

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Faculty Publications, UNL Libraries

Libraries at University of Nebraska-Lincoln

---

10-2010

## GEOSPATIAL DATA PORTALS: LIBRARIANS ADD EXPERTISE IN THE DEVELOPMENT OF GIS METADATA CATALOGS

Adonna Fleming

University of Nebraska-Lincoln, [dfleming2@unl.edu](mailto:dfleming2@unl.edu)

Follow this and additional works at: <https://digitalcommons.unl.edu/libraryscience>



Part of the [Geology Commons](#)

---

Fleming, Adonna, "GEOSPATIAL DATA PORTALS: LIBRARIANS ADD EXPERTISE IN THE DEVELOPMENT OF GIS METADATA CATALOGS" (2010). *Faculty Publications, UNL Libraries*. 297.

<https://digitalcommons.unl.edu/libraryscience/297>

This Article is brought to you for free and open access by the Libraries at University of Nebraska-Lincoln at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, UNL Libraries by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# GEOSPATIAL DATA PORTALS: LIBRARIANS ADD EXPERTISE IN THE DEVELOPMENT OF GIS METADATA CATALOGS

Adonna C. Fleming  
Geology Library  
University of Nebraska - Lincoln  
dfleming2@unl.edu

*Abstract* — Searching a geospatial data catalog can be frustrating for many GIS data users. Geospatial data catalogs are built on records created to a specific metadata standard, such as ISO 19115. The search query is often limited to searching a few tags within the record, such as title, publisher, spatial extent, content theme, and content type. GIS personnel tend to create metadata records with little thought into how it will be discovered by others. On the other hand, library catalogs are developed for a broad spectrum of users with varying knowledge of the subject. Librarians can bring this expertise in creating user-friendly catalogs to the GIS profession.

In the development of NebraskaMAP, a statewide geospatial data sharing portal and web services network, the GIS Librarian from the University of Nebraska-Lincoln used knowledge of Boolean operators, Lucerne query syntax, and the understanding of geospatial metadata standards to improve the searching, retrieval, and display capabilities of this geospatial data portal's metadata catalog.

This paper describes the process of building a statewide geospatial data portal in Nebraska, and how the UNL Libraries contributed to the organization of the portal's metadata records. The portal uses ESRI's Geospatial Portal Extension software.

## INTRODUCTION

Wording used to describe geospatial portals frequently includes terms such as gateway, discovery, and user-focused access to spatial data. For example, Maguire and Longley (2005, p.7) state, "Portals are web sites that act as a door or gateway to a collection of information resources." Tait (2005, p. 34) continues, "A geographic portal is a web site where the discovery of geographic content is a primary focus," and the ESRI GIS Dictionary (2006) states, a GIS portal is a "Web resource that provides access to a broad array of related resources and services." The question becomes how well are geospatial portals living up to this concept of *access* and *discovery* of spatial data? Not very well, according to Comber, et al. (2008, p. 287), "In the domain of spatial information semantics are poorly treated by metadata and data standards."

