IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Hershey PA 17033, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com
ITP4297

Spatial Metadata Server: Supporting Data Clearinghouse in the Hindu-Kush Himalayan Region

Iftikhar U. Sikder
Department of Information Systems, University of Maryland Baltimore County
Tel: (410) 455-2620, isikde1@umbc.edu

Nagma Yasmin
Department of Computer Science, University of Northern Iowa
Tel: (319) 222-5931, nag1211@uni.edu

Aryya Gangopadhyay
Department of Information Systems, University of Maryland Baltimore County
Tel: (410) 455-2620, gangopad@umbc.edu

ABSTRACT

This paper outlines an ongoing research on the implementation of spatial metadata server for the Hindu-Khus Himalayan region to facilitate spatial data sharing among different regional agencies. As a mainstream of information technology, Geographic Information Systems (GIS) is getting into the core of natural resource planning and decision-making in the mountain region. In order to communicate and interoperate with different spatial data warehouse in the region, a metadata sever is an essential and an integral part of the spatial data infrastructure. We identify issues on metadata sharing standards and the institutional position in the Hindu-Khus Himalayan region. As a part of the FGDC(Federal Data Geographic Committee) compliant metadata tools, we illustrate a spatial search tool to allow for cataloging, indexing and searching spatial data. We also outlined a planned distributed architecture for metadata server and implementation schema to map the existing differences in the metadata standards.

INTRODUCTION

The Hindu Kush Himalayan (HKH) Region covers more that 4 million sq. Km in the mountain range of the Himalaya. The region consists of up to eight different countries (Afghanistan, Bangladesh, Bhutan, China, India, Myanmar, Nepal, Pakistan) or mountainous parts thereof. The area is one of the most diverse ecosystems in the world with rich biodiversity and natural beauty. Most its parts are inaccessible due to difficult topography and political unrest. However, intensive population pressure and rapacious use of natural resources have generated a serious concern on this fragile ecosystem. Considering its ecological sensitivity and increasing importance of the region to the surrounding ecosystem, various scientific research institutes and development agencies are working on sustainable mountain development research programs. These research institutes have already gathered a substantial amount of information in various formats. However, there is no accepted standard or protocol that could bring these national or local institutes to a common agreement to exchange information and services. Inter-organizational impasse in communication and cooperation has always been a principal factor in the lack of data exchange in the region. The International Center for Mountain Development (ICIMOD) was established in 1983 in response to the widespread concern for environmental degradation as a forum for research, scientific exchange and documentation in the region. The role of ICIMOD is to coordinate research initiatives of the national and regional institutes. As a thematic division of ICIMOD, Mountain Natural Resources Information Service (MENRIS) was established in 1991 to foster network of nodal agencies to support information exchange. The objective of MENRIS is to develop a database on geomorphology, soil, land use, vegetation and related factors through remote sensing techniques. MENRIS has been leading the effort to bring consensus on data exchange protocols and standards especially in the spatial database. As a part of its initiatives MENRIS is undertaking steps to improved spatial data access by providing metadata services to user community. The role of MENRIS is to define pointers to data servers of different institutes in the region by developing a metadata server that would document the following data type:

- GIS data sets on topography, geology, land cover, hydrography, land cover, soil, administrative and socio-economic statistics
- Raw or georeferenced satellite images or air photo in hold by MENRIS or partner agencies.

In addition to technical specifications, the metadata server is expected to have general reference information on data quality, history, retrieval system, location of GPS base stations, satellite referencing systems. The metadata can be used to leverage data resources and to avoid redundant data collection and preparation in the region. This paper reports on the ongoing effort of MENRIS to facilitate the spatial data exchange mechanism with reference to spatial the metadata framework. The paper outlines an approach to metadata search engine as an active component of GIS applications.

