



LTER Databits

Information Management Newsletter of the Long Term Ecological Research Network

01001100 01010100 01000101 01010010

Spring 2012

The Spring 2012 issue of Databits is focused on highlighting geospatial activities at both individual and cross-site scales, as well as identifying resources to help manage the large collections of GIS and imagery data at LTER sites. The Information Management GIS working group provides updates on their recent workshops and their 2 cross-site projects: LTERMaps and the GeoNIS.

Feature articles include coverage of the Maps and Locals (MALS) cross-site project, the Malpai Portal in New Mexico (a prototype for the organization and discovery of geospatial data), and a collection of selected geospatial projects from across the network. In addition, there are two related articles that provide background information on geospatial metadata standards and resources to help LTER sites document their spatial data and contribute it to the network's data catalog. Hopefully these articles will help to inform you on the variety of projects happening in the spatial side of data management and research.

We hope you enjoy the images and internet mapping websites we found as examples of some of the latest technologies in visualization, real time data, space imagery, and other interesting cartography/art using GPS and on-line maps. The Interactive Cartographic Almanac provides a tool for LTER members to make cartographically pleasing map images for talks, publications, and websites. Jamie Hollingsworth did a nice job putting this package together for the LNO.

The IM Committee Co-Chairs have prepared a commentary about the recent discussions around the network and NSF about online availability of LTER Data. They provide some recommendations from IMExec for making data more easily discoverable, accessible, and usable.

While not geospatial centric, we have included two articles on content management systems. One is about general organization and layout of websites, and the other highlights a cross-site effort to develop a framework for LTER website development using Drupal software. This effort is helping many of the LTER sites move to a database driven development framework for their webpages, and the effort is paying off in shared expertise, tools, and connection to an Open Source community broader than LTER.

Lastly, we offer a listing of several workshops and meeting that are coming up in the next few months, including several of interest to those who would like to learn more about GIS.

Theresa Valentine (AND) theresa.valentine@oregonstate.edu and Adam Skibbe (KNZ) askibbe@ksu.edu

Spring 2012 Co-Editors

Featured Articles

Maps and Locals (MALS): A Cross-Site LTER Comparative Study of Land-Cover and Land-Use Change with Spatial Analysis and Local Ecological Knowledge
 The Malpai Portal
 Preparing Spatial Data and Associated Metadata for the GeoNIS
 Considerations for making your geospatial data discoverable through the LTER metadata catalog.
 An update on LTERMapS: Phase 2
 The GeoNIS: Adding Geospatial Capabilities to the NIS

Commentary

Are LTER data online?

News Bits

Selected Geospatial Data Projects and Site News

Good Tools And Programs

Interactive Cartographic Almanac
 The Drupal Ecological Information Management System (DEIMS) As A Tool For Many Tasks
 Organization and Layout for an LTER Web site using a Content Management System

Good Reads

Images and Internet Mapping Websites
 How Institutional Factors Influence the Creation of Scientific Metadata
 A Comparison of two information management systems

Calendar

Events: Summer/Fall 2012

Featured Articles

Maps and Locals (MALS): A Cross-Site LTER Comparative Study of Land-Cover and Land-Use Change with Spatial Analysis and Local Ecological Knowledge

edit

Hope C. Humphries (NWT) and Patrick S. Bourgeron (NWT)

The Maps and Locals (MALS; http://www.lter.uaf.edu/bnz_MALS.cfm) project started in 2009 as a collaborative effort funded as social science supplements for 11 participating LTER sites (AND, ARC, BNZ, CCE, CWT, GCE, JRN, KBS, KNZ, LUQ, and NWT) to investigate changes in socio-ecological systems using a mixed methods comparative approach. Other LTER sites and groups have been involved at various levels. MALS coordinators are Gary Kofinas, BNZ & ARC; Robert (Gil) Pontius, PIE; and Nathan Sayre, JRN. The specific objectives of MALS are to: (1) use spatial representations of land cover and land use to identify patterns of landscape change in regions in and around LTER sites; and, (2) integrate local ecological knowledge (LEK) and other existing social data into theories and models of social ecological change and their implications for human livelihoods. LTER sites participating in this program of research emphasize these activities to varying degrees. Cross-site comparisons are being conducted to develop methods and questions, test hypotheses over larger scales, and set the stage for cross-site comparative studies. Three workshops were held to coordinate MALS activities. A three-day training/planning workshop will be held in 2012 that provides LTER investigators, LTER graduate students, and others with a theoretical orientation, practical skills, and the research tools to document local ecological knowledge and integrate that knowledge with spatial analysis and other forms of scientific data to understand socio-ecological resilience.