Geospatial portals, like libraries, build their cata-

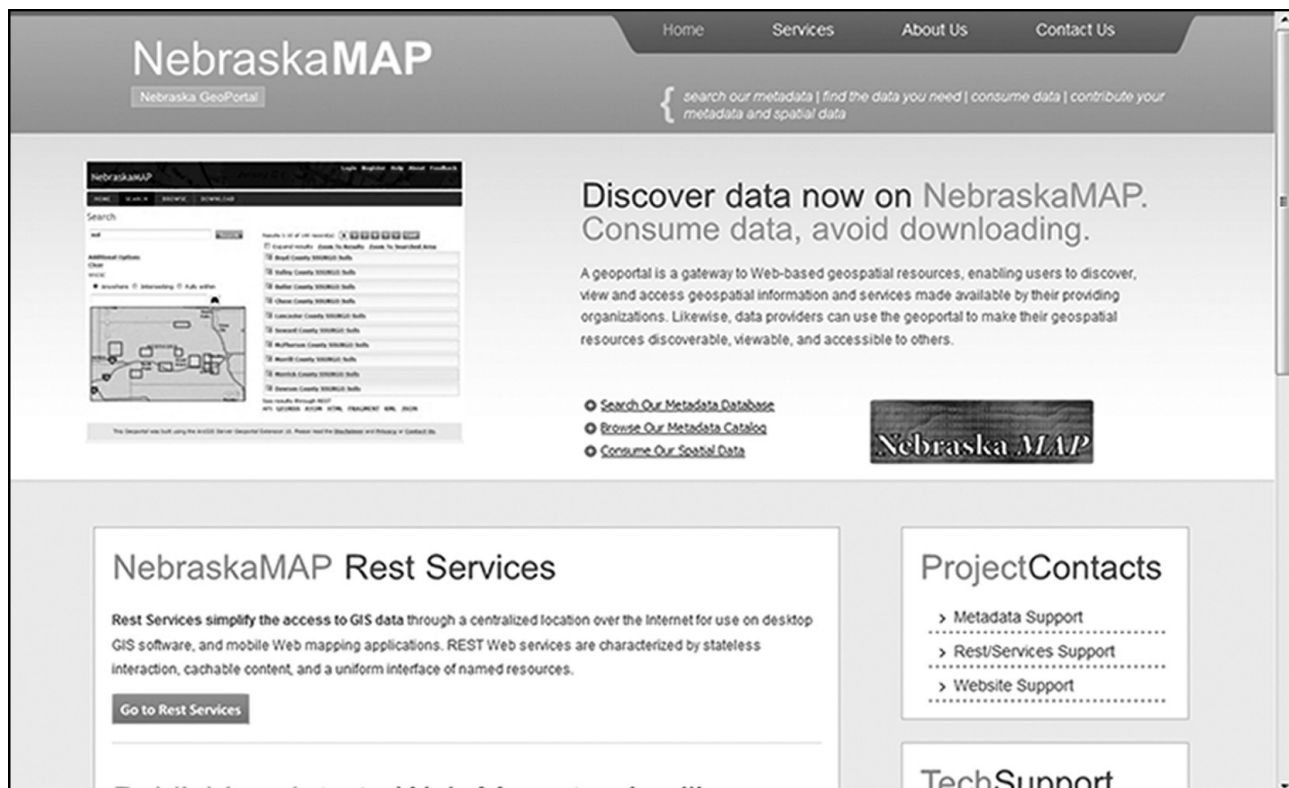
logs based on metadata records. The records are stored in an online database that is accessed when a user enters a search query through the portal's client software interface. This software is built on World Wide Web technology utilizing a web browser to search and display query results. World Wide Web technology enables geospatial portals to become interactive sites where users search a metadata database for information about spatial data and services, as well as directly connect to these services for live mapping, geocoding, and routing. The technology also allows for the downloading of spatial data to be used in more robust desktop GIS applications such as ESRI's ArcMap or Google Earth.

Metadata, often described as "data about data," is the roadblock to intuitive, user-friendly searching of spatial data catalogs. Encoding within metadata records is standardized to help facilitate the searching of spatial data produced by different govern-

mental agencies and commercial developers. The standard recommended for use in the United States for spatial data is the North American Profile (NAP), adopted in 2009 by the Federal Geographic Data Committee (FGDC) and based on International Standard (ISO) 11915, which was approved by the International Standards Committee in 2003. These standards dictate which fields in the record are searchable and what information is required in each field. Geospatial portal developers often use commercial software, such as ESRI's Geoportal Server Toolkit Extension, for managing the spatial data catalog. The ESRI product indexes the meta-data database to allow for searching by three parameters: spatial, temporal, or thematic. The NAP standard dictates that elements within the metadata record describe spatial data in terms of content, quality, condition, and origin. However, no required element defines the purpose or use of the data, or allows for "free term" word searching. While technology has made spatial data available to a wide audience through the World Wide Web, discovery by non-specialized users is frustrating at best. Either the terminology the general populace uses does not access the desired records, or the retrieved data is inappropriate for their needs.