SPATIAL METADATA IMPLEMENTATION GUIDELINE

The institutional approach to spatial data sharing initiative was formalized in response to an Executive Order 12906 (1994), *Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure (NSDI)*. The vision of the NSDI strategy was that "current and accurate geospatial data will be readily available to contribute locally, nationally, and globally to economic growth, environmental quality and stability, and social progress". The FGDC (*Federal Data Geographic Committee*) is in the process of implementing the *Content Standards for Digital Geospatial Metadata* and the *National Geospatial Data Clearinghouse*. The content standards identify the elements that are needed to characterize the data content and quality. The objective of the *National Geospatial Data Clearinghouse* was to facilitate large-scale access to geospatial metadata through distributed, electronically linked stores of information about geospatial data (FGDC 1995). The vision of NSDI was echoed in the HKH region

This paper appears in *Issues and Trends of Information Technology Management in Contemporary Organizations,* the proceedings of the Information Resources Management Association International Conference.

in various agenda (UNCRD, 1997). NSDI is viewed as a framework to bridge the gap between information "haves" and "have-nots" in the HKH region. So far, at the regional level, there are little agreement on spatial data content standards. ICIMOD's position, in this respect, is not to impose any standard on the national institute rather to accord with international standards to promote mutual understanding on data sharing mechanism in the region. Informally, MENRIS has adopted FGDC standard. However, this would mean an interoperation with other metadata parsers (e.g., DIF standard) currently being used in organization (e.g., Directory Interchange Format of NASA, 1998; Bitter, 1999). There is still a need for regional level dialog and understanding of the FGDC standards especially with reference to the metadata documentation fields, which are related to institutional jurisdictions and policies. Regional acceptance and technical capability is not sufficient to customize the interface with local needs. Metadata creation and sharing inevitably requires rigorous stuff training and implementation – which means additional overhead on project budget. Moreover, there is hardly any project specific requirement or commitment to documents metadata because the benefit of metadata is not yet realized at project planning level. In order to illustrate the significance and relevance of metadata search tool in the scientific research, we developed a prototype system for cataloging and searching metadata in the HKH region. The tool offers cataloging, documenting and searching functionalities in a customizable user interface.

IMPLEMENTING METADATA TOOLS

Early attempts to implement FGDC standard included creation of word processing forms to layout documents, which was often cumbersome. USGS developed DOCUMENT.AML – an Arc/Info based macro language interface to collect and document metadata following FGDC standard. Majority of the tools available to document metadata involves using DOCUMENT.AML system where the data specifications can be extracted from the dataset itself using AML (Arc Macro Language) directives &DESCRIBE. The information extracted is stored in ASCII text file. The FGDC has developed ASCII based metadata compilation tools for different platforms, to support the development of the NSDI Clearinghouse network. In order to search through the data set in ASCII file user still need to organize, track them down and normalize them for further analysis. The database for metadata, therefore, must allow standard database query and indexing technique to make the search process efficient. User must be enabled to make visual query over a thematic layer. For example a typical visual query could be "How many satellite images are available within my area of interest?" or "What are the neighboring countries that have less cloud cover? "

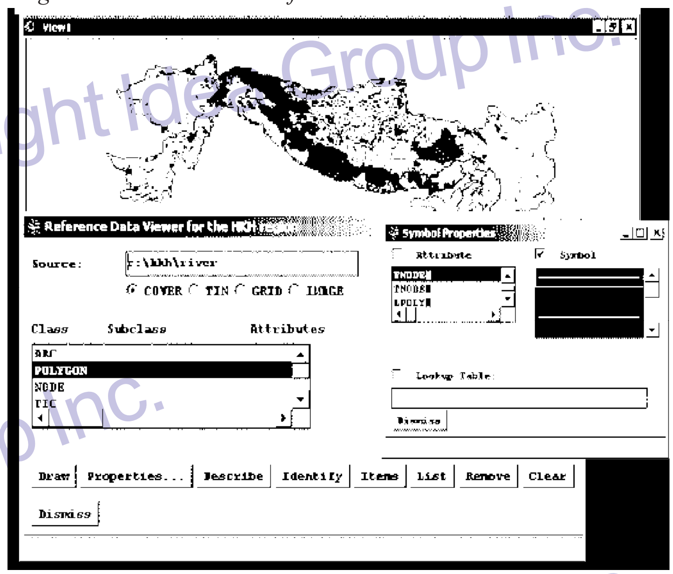
SPATIAL METADATA SERVER FOR THE HKH

The HKH metadata server is a test bed prototype system for the HKH region to catalog, document and search metadata in a user-friendly graphical interface. The system offers a menu driven visual interface where user can make different types of visual query. The system was designed in ESRI's Arc/Info environment using AML with sophisticated search options. Currently, the system is in the process of migration to ArcCatalog and ArcIMS to ensure easy access and search metadata from a standard Web-browser.