Mapping

In the initial phase of MALS, research was primarily focused on cross-site land-use land-cover change analyses. Each site has assembled or is assembling a time series ($n=2$) set of maps that represent known biophysical, infrastructural, and land-use changes in its region. At each site, maps are used both as corroborating data and as research tools for use to collect local ecological knowledge. At the network scale, maps are collected from several sites for development and application of methods for spatial/GIS analysis. A database of land category maps is being compiled from each participating site from at least two points in time. Maps from each site show two or more land categories that are overlaid on a single raster grid to facilitate statistical analysis. If maps are available from only two points in time, changes are characterized over one time interval; if maps are generated from more than two points in time, analyses can detect whether the process of land transformation has been stationary across more than one time interval. Metrics are also being developed for cross-site analysis; for example, to measure the level of stationarity across sites in a manner that sites can be ranked from less stationary to more stationary. The maps will be designed for use in conjunction with the LTER's web-based map browser, LTER MapS (<http://www.lternet.edu/map/>). In interpreting MALS results, Pontius and Millones (2011) concluded that it is more useful and simpler to compare maps in terms of two summary parameters, quantity difference and allocation difference, than to follow the previous paradigm that compares the agreement between maps to the agreement that could be expected due to randomness (Kappa indices). In this respect, MALS is inducing a transformative shift in the conceptualization and mathematics of map comparisons. The MALS effort advances the LTER agenda in two respects: 1) creation of the database, and 2) development of methodology.

Local knowledge documentation and social data

For the documentation of local knowledge, individuals and classes of informants are identified for each site, with an emphasis on people who have had continuous or regular familiarity with specific places over long time periods (10-50 years, or potentially more through ancestors). Such familiarity may involve direct management of a property or repeated regular visits for specific purposes. Considerable information about local knowledge has already been documented in many regions through past and current research projects, and to the extent possible, participating LTER sites draw on these data to meet the needs of this comparative study. Additional new data are being collected where opportunities and resources are present. The compiled data serve both to generate hypotheses and as a means of corroborating and illuminating existing data.

MALS work at NWT LTER

The MALS project at NWT integrates spatial data and social data into ongoing efforts that test components of the ISSE framework (<http://www.lternet.edu/planning/>). For this work, we have identified an area that ranges from mid-elevation (2500 m) to the alpine zone (up to 4000 m), which encompasses NWT, including the alpine and subalpine zones. In addition, a fairly large portion of Colorado Front Range forests falls within this elevation range, where substantial social and economic changes are taking place. Our initial MALS efforts focused on the area surrounding Niwot Ridge and the adjacent Green Lakes Valley. An additional site, the nearby small town of Nederland and its surroundings, is representative of many mountain towns in the Colorado Front Range that had an early history in resource extraction but have recently experienced residential development in the wildland-urban interface.

NWT human impacts and local knowledge

Although the subalpine, treeline, and alpine ecosystems of Niwot Ridge and the Green Lakes Valley appear pristine, they have experienced impacts from a variety of post-settlement human activities. A chronicle of such activities at NWT was compiled based on historical documents and interviews with long-time residents in the area (Komarkova et al. 1988). To identify key drivers of socio-economic changes in both areas, we will continue the process of acquiring census data, land transfer data, and land survey notes.

Settlement of the area began in 1861 with mining in the Green Lakes Valley, including the area surrounding Lake Albion, where the Albion townsite housed as many as 200 people in the late 1800s. The town was one of the country's highest and most remote settlements at the time, but was abandoned by 1910; some of its buildings are still standing (Figure 1). Dam construction occurred in the Green Lakes Valley and the valley to the north of the ridge beginning in the late 1800s, with impacts that included road building, tree cutting, borrow pit excavation, and subsequent dam enlargement (Figure 2). Forests in the area, primarily at lower elevations, were cut to supply timber for mining activities. The early history of the town of Nederland and the surrounding area was associated with resource extraction, primarily mining, including several cycles of booms and busts. The economy now depends heavily on recreation and tourism. In recent years, housing density has greatly increased in the wildland-urban interface in Nederland and throughout the Colorado Front Range.



Figure 1. Albion townsite just below Lake Albion in Green Lakes Valley. Lower left photo taken pre-1890 (source: Boulder Carnegie Library, 213-1-3). Upper right photo shows remaining structures in 2008.