Thus, geospatial portals, although intended to be one-stop shopping for everyone, are marginally successful, and often are accessed and used only by specialized GIS users who already know what they are looking for before they begin their search.

Librarians, whose profession is founded in the organization and access of information, can provide a broader perspective to the design of geospatial portals. Library online catalogs are built on the same technology as geospatial portals and librarians are experts in working with patrons who have varying degrees of experience in navigating their catalogs. Having a librarian involved in the development and management of a geospatial portal brings a unique perspective often over-looked in geoportal development. Librarians bring to the project their expertise in indexing and catalog development. They understand how coding within the metadata record defines search parameters, and determines retrieval and display of the results. This, in turn, has the capability of improving the functionality of the portal through the creation of "search savvy" metadata records. As members of the project team, librarians can use their influence to advocate for applying their expertise in the education of GIS metadata creators. This was the case in the devel-



**Figure 1.** NebraskaMap is a gateway to Nebraska-related geospatial data and services from a wide range of state, local and federal agencies.

opment of NebraskaMap (<http://nebraskamap.gov>, Figure 1), which included a librarian from the University of Nebraska-Lincoln (UNL) on the development taskforce.

## NEBRASKAMAP

NebraskaMap is a collaborative project initiated by the Nebraska GIS Council under the guidance of the Nebraska Information Technology Commission (NITC). The goal of the NebraskaMap project is to develop a geospatial portal to enable finding and providing online access to Nebraska-related geospatial data and services available from a wide range of state, local, and federal agencies.

In 2009, the NebraskaMap Partners Committee, an interagency taskforce, was formed to help advise and guide the portal's development and implementation. Members of the committee included: Nebraska GIS Council, Nebraska Office of the Chief Information Officer, City of Lincoln/Lancaster County, City of Omaha/Douglas County, Nebraska Department of Natural Resources, Nebraska Department of Health and Human Services, Nebraska Department of Roads, Nebraska Emergency Management Agency, Nebraska Game and Parks Commission, Sarpy County, University of Nebraska-Lincoln Libraries, University of Nebraska-Lincoln School of Natural Resources, UNL Center for Advanced Land Management Information Technologies (CALMIT)/NebraskaView, and the U.S. Geological Survey.

Two-year grant funding enabled the hiring of a project administrator, and for CALMIT and UNL to provide project management and technical leadership. Along with the NebraskaMap Partners Committee, these agencies were responsible for day-to-day management of the project implementation. The UNL librarian on the NebraskaMap Partners Committee was responsible for the metadata component of the portal.

## THE LIBRARIAN'S GOALS

1. Enhance the "out-of-the-box" portal development software to improve data searching and results display.

2. Develop a customized standard for metadata entry (Nebraska FGDC).
3. Through a series of workshops, teach NebraskaMap Partners Committee members and other metadata contributors how to create metadata in the portal that is enhanced for user-friendly searching and is compliant with the Nebraska FGDC standard.

NebraskaMap is built on ESRI's Geospatial Server Toolkit Extension version 10. Without any customization, the out-of-the-box editing tool for the NAP standard requires the following fields to be populated:

*General information* – Provides information about the metadata. For example, the organization which created it, who to contact for more information, and the date it was created. These fields are automatically populated in the NebraskaMap portal based on information from the registered users' profiles.

*Identification* – Includes the title of the spatial dataset or service, an abstract, and the dataset language.

*Citation date* – Includes the time frame the spatial data was created or revised.

*Dataset development phase* – Describes the currency of the data.

*Contact for resource* – Includes who or which organization to contact about the spatial data. These fields are automatically populated from the registered users' profiles.

*Data theme* – Describes the spatial data using a controlled vocabulary, similar to Library of Congress Subject headings. The NebraskaMap metadata editing tool allows the metadata creator to choose from a pick list of 19 thematic categories. These categories were derived from the ISO 11915 standard.