Thematic Data Viewer

This option offers the easiest way to view spatial data from the thematic classifications. The classifications hierarchy includes different ecological variables (e.g., rainfall, river system, transport system etc.). Once a thematic layer is selected the user is prompted to regional, national, and administrative level hierarchy to select the lowest level of spatial resolution/scale. The spatial data type supported includes Arc/Info Cover, TIN, GRID and Image. The layer can be visualized and overlayed for different feature classes with different symbol properties. However, the definition of the classification hierarchy is still subject to further modification. At present, there is a need come up with a flexible and customizable hierarchy unique to user type so that organization-centric view is reflected in the hierarchal structure. Future version of the server is planned to address this problem.

Figure 1: Thematic viewer of the HKH metadata server



Query Frame

Query frame allows visualization of dataset from its "footprint" or data frame in response to visual query. For instance, a user can search the list of satellite images from the footprints in a region. Once a footprint is selected a window pops up with the list of satellite images and corresponding attributes.

User can further narrow down the search result and select a particular image (in degraded resolution) to visualize and check its quality in a "quick-look" frame. Search frame also allows text-based search from user-defined criteria in a search window. The search can be made with geographic extent (latitude/longitude), satellite sensor type, path/row, and percentage of cloud cover, date, storage type and many other criteria.

Editing Frame

This feature allows adding, updating, and removing a frame or a quick-look in the existing catalog. Adding a frame involves creation of a unique identifier and updating the catalog index. Frames can be added following a thematic hierarchy of the metadata.

For instance, when user adds a new frame for SPOT-PAN image, sensor specific information of SPOT-PAN is retrieved automatically from the pre-existing database. User only needs to add geographic references and corresponding attributes (e.g., Date, Row, Path, Center Lat, Center Lon, etc.). In order to change or delete a frame user need to identify the frame either visually or through text search options. A dialog box appears to help user if any change or update is necessary.

Metadata Documentation

This feature allows documentation of metadata in compliance with the FGDC standards. From the implementation point of view, the system uses the DOCUMENT.AML and integrates it with the frame based cataloging interface. Several FGDC fields are automatically filled out from the frame catalog so as to avoid redundant entry. However, there are fields, which are still empty and need to be filled out with consultation with partner organizations. A properly documented