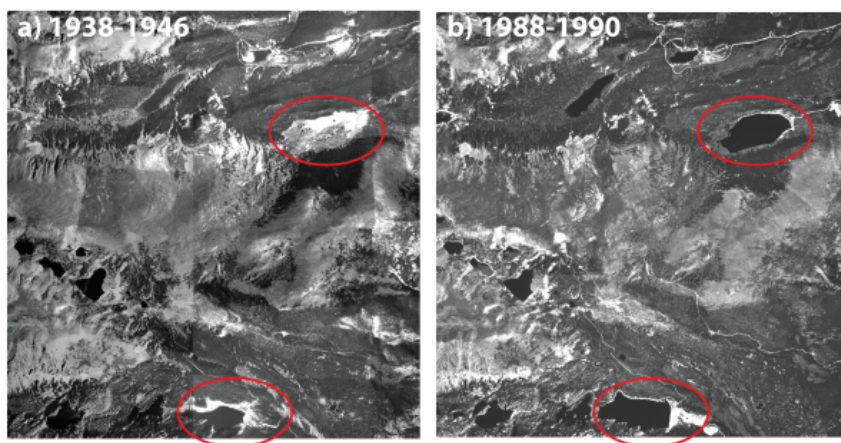


Figure 2. Composite aerial photos of Niwot Ridge, showing changes over time due to road and dam construction in upper right (Lefthand Reservoir) and lower center (Silver Lake).

Climate change impacts to the area include earlier snow meltout, which changes the timing and amount of runoff; this is important because the Green Lakes Valley provides water to the city of Boulder. Documentation of changes in land cover over time in our study will become especially important in the future as the current mountain pine beetle outbreak in the Colorado Front Range intensifies (<http://www.rockymountainnews.com/news/2009/jan/17/colorado-pine-beetle-i...>). The expected widespread mortality of pines will have a pronounced impact on land cover in the region, which in turn is likely to have a strong effect on future land use decisions.

NWT land cover mapping and analysis

At the beginning of the MALS project, Niwot had little in the way of orthorectified imagery, and no satisfactory land classification that covered the area of interest. We have put a lot of effort into acquiring and orthorectifying imagery for a number of time slices ranging from 1938 to 2008. However, they differ in type, resolution, and quality, and there is a lack of consistency in pixel values within and between images. For this reason, we are generating land-cover polygons by digitizing them. We are in the process of acquiring aerial photographs for Nederland to add to existing imagery. Polygon digitizing is underway for three time periods in each location to produce the land-cover categories forest, non-forest vegetation, rock/soil/ice, water, and anthropogenic features. Anthropogenic features and water in the area surrounding NWT have increased in cover over the period 1938/1946 to 1988/1990 due to dam construction and enlargement (Figure 2). MALS spatial statistical methods (Pontius et al. 2004) will be applied to analyzing changes in land cover over time for categorized land-cover maps in both NWT and Nederland.

Komarkova, V., A. Peters, G. Kamani, W. Jones, V. Howard, H. Gordon, and K. Southwick. 1988. Natural recovery of plant communities on disturbance plots and history of land use in the Niwot Ridge/Green Lakes Valley, Front Range, Colorado. Niwot Ridge Long-Term Ecological Research Working Paper 88/1.

Pontius, R.G., Jr. and M. Millones. 2011. Death to kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing* 32:4407-4429.

Pontius, R.G., Jr., E. Shusas, and M. McEachern. 2004. Detecting important categorical land changes while accounting for persistence. *Agriculture, Ecosystems and Environment* 101:251-268.

The Malpai Portal

edit

Ken Ramsey (JRN)

The Malpai research project uses ESRI open source geoportals to share and publish geospatial data and services with project members and the public. A public geoportal is used to share data and services that are publically available. A private geoportal is used to share restricted access data and services with selected project members. The private portal web services require authentication and authorization (group membership) prior to accessing its data and services. The Malpai Portal is an example of a public geoportal.

Geoportals can be used to provide access to all forms of data and web services, not just geospatial data and services. Any URL accessible resource can be cataloged and made available using spatial or keyword search. Geoportals are populated using XML formatted metadata files. Content metadata files can be formatted using FGDC or ISO XML formats. If content metadata contains bounding coordinates, the published content can be accessed using spatial search capabilities. Otherwise, the content is only accessible using keyword searches and has no spatial functionality (.kmz, dynamic maps, etc.).