*Spatial domain* – Describes the geographic location of the spatial data entered in latitudinal and longitudinal decimal degrees.

*Metadata reference information* – Names the standard used to create the metadata.

The out-of-the-box settings for searchable fields are title, abstract, and data theme. The default setting for the display of the results is “relevancy,” with the ranking based on a ratio of characters in the matching term/terms to overall number of characters in the field.

### **Outcomes of Goal One – Searching and Display**

The NebraskaMap Partners Committee’s portal development taskforce began by exporting approximately 400 metadata records from the Nebraska’s Department of Natural Resources spatial data repository. The taskforce’s librarian used these records to demonstrate how the relevancy ranking worked, and how to enhance search skills by the use of Apache Lucerne query syntax and Boolean operators. To address the issues of discovery by non-specialized users, the librarian encouraged metadata creators in the group to include common descriptive terms in the metadata’s abstract, as well as spell out initialisms used for agency names. For example, include “soil” or “soil survey” in the description of a SURRGO record, and spell out what it stands for (Soil Survey Geographic database).

### **Outcomes of Goal Two – Customized Metadata Editing Tool**

The reasoning for developing a customized metadata editing tool instead of using the out-of-the box configuration was twofold. The first was to make it easy to contribute metadata directly into the portal’s metadata database. Second, the librarian wanted to enhance searching and retrieval functionality.

#### Creating Metadata Made Easy

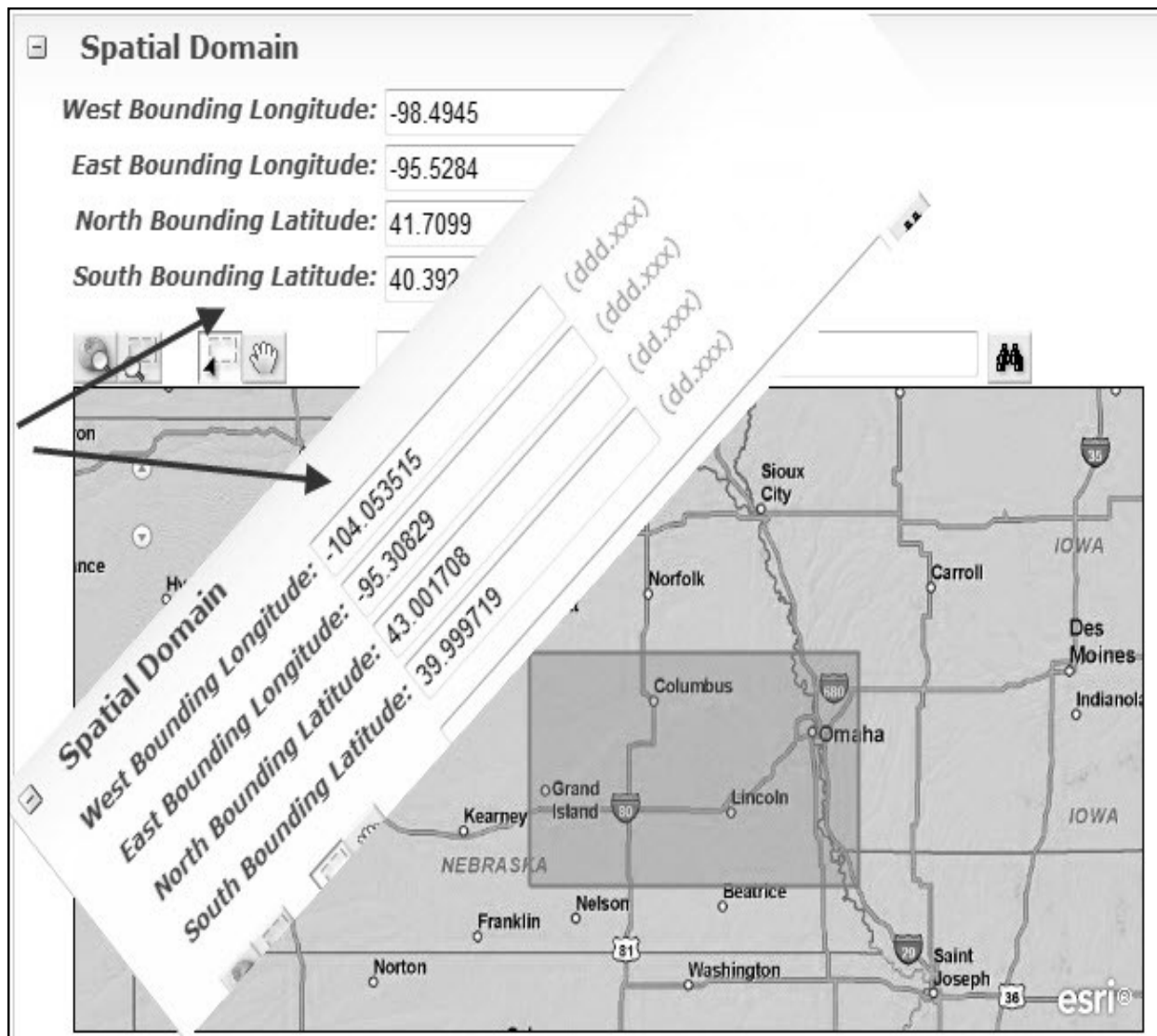
The ESRI portal software allows registered users to pick from a list of GIS metadata standards. Upon selection, the metadata editing tool activates and is similar in functionality to the editing tools included in desktop ESRI products, such as Arc-Catalog. The portal editing tool includes a validation function that runs a software application developed by Peter Schweitzer, the U. S. Geological Survey’s metadata guru. The application verifies compliance to the standard by parsing the created metadata against a template of the standard. If required fields are missing, or there are errors, the

