National Spatial Data Infrastructure

Cooperative Agreement Program

Metadata Training for Earth Science and Cyberinfrastructure Communities

Final Project Report

Submitted to:

Sharon Shin, Metadata Coordinator, and
Brigitta Urban-Mathieux, NSDI CAP Coordinator
MS 590 National Center
Reston, VA 20192

Submitted by:

Chaowei Phil Yang
The Joint Center of Intelligent Spatial Computing
George Mason University
4400 University Drive, MS 4C6, Fairfax, Virginia 22030
http://cisc.gmu.edu

Final Report Summary Information

Date:
Aug 31, 2009

Agreement Number:
08HQAG0015

Project title:
2008 CAP-Category 1: Metadata Training for Earth Science and Cyberinfrastructure Communities

Interim or Final report:
Final Report

Organization:
Joint Center for Intelligent Spatial Computing (CISC),
George Mason University (GMU),
4400 University Drive, MS 4C6, Fairfax, Virginia 22030
http://cisc.gmu.edu

American Association of Geographers (CISG, AAG)

Federation of Earth Science Information Partners (ESIP)

Project Leader:
Chaowei Yang, (703)-993-4742, cyang3@gmu.edu
Co-investigators:
Rob Raskin AAG CISG, raskin@jpl.nasa.gov
Carol Meyer, ESIP, carol.meyer@earthsciencefoudation.org
Hannes Wu, GMU CISC, hwu8@gmu.edu

Final Report

Executive Summary

The project is to conduct metadata training for the Federation of Earth Science Information Partnership (ESIP) and the Cyberinfrastructure Specialty Group (CISG) within the American Association of Geographers (AAG). With the support of this project, in addition to the accomplishments described in the interim report, we have provided:

    1) Metadata training and demos in the ESIP winter semi-annual meetings of 2009;

    2) A metadata workshop in the AAG 2009 annual meeting;

    3) Metadata training in the ESIP summer semi-annual meetings of 2009;

    4) Integrate metadata training of distributed GIS courses in GMU in spring semester 2009;

    5) The 2009 CISC international summer training program on Fairfax campus, GMU;

    The project's outcomes include:

    1) ESIP, AAG CISG, and other organizations and individuals received training within the communities of Earth Science and Cyberinfrastructure are competent in creating compliant metadata;

    2) Metadata files and services built in Air Quality and water cluster in ESIP;

    3) A Catalogue Service for Web (CSW) that can be discovered and harvested through GOS, Z39.50 Clearinghouse server, and the FGDC Browse-enabled Web Directory;

    4) Relationships created with ESIP, CISG, and other organizations to sustain metadata activities beyond the performance period;

Project Narrative
Activity I: Publish, Harvest, and Query Metadata

This activity is trying to address OGC interoperability standards including Web Map Service (WMS), Web Feature Service (WFS), and Web Coverage Service (WCS) by building on a community catalog for air quality working group to publish, harvest and query metadata. The development of the community catalog is to utilize OGC Catalogue Service for the Web (CSW) standard to build both catalog servers for publishing, harvesting and conducting transaction of metadata. Meanwhile, CSW client is developed to train domain users to understand the working mechanism of catalogue service as well as how to correctly format and send a metadata query from the metadata catalogue. Currently, the operational CSW server is hosted at GMU site (http://gesg.gmu.edu:8083/aq201/srv/en/csw?request=GetCapabilities&service=CSW&version=2.0.1). Metadata training in this activity includes guiding AQ community researchers on ISO19115/ISO19139 metadata creation, metadata and OGC service publication, CSW client development.

Activity II: Z39.50 Metadata Service Setup and Training for Air Quality Community

A full background study was conducted on both open-source and commercial software for establishing Z39.50 metadata server for geospatial information retrieval. The software studied included open source ones such as Isite and Geonetwork; commercial toolkits including Compusult MetaManager Toolkit, ArcIMS metadata service (ArcCatalog, ArcSDE and Z39.50

Connector) and Intergraph Geoconnect Metadata Management Server. We also evaluated the full support of the software to various metadata standards such as FGDC metadata, ISO 19115(19119), NASA DIF, ANZLIC metadata and Dublin Core. And FGDC recommended metadata service Isite was selected for Z39.50 metadata training because of its stability and easy-to-use features. The Z39.50 server is 129.174.65.103 at port 6668.

We conducted two training sessions in 2009, the ESIP Winter Meeting in D.C. and a Summer Meeting in California. The training session includes hands-on experience on how to put imported metadata files into Z39.50 server; how to build spatial index to speed up spatial query on the client side; and how to conduct regular harvest, full harvest, validation and publishment of existing Z39.50 services into Geospatial One Stop (GOS) portal. This activity mainly orients researchers in Air Quality (AQ) community who are trying to build up and maintain metadata information about emissions, human dimensions and other impacts on environment and atmosphere.

Activity III: Metadata within Cyberinfrastructure

A Metadata workshop was conducted during the AAG annual meeting on March 22nd, 2009. We integrated our metadata research experiences from several projects like NASA ESG (Earth Science Gateway), ESIP EIE (Earth Information Exchange), and WECHO (WaterCycle for Earth Clearing HOuse) to instruct the attendees on FGDC metadata standards and answered several important questions; how can we use the metadata, what are the online tools available for us to create metadata, how can we search data by using metadata catalogs, how does metadata relate to knowledge representation and ontology, how can metadata help with knowledge-based reasoning and computing scheduling, as well as data management and visualization, how can metadata be utilized within grid computing, cloud computing, and other advanced CyberInfrastructure topics? Ms. Sharon Shin participated in the training workshop by introducing metadata and other FGDC CAP programs.