The Jornada LTER site is implementing additional geoportals to provide access to Jornada data and services. The services will provide access to 'live' data using data, metadata, image, and map web services. Data and metadata services will be delivered using the Drupal content management system and the image and map services will be delivered using ArcGIS Server.

Potentially, the LTER NIS web services and workflows (<http://nis.lternet.edu/>) could be integrated with a geoportal and LTERmapS (<http://www.lternet.edu/map/>) map services and the LTER SiteDB (<http://www.lternet.edu/sites/>) to serve as the data portal for the LTER NIS. The Geo.Data.gov website provides access to national geospatial data and services and is a good example of a customized geoportal interface that provides access to multiple types of data (see Browse tab @ <http://geo.data.gov/>). Additionally, geoportals can reference other portals and geoportals to provide seamless access to data and services from one interface.

The remainder of this article describes the public Malpai Portal (<http://jornada.nmsu.edu/data-catalog/malpai/portal>). Feel free to explore the Malpai Portal online, but be aware that the portal is being populated and enhanced, so the content and functionality may not match the descriptions below. We encourage you to register, login, and provide feedback on the portal and/or content.

Home tab:



From the Malpai Portal home page, you can view all records (currently 27) in the portal by selecting the Search button with an empty search term box.

Search tab:



Alternatively, enter a search term (e.g., ecological) and select the Search button to search by keyword.



Another alternative is to perform a spatial search using the graphical interface. By selecting the up symbol (⬆) or down symbol (⬇) a user can zoom in or out on the map, respectively. By left clicking on the map and dragging the mouse a user can pan the map. By selecting the WHERE clause (Anywhere, Intersecting, Fully within), the user can specify the spatial relationship between the search results and the map display. The text box and related binoculars icon can be used to quickly navigate to a place name (e.g., city) within the map.



By selecting the plus symbol (+) next to 'Record shown from ...', the user can select additional portals to search in addition to the Malpai Portal. The referenced sites (portals) are defined in the geoportal configuration. Once the search has been performed, the user can select the additional sites individually to update the search results pane.

Firefox - Malpai Portal | The Jornada

Home - Malpai Portal

Malpai Portal

Tags: Data Catalogs

Malpai Borderlands Portal

Login Register Help About Feedback

HOME SEARCH BROWSE DOWNLOAD LAUNCH MAP VIEWER

Search

ecological Search

Results 1-10 of 1244 record(s)

Expand results Zoom To Results Zoom To Searched Area

USFS Ecological Subregions

Ecological Footprints of Nations

USFS Ecological Subregions

Canada's Ecological Values within Intact Forest Landscapes (Data Basin Dataset)

USDA Forest Service, California Ecological Sections (Data Basin Dataset)

Commission for Environmental Cooperation (CEC) Ecological Regions of North America - Level III (Data Basin Dataset)

Commission for Environmental Cooperation (CEC) Ecological Regions of North America - Level I (Data Basin Dataset)

Ecological Condition Index 2000 (Data Basin Dataset)

Forest Park, Portland, Oregon Ecological Health (Data Basin Dataset)

USA Ecological Subregions

See results through REST
API: GEORSS ATOM HTML FRAGMENT KML JSON

Records shown from: ArcGIS.com
Click here to select different site or configure search.

Malpai Borderlands Portal 3 Results

ArcGIS.com 1244 Results

Select sites where the search will be distributed
You may select up to 5 sites.

☒ Malpai Borderlands Portal

☒ ArcGIS.com

Additional Options
Clear

WHERE

☐ Anywhere ☒ Intersecting ☐ Fully within

No matching place was located.