Figure 2: Sample query frame for IRS-WIFS and Quick look views

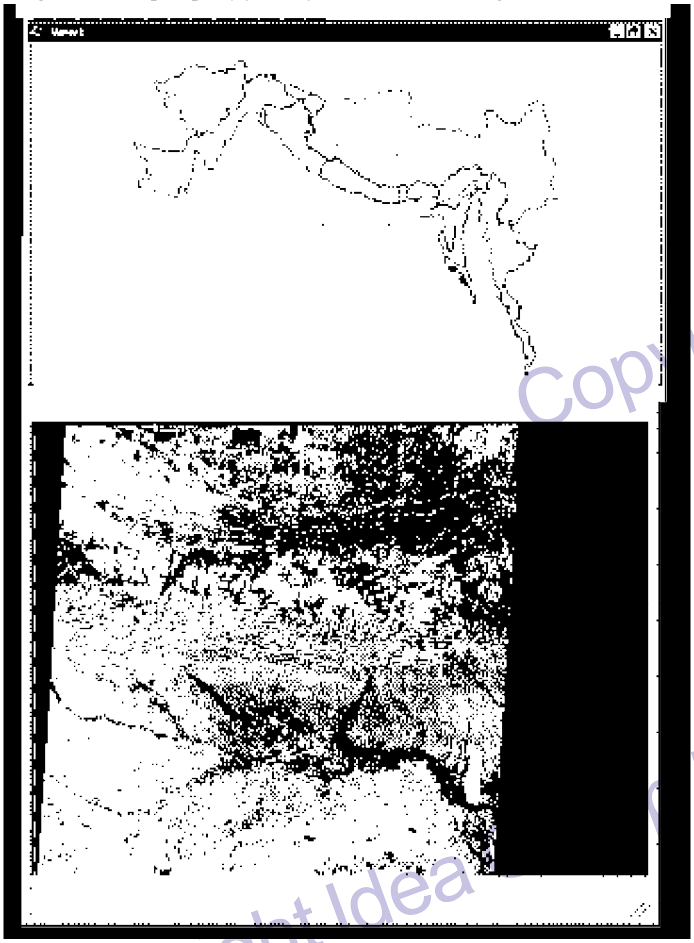
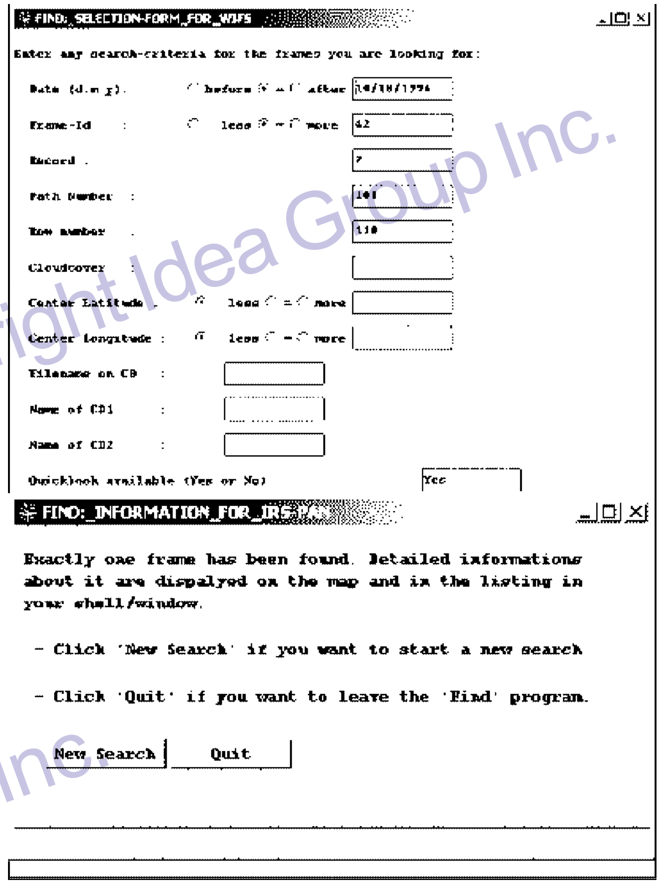


Figure 3: Text based serach for image frame



Results of Query Frame:

C_WIFS region.wif : 2 of 2 selected.

Record	SENSOR	DATE	FRAME-ID	PATH	ROW	CLOUDCOV
1	Wifs	10/18/1996	42	104	52	20/100
2	Wifs	11/29/1996	2	110	52	5/100