portal will not accept the metadata.

At this stage of the development, the Federal Geographic Data Committee (FGDC) was still recommending metadata creators use the Content Standard for Digital Geospatial Metadata (CSDGM), the predecessor to the NAP standard. The CSDGM has over 300 elements, and many in the GIS community consider it bulky and out-of-date. The librarian was very aware that many metadata creators found it exhausting to try to validate so many fields using any of ESRI’s metadata editing tools.

With this in mind, the librarian encouraged the NebraskaMap Partners Committee portal taskforce to develop a customized metadata editing tool that required a minimal number of fields to validate, but would be interoperable with other geospatial portals that were still using the CSDGM on their sites. Knowing what would become the minimum required fields in the NAP standard, the librarian selected fields in the CSDGM that were similar. These then became the required fields in the customized Nebraska FGDC standard. The remaining CSDGM fields were available as optional fields. Thus, the metadata creator has the option to provide as complete a metadata recorded as needed, without being discouraged by validation problems. Knowing the NAP standard was on the verge of approval, the portal developers planned to offer both standards in their metadata editing tool, making sure that any added functionality would continue to be available in both options.

To simplify the validation process even further, the NebraskaMap Partners Committee portal taskforce wrote coding that filled in as many of the required fields as possible, either with information from the user’s profile, or from a pull-down menu. For example, the ISO 11915 required thematic categories are listed in a pull-down menu from which the metadata creator selects the categories that best describe the spatial data. Fields for status and completeness are also created in this fashion. Fields that automatically populate are the same as in the NAP version of the standard, using information from the users’ profiles. The required field for spatial domain automatically populates with the latitude and longitude bounding coordinates for Nebraska, a feature unique to the Nebraska FGDC standard. The editing tool (Figure 2) allows a



**Figure 2.** The NebraskaMap editing tool automatically populates with the latitude and longitude bounding coordinates as the metadata creator zooms in on the data's location.

metadata creator to zoom-in on a web mapping service, select the data's location, and the bounding coordinates will change accordingly.