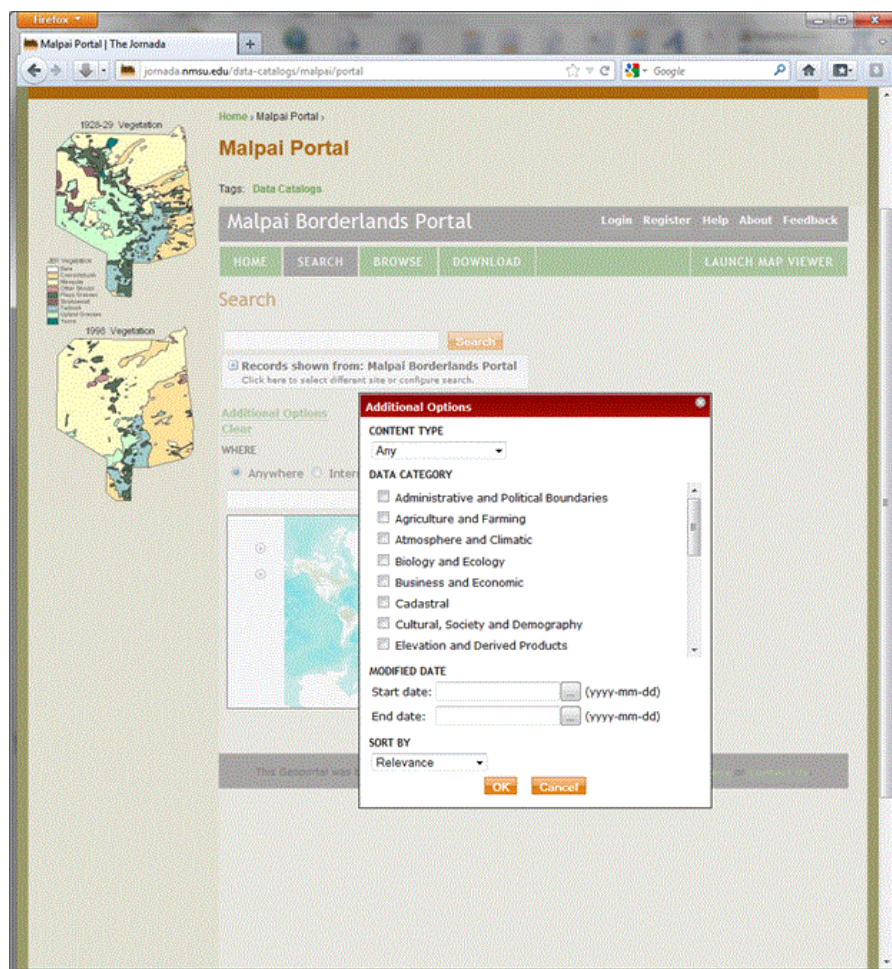
Kansas Missouri

New Mexico Oklahoma Arkansas

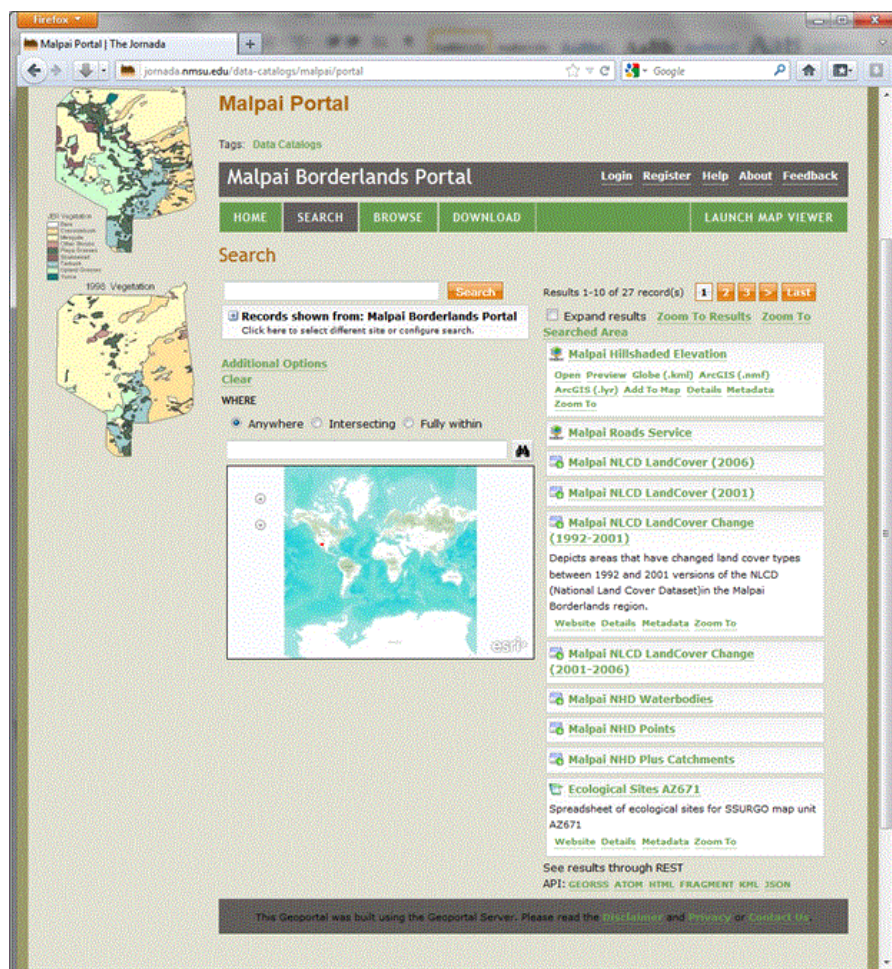
Texas

This Geoportal was built using the Geoportal Server. Please read the [Disclaimer](#) and [Privacy](#) or [Contact Us](#).

