include spatial data.

The tendency of Mongolian scientists to participate in international scientific projects leads to a need to develop existing informational resources of the MAS. Particularly, this development implies the creation of digital representations of spatial information resources, need for remote access and creation of metadata catalogues.

Related Projects

The creation of integrated informational resources for spatial data processing is attainable. The most popular re-sources have a three-tier architecture and are realised as web-resources. Some of them form infrastructures. Re-source of this type are often named Geoportals. This type of resource is widely distributed throughout the world.

Examples of geoportals in the:

- United States of America — http://geo.data.gov
- Canada — http://geoconnections.nrcan.gc.ca
- New Zealand — http://geodata.govt.nz
- European Union — http://epp.eurostat.ec.europa.eu

The entire set of free access resources of this type is divided into two categories: cartographic services for information search and thematic-oriented services.

Most of the free-access geospatial data informational systems are cartographic services, such as:

- http://www.fgdc.gov/dataandservices — geoportal of the Federal Geographic Data Committee for the USA. A user is able to find shared geographic data, maps, and some online spatial data processing services.
- http://www.geosrbija.rs — (in Serbian; an English version is accessible) Serbian government geoportal. It provides access to discovery and view services for selected metadata, spatial data sets, and services via the Internet for both professional users and the public.
- https://data.linz.govt.nz — Land information New Zealand geoportal. It provides open access to land and seabed data for use, reuse and sharing.
- http://map.2gis.ru — (in Russian) 2GIS is regularly used for personal and business purposes in over 200 cities of Russia, Ukraine, Kazakhstan and Italy. Users can obtain information about geographical objects in city areas, build routes between objects and calculate distances [6].

Information Resources

Currently, because of localization and an absence of remote access, informational and spatial resources of the MAS cannot always be used in scientific research, educational processes and production. It is proposed to perform the integration of these resources using the modern international metadata standard. This will make the resources more open and available.

1.1 Forming Spatial Data and Metadata

The MAS has unique information resources with spatial features [7], such as a Mongolian insects catalogue (7800 records) with their habitat (9000 records), catalogue of birds (4000 records), perennial herbarium (75000 items), weeds (10000 items), vegetation of marshes (3500 items), mosses (10000 items), fungi (2500 items), folk song collection (600 items) and artwork (300 items), formed over many years of research at the MAS. There are also the following spatial data formed at the MAS: the map of forests, soils, vegetation, pasturage and hayfields, and Landstat and SPOT satellite data. The main portion of the spatial data was created in the institutes of Botany, Geo-ecology, Geography, Biology, Informatics, and Geology. These resources may be represented in the

SDI as thematic spatial data. The basic spatial data may be formed by the Administration of Land Affairs, Geodesy and Cartography of Mongolia.

Today, in the MAS, the basic spatial data are presented as maps of aimag boundaries, ecosystems, fodder, geology, groundwater, hayland, landscapes, landuse, permafrost, soil, vegetation, and so on. In addition, under the development of SDI, a database containing statistical data of the spatially referenced socio-economic state of Mongolia is formed. Now, the database contains over 15000 entries in over 40 categories. These resources are information about population, crude birth and death rates, income, and so on. These statistics are grouped by aimags and years. The database can be used for the rapid construction of thematic maps.

Figure 1: A screenshot of the MAS metadata catalogue with web interface

The collaborative work of the SB RAS and the MAS produced a metadata catalogue based on two pieces of free software, the "GeoNetwork" metadata catalogue management system (http://geonetwork-opensource.org) and "GeoServer" spatial data management system (http://geoserver.org) (see Fig. 3). Information resource cataloguing is used in the created system. This cataloguing is based on the modern international metadata standards: Dublin Core [8], FGDC-STD-001-1998 [9], ISO 19115/19139 [10, 11]. The metadata of 155 informational resources include information describing the map of forests (38 records), soil and vegetation (39 records), and Landsat and SPOT satellite data (12 records) of Mongolia. These metadata are stored in the catalogue.
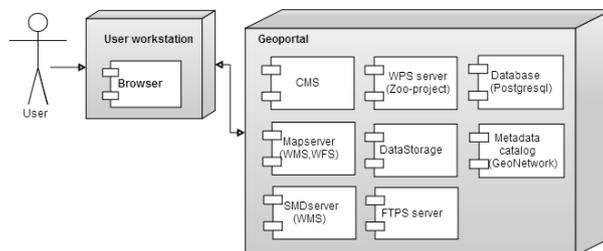

Figure 2: Architecture of the Geoportal of the MAS

1.2 Geoportal of the MAS

The created system of the MAS information resource is the main component of the SDI MAS to form and man-age metadata. Interactive maps based on the Open Geospatial Consortium (OGC) Web Map Service (WMS) and Styled Layer Descriptor (SLD) standards are formed for spatial data described in the catalogue. This process includes a deployment of spatial data on a geodata server, a formatting of visualization styles and a binding of the resources to their metadata in a catalogue.

This Geoportal is based on OGS OpenGIS open international standards, and it has a client-server architecture (see Fig. 2). Its main goals are to provide the remote control of spatial data, remote control of metadata, geoprocessing services, and publishing and visualization of spatial data. The main user side component is a web browser. The main server components are discussed more precisely below.