FOR MORE DETAILS ENTER REC-NO OF REQUIRED FRAME, ELSE PRESS <CR>: 2

SENSOR = Wifs

DATE = 11/29/1996

FRAME-ID = 2

PATH = 110

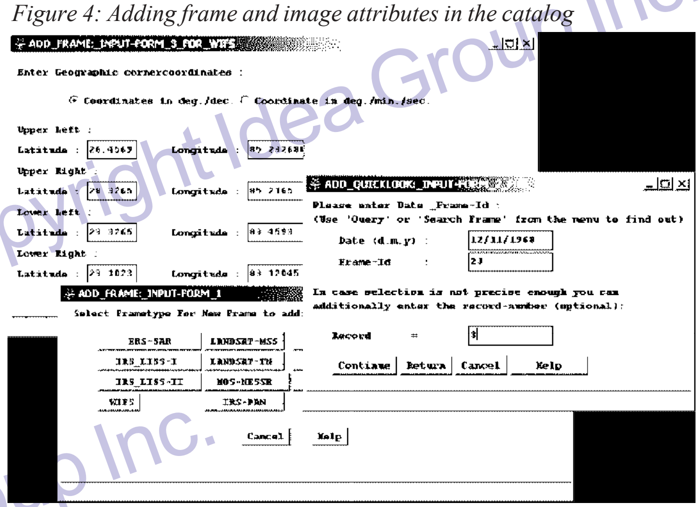
ROW = 52

CLOUDCOV = 5/100

CENTERLAT = 27.402618

CENTERLON = 91.944610

Figure 4: Adding frame and image attributes in the catalog



metadata can be exported to ASCII format. In addition, the system offers full view of the FGDC metadata, view of references, attribute description, narrative description and log file etc. However, in the future version, use of DOCUMENT.AML is still subject to further consideration, for several literatures have identified its shortcoming (Nelson, 1997) with regard to its compatibility with FGDC standard and different versions of Arc/Info.

Next Step-Distributed Architecture

The Metadata Server for HKH needs to be derived from the data itself, i.e., data provider must assume the responsibility of generating metadata. Therefore data and metadata cannot be separated from each other. Thus a distributed architecture is necessary where individual organization can generate metadata in a customizable interface and standards and at the same time allow different stakeholders to access and search the metadata from a standard web browser. There is a need to automatically derive metadata information from the data. ESRI's recently released out of box ArcCatalog is somewhat directed towards these capabilities. The application can be used for locating, browsing, and managing geo-dataset. It maintains a strong automatic association of metadata with data set and supports the FGDC standard for spatial metadata. It supports Arc/Info as well as other data format (e.g., CAD, text, images, scripts etc). The metadata is stored in XML file and can be viewed in any XML-aware environment. Different organization can develop their native custom environment in ArcCatalog. This involves designing a logical model of metadata content, an implementation schema, an editor and style sheets. Individual organization can have the freedom to decide the metadata elements they want to incorporate. Then, the logical model needs to be translated into implementation schema in XML resulting attributes as XML tags that follows the FGDC content standard. The metadata editor would allow user to edit metadata following a standard or other defined model.

Figure 5: Editor mediated metadata schema generation

However, ArcCatalog does not offer any distributed metadata search engine. However, it generates open format XML schema that could be used in any standard browser to create customized query tools. This could be linked with ArcIMS HTML viewer to search and displays a thematic map. This page could be created using Arc XML (AXL) rendering code. User request can be sent to a servlet, to dynamically render map service using the stored AXL code. The exploration environment could be further enhanced by the ArcIMS Java Custom Viewer ensuring easy and smooth access to spatial dataset.

CONCLUSION

The demand for metadata sever in the HKH region is crucial for circulating knowledge and scientific research results in the region. By providing the resources to document large numbers of legacy datasets, we plan to enable different organizations to share spatial data and bridge institutional gaps by stimulating the need to create metadata. The tools that we developed needs to be incorporated in an open extensible distributed architecture to reach its design goal. However, this will also require organizational commitments and understanding to foster the data-sharing environment in the region.

REFERENCES

Bitter, P. (1999). A Spatial Data Infrastructure for the Hindu Kush Himalayan Region. In Interoperating Geographic Information Systems (Eds. A. Vckovski, K. Brassel and H. Schek) Second International Conference 1999.

Federal Geographic Data Committee (1995). Content Standards for Digital Geospatial Metadata Workbook (March 24), Federal Geographic Data Committee, Washington, D.C.

Nelson, Daniel O., Krumm, Robert J., Denhart, Sally L. and Beaverson, Sheena K. (1997). ArcInfo Solutions to Metadata Problems: Building a Solid NSDI Clearinghouse Node on a Shifting Metadata Landscape. ESRI User conference Proceedings, 1997.

UNCRD (1997). Space Informatics for Mountain Resources Management. Proceedings of the Second Space Informatics Information Seminar for Sustainable Development: United Nations Centre for Regional Development.