### Improving Discovery

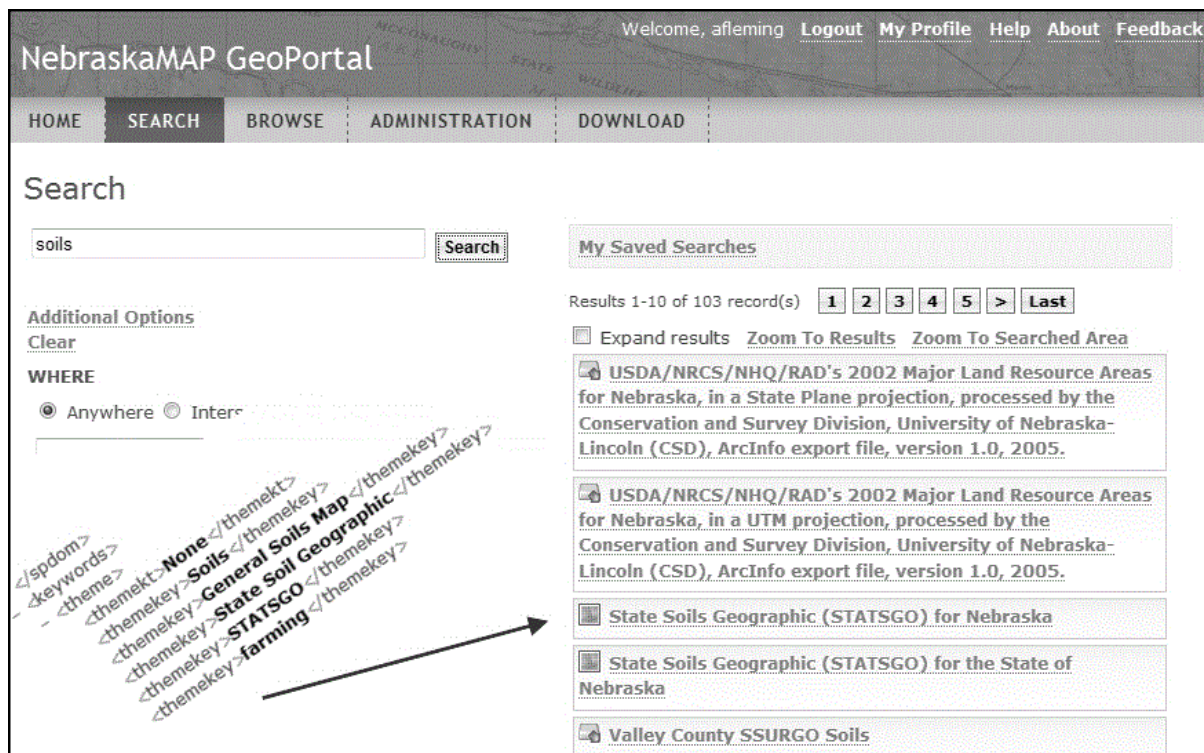
The librarian wanted to enhance searching and retrieval properties of metadata created with the Nebraska FGDC standard. To improve discovery by non-specialized users, two new elements were added to the searchable fields of the customized standard.

*Purpose*, the “why” the spatial dataset was created, is the field in which the metadata creator describes how the spatial data may be used. Metadata creators are encouraged to give practical examples, such as, “centerline street data may be used by realtors for the development of a web mapping

service showing homes for sale.”

The purpose field from the metadata record for SSURGO soils is another example of how to give practical use information. Examples of information that can be queried from the database are: available water capacity, soil reaction, salinity, flooding, water table, and bedrock, building site development and engineering uses, cropland, woodland, rangeland, pastureland, and wildlife, and recreational development.

The second element is a “free-term” keyword option. This option allows for the expansion of descriptive terms beyond the 19 required by the ISO 11915 standard. For example, adding the terms “soil,” “soil type,” or even “dirt” to the searchable keyword field of the metadata record for the SURGO database increases the likelihood that non-



**Figure 3.** Free-term keywords create “search savvy” records.

specialized users would retrieve the record (Figure 3). On the other hand, retrieval by browsing records indexed under the controlled vocabulary heading of “agriculture and farming” is less likely by a non-specialized user, and would be too time consuming for efficient retrieval.