A content management system (CMS) is based on the open-source content management system, Calypso (http://calip.so). It manages user roles, services and spatial data. Reliable user data storage is ensured by the data storage system. The Geoportal's file manager provides access to the data storage system, including spatial data uploading function. Herewith, the Geoportal allows for defining the access rules. An FTPS (File Transfer Protocol + SSL) server provides access to the CMS file system and simplifies up- and downloading large amounts of data. CMS manages the FTPS server. The PostgreSQL database management system (with an extension for PostGIS) is designed to store user data sets and Geoportal service information. The geoprocessing services' management system registers and executes services. These services are separate program units performing different functions of geoprocessing. These services also provide an interface complying with the OGS OpenGIS Web Processing Service (WPS) standard. The Geoportal enables access to registered services, where the input data may consist of strings or files and the output data (geoprocessing results) are saved in a storage database system. The MapServer is used as a geodata server to show user maps.

A service for spatial data visualization, a "SMD Server", developed by Geoportal's authors is used in the formation of SDI. In visualization, the SMD Server, which supports the WMS standard, should be use. In this service, an original vector data format named SMD (Static Map Data) is used, which saves time in massive data visualization. Spatial indices and object lists are sorted by range scales. Data structure compilation is used in this format to store static information more effectively. Spatial data must be changed or prepared by the third party software. There are many programs that convert information to the SMD format from the following formats: SXF/RSC (Panorama), APR/Shape (ArcGIS), and WOR/TAB (MapInfo). The Geoportal allows not only metric and semantic data importing but also importing visualization properties. During the loading of geographical data of these formats, they are converted to the SMD format. Then, they may be rendered using the WMS service of SMD Server.

The Geoportal allows inputting and editing relational data. This system consists of the following components. The first subsystem supports loading and normalization data from Excel or CSV files in the PostgreSQL database. The second subsystem serves to generate table data structures. It creates new tables and uses existing tables for viewing and editing their metadata. The third subsystem provides support for the inputting, editing, and printing of tabular data (see Fig. 3).
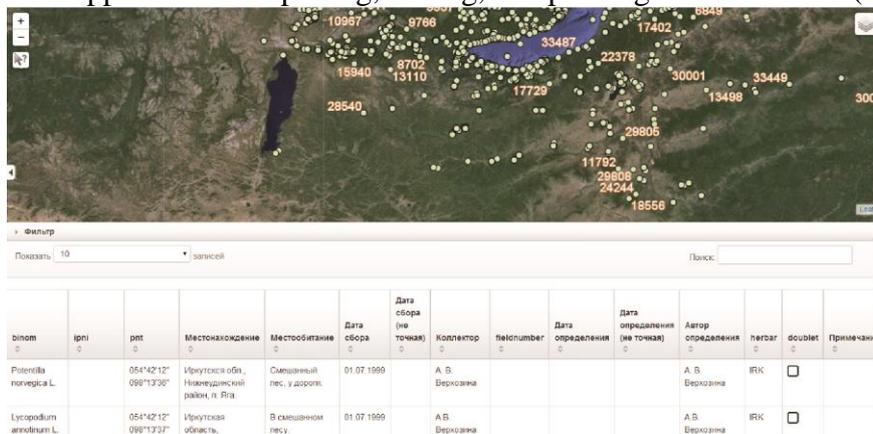
Figure 3: Viewing and editing tabular spatial data on the Geoportal

Conclusions

There is a large amount of spatial data (thematic maps, layouts, space photos and others) gathered in the scientific institutes of Mongolia. These data, however, re used in the SDI-contrary not

distributed architectural geoinformatics paradigm. All of the functions for spatial data processing are used in a unified GIS environment according to this paradigm. This paper proposes a technological solution aimed at forming a local spatial infrastructure in institutes of the MAS. It includes the modernization of informational resources and also the technologic development of visualization and geoprocessing services. The creation of this SDI will make it possible to manage all of the scientific spatial data and use these spatial data in research at the MAS institutes more effectively.

Acknowledgements

References

[1]   Masser, I. *GIS Worlds: Creating Spatial Data Infrastructures*. ESRI Press, 2005.

[2]   Masser, I. All shapes and sizes: the first generation of national spatial data infrastructures. *Int. J. of of Geographical Information Science*, 13(1), pp. 67-84, 1999.

[3]   Craglia, M., Annoni, A., INSIRE: An innovative approach to the development of spatial data infrastructures in Europe. *Res. and Theory in Advancing Spatial Data Infrastructure Concepts*, 2007.

[4]   Koshkarev, A.V., Ryakhovsky, V.M., Serebryakov, V.A. Infrastructure of distributed environment of spatial data storage, search and processing. *Otkritoe Obrazovanie*, (5), pp. 61-73, 2010.

[5]   *Space Technology Application Capabilities, Facilities and Activities in Asia and the Pacific: A Regional Inventory*, United Nations ESCAP, 2007.

[6]   Paramonov, V.V., Fedorov, R.K., Ruzhnikov, G.M., Shumilov, A.S. Web-based analytical information system for spatial data processing. *Communications in Computer and Information Science*, 403, pp. 93-101, 2013.

[7]   Nergui, B., Bulgan, G. Status of e-science activities in Mongolia. *Proc. of the Int. Symposium on Grids and Clouds*, Academia Sinica: Taipei, 2012.

[8]   Dublin Core Metadata Element Set, dublincore.org/documents/dces

[9]   FGDC-STD-001-1998 Content Standard for Digital Geospatial Metadata. www.fgdc.gov/metadata/csdgm

[10] ISO 19115:2003 Geographic Information Metadata.

[11] ISO 19139:2007 Geographic Information Metadata XML Schema Implementation.