Additional search options can be selected using a popup window. Once the additional options have been defined, the options can be applied by selecting the Search button. Additional options include date ranges, categories, and sort options.



All search methods and options can be combined to further refine a query. The search parameters and additional options can be reset by selecting the Clear hyperlink.



Within the search results, the content type is indicated using the following icons:



Documents



Downloadable Data



Live Map Services

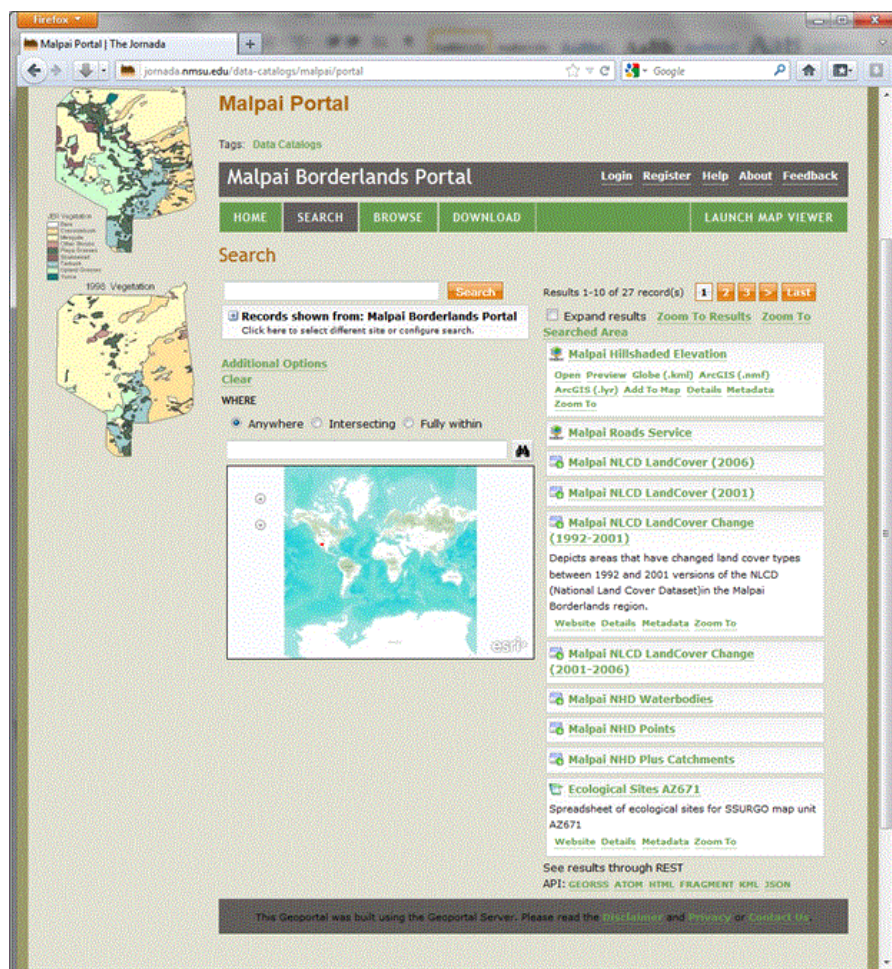


Static Map Images



Resource (e.g., shapefile)

If the title of the content is selected, the content description is expanded. Depending on the content type, different information is available by hyperlink. Alternatively, all content descriptions can be expanded by selecting the Expand Results checkbox. Conversely, all content descriptions can be hidden by deselecting the Expand Results checkbox.