This training was also conducted during the CISC summer training program where 18 international students were instructed with detailed information about metadata standards and catalogue services.

Activity IV: Semantic-Enabled Meta-Catalogue Research

Semantic enabled metadata-catalogue research was deployed during this project including research activities and training. The cutting-edge semantic web technology is utilized in order to seamlessly connect various geo-catalogues such as FGDC's Geospatial One Stop portal, NASA's Earth Science Gateway Portal, NOAA's National Climate Data Center Portal, NASA's Earth Clearinghouse Portal, and NASA's Global Climate Master Directory was applied to form an integral portal for geospatial data and metadata discovery. Semantic technology is added to the query layer to improve the accuracy of traditional keyword based search functionality. We presented this new research outcome during the 2009 summer training program to international researchers in geospatial fields. The current available operational system can be found at:
http://eie.cos.gmu.edu/c/portal/layout?p_1_id=PUB.1.113

Training and outreach assistance:

I: Metadata Training in the ESIP Winter Semi-annual Meeting

We provided an hour long metadata training session on Jan 6, 2009 during the ESIP winter semi-annual meeting held in the Embassy Suites Convention Center in Washington D.C. 50 people from ESIP member institute and some other professionals and students from academia joined this training. A live demo was given during the poster session.

II: AAG 2009 Pre-conference Metadata Workshop

A metadata training workshop was organized on March 22, 2009 within the AAG 2009 conference in Capri Meeting Room 109, Reviera Hotel, Las Vegas. Cyberinfrastruture Specialty Group (CISG) of AAG sponsored this workshop. Phil Yang, Wenwen Li, Zhenlong Li, Rob Raskin, and Sharon Shin gave presentations and online demos. Students experienced the demos under the guidance of the instructors. This workshop integrated several examples of GOS, EIE, WECHO, and Interoperability Testbed to introduce the metadata issues including FGDC metadata standards, Catalogue Service for Web, online tools available for metadata, knowledge representation and ontology for metadata, and metadata visualization. Most of the participants brought a laptop for practicing the operations of metadata. Around 25 people joined the workshop.

III: Metadata Presentation in ESIP Summer Semi-annual Meeting

We gave a half-hour presentation on metadata search, access and visualization on July 8, 2009 in metadata showcase session in the ESIP summer semi-annual meeting held at University of California, Santa Barbara, California. 30 people joined this session. A live demo was given during the presentation.

IV Integrated Metadata Training in GMU Course;

The course Distributed GIS was designed for senior undergraduate students and graduate students in Fairfax campus, GMU. Ten students selected this course in the spring semester. A 3-hour course about metadata standards, catalogue service, search and visualization is given.

IV: CISC International Summer Training Program

The 2009 Advanced Certificate for GeoInformation Sciences (ACGIS) program was successfully accomplished on Aug. 20, 2009. Eighteen professionals from the surveying and mapping bureaus

of Heilongjiang, Guangxi, Shanghai, Jiangsu, Zhejiang and Hubei provinces in China participated in the 8-week training from July 19 to Aug 20, 2009. A half-day training was conducted about metadata standards, catalogue service, metadata clearinghouse, search, access and visualization.

Status of Metadata Service

The metadata is harvested by Geospatial One Stop portal and Global Earth Observation *System of Systems* GOESS system.

In Activity I, we have maintained and harvested 60 metadata files.

In Activity II, we have maintained and harvested 18 metadata files.

In our Earth Information Exchange system, we have integrated 5 catalogue services, over 900 WMS servers and more than 10000 registered layers of Web Map Services, Web Feature Services and Web Coverage Services. (http://eie.cos.gmu.edu/c/portal/layout?p_l_id=PUB.1.92)

Next Steps:

This project helped us conduct training in ESIP, AAG communities and international geospatial professionals. We will be happy to work with FGDC to conduct more metadata-related training, technology development and research to advance the leverage of our expertise on geospatial metadata standards, creating metadata catalogue, metadata service, integration, and visualization.

NSDI CAP Feedback

1. What are the program strengths and weaknesses?

The program gave us a great opportunity to interact with beginners to advanced metadata users. We were able to provide basic information about metadata to the beginners such as: catalogue, search and query, and how metadata can be applied to their own domains. For the advanced users, we are able to teach them how to create a catalogue service for their own data set and integrate it with other catalogues. This program also helped us to promote FGDC metadata standards to students from another country. The program is very operable and is directly linked to real applications across a wide-range of communities.

While the metadata technology is relatively stable, the major weakness of this program is that we only had a one-year period to conduct the training. A longer period, for example, 3-5 years, would be helpful for us to form a comprehensive training series of training and contribute to different aspects in formulating a metadata science through collaboration with various communities and the FGDC.

2. Where does the program make a difference?
The program helped us reach a wide variety of audiences based on the PI's network. On another hand the training also helped us have a more improved position of ourselves in the community.

3. Was the assistance you received sufficient or effective?
The assistance we received was sufficient and effective for the goal of our project. A longer term and more support would help us identify research issues with the community for metadata science

4. What would you recommend doing differently?
We hope that a longer term is available from FGDC for us to work with the community to maintain a set of online demos relevant to the training tasks/modules.

5. Are there factors that are missing or additional needs that should be considered?
Same as stated in point 4.

6. Are there program management concerns that need to be addressed, such as the time frame?
Same as stated in point 4.

7. If you were to do this again, what would you do differently?
The project was a big success for us in our collaboration with FGDC to conduct training for different domains. With more experience in working with multiple domains on materials, knowledge, systems of metadata, we would be able to expose attendees to multiple domains, multiple catalogues, and multiple frontier technologies, such as semantic search, technologies such as semantic search to help interact with metadata database.