### Outcomes of Goal Three – Metadata Workshops

As part of the two-year project, the portal development taskforce planned a series of training workshops using NebraskaMap. The project administrator, a representative from CALMIT, and the librarian developed and conducted the workshops. Each training session lasted roughly two hours with another hour set aside to assist attendees with introducing their metadata to the portal. Attendees were asked to provide assessments of the software, portal, and web. These sessions were held on July 28, July 30, August 3, and August 5 of 2010 on the UNL campus in Lincoln, Nebraska. Topics addressed included registering for the portal, editing metadata in ESRI’s ArcCatalog for the purpose of importing into the portal, the Nebraska FGDC standard, and searching for metadata. The last part of the training allowed time for attendees to import

or create their own metadata.

To address the issues of discovery and improved search results by non-specialized users, the taskforce’s librarian introduced additional components to the traditional metadata training module. These included information on how metadata works, how to construct a search that will get the information wanted, and the importance of standards. Each field in the Nebraska FGDC standard was described and explained. The librarian constructed different search scenarios to demonstrate how relevancy ranking works, and then gave examples of how to construct the title, abstract, purpose, and keyword fields to improve discovery. In addition, the librarian gave attendees refresher tips on how to use Apache Lucerne query syntax and Boolean operators to improve search techniques. The goal of the training sessions was for attendees to have enough experience to be able to import a metadata record from ESRI’s ArcCatalog, or to create their own and submit the metadata for approval. To date, 77 additional metadata records have been added to the NebraskaMap metadata database by workshop attendees.

## CONCLUSION

Librarians are experienced in the development of online library catalogs. This experience brings a unique perspective to the development of spatial data catalogs. Librarians provide expertise in three areas. First is their understanding of how coding within the metadata records effects the display and retrieval of query results. Second, their knowledge of Apache Lucerne query syntax and Boolean operators improve searching skills. Third, their familiarity with cataloging standards for library resources, such as Anglo-American Cataloging Rules (AACR2), leads to the understanding of geospatial metadata standards. Incorporating a librarian's perspective during the development of a geospatial portal increases the functionality of the metadata catalog, and thus, increases the chances of discovery of spatial data by both experienced and novice users.

## WHAT'S NEXT?

The NebraskaMap Partners Committee continues to support the portal on a volunteer basis, with hopes of permanent funding from the State of Nebraska during the next budget cycle. The UNL Libraries continue to participate in the project through the teaching of workshops on metadata, customizing the portal's metadata editing tools, and monitoring contributed metadata.

## REFERENCES

- Maguire, D.J., and Longley, P.A., 2005, The emergence of geoportals and their role in spatial data infrastructures: Computers, Environment and Urban Systems, v. 29, p. 3-14.
- Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture, Soil Survey Geographic (SSURGO) Database: [http://gep.frec.vt.edu/pdfFiles/Metadata\\_PDFs/1.3Metadata-SSURGO%20SOILS-Section1.pdf](http://gep.frec.vt.edu/pdfFiles/Metadata_PDFs/1.3Metadata-SSURGO%20SOILS-Section1.pdf) (accessed July 6, 2012).
- Tait, M.G., 2005, Implementing geoportals: applications of distributed GIS: Computers, Environment and Urban Systems, v. 29, p. 33-47.
- Wade, T., and Sommer, S., 2006, A to Z GIS: An illustrated dictionary of geographic information systems: ESRI GIS Dictionary: <http://support.esri.com/en/knowledgebase/Gisdictionary/browse> (accessed July 6, 2012).
- Zink, L. (ed.), NebraskaMap: <http://www.nebraskamap.gov> (accessed July 6, 2012).



Published in GEOSCIENCE INFORMATION SERVICES: "Peak" Performances. Proceedings of the 45th Meeting of the Geoscience Information Society, October 31-November 3, 2010, Denver, Colorado USA. Edited by Janet E. Dombrowski; Proceedings, Volume 41 (Geoscience Information Society, 2010)

Copyright 2012 by the Geoscience Information Society. Used by permission.