The following listing describes the information currently available, by content type:

Live Map Service:

- Open – opens the REST service page
- Preview – online preview of the map service, including the map service metadata, URL, and HTML code to needed to allow embedding the map service within a standalone web page, as well as user comments (review) and content relationships
- *Note: We have not yet implemented Relationships, which allow content to be related to other content and shared using Preview*
- Globe (.kml) – open or download the GIS layer (.kmz: compressed .kml) used by Google Earth (free software)
- ArcGIS (.nmf) – open or download the GIS layer (.nmf) used by ArcGIS Explorer (free software)
- ArcGIS (.lyr) – open or download the GIS layer (.lyr) used by ArcGIS Desktop software
- Add to Map – opens an ArcGIS Viewer for Flex API internet map with the selected layer added to the interactive map

All search results other than Live Map Service:

- Website – open or download data or document file

All search results:

- Details – metadata in readable format
- Metadata – metadata in XML format
- Zoom To – zoom the map to the extent of the search results

All search results (while logged in):

- Thumbs(👍) – to indicate users comments and approval of the contents

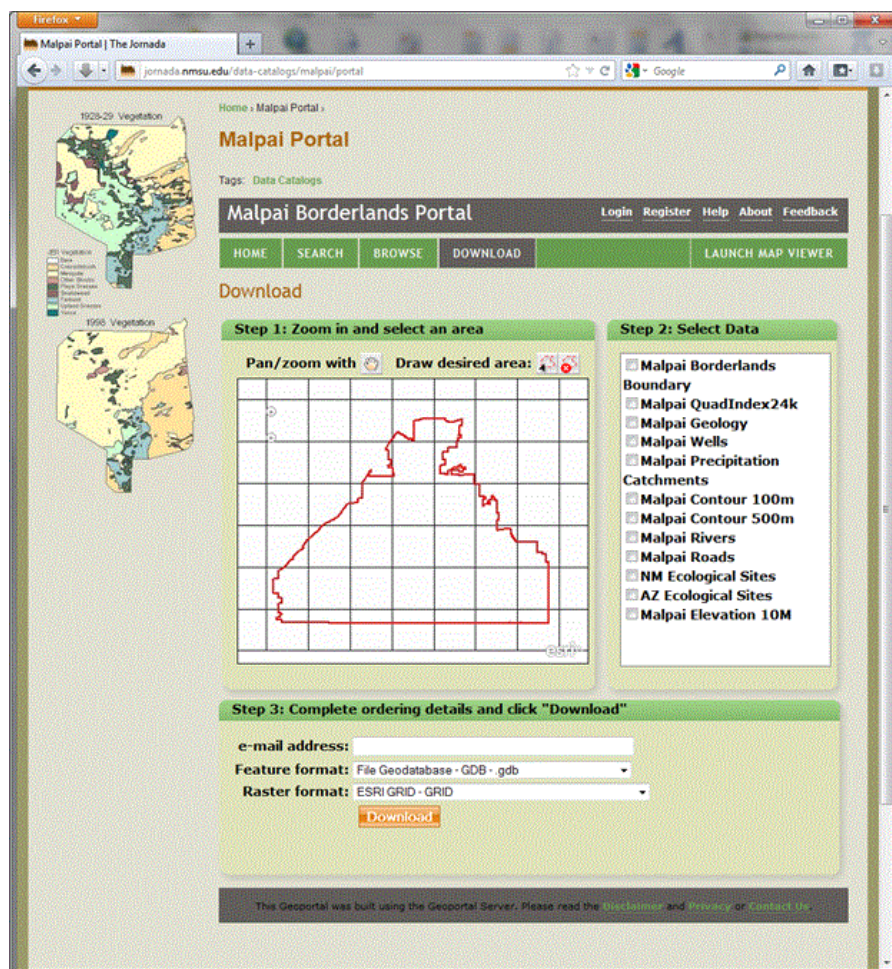
The search results are also available in other formats (GeoRSS, ATOM, HTML, fragment, KML (Google Earth), JSON) using the REST API hyperlinks under the search results.

Browse tab:



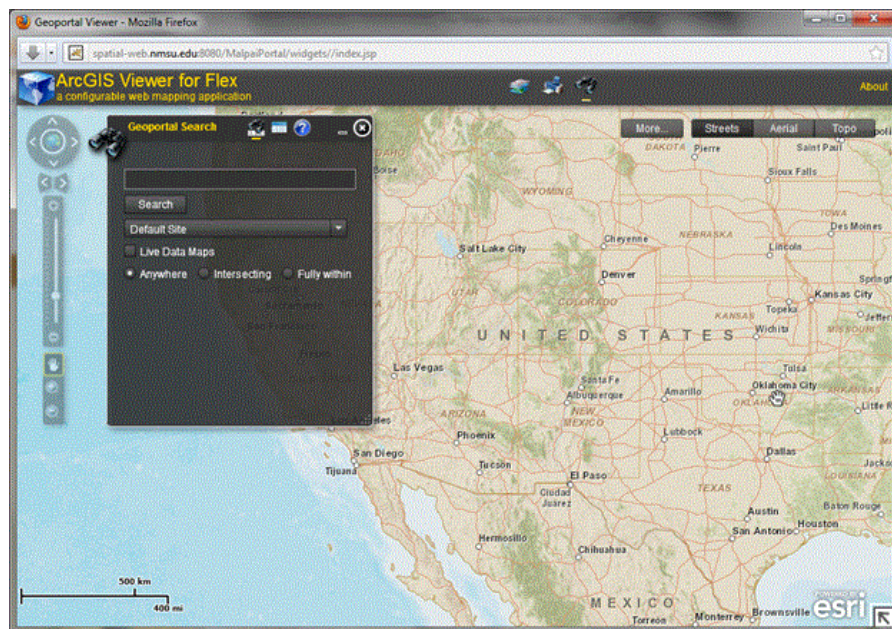
The browse tab allows the user to browse content within the portal by content type or ISO category and filter the results by keyword.

Download tab:



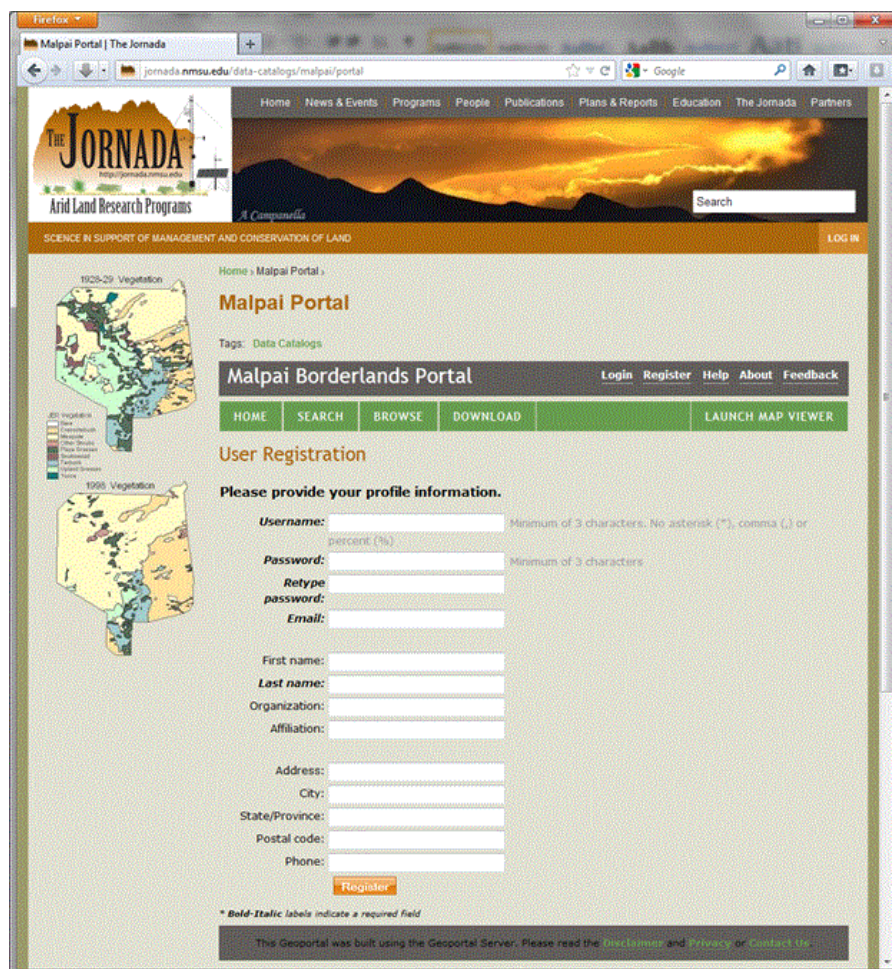
The download tab, allows the user to request data in various formats. The data is automatically emailed to the email address provided by the user. Depending on the data selected and system resources, it may take time to process the request.

Launch Map Viewer window:



An ArcGIS Viewer for FLEX interactive map is available by either selecting the Launch Map Viewer from the portal menu bar or by selecting Add to Map from the search results. The interactive map uses the FLEX API to access data from live map services described within the portal in combination with base layers from ESRI and other online resources (web services).

Register and Login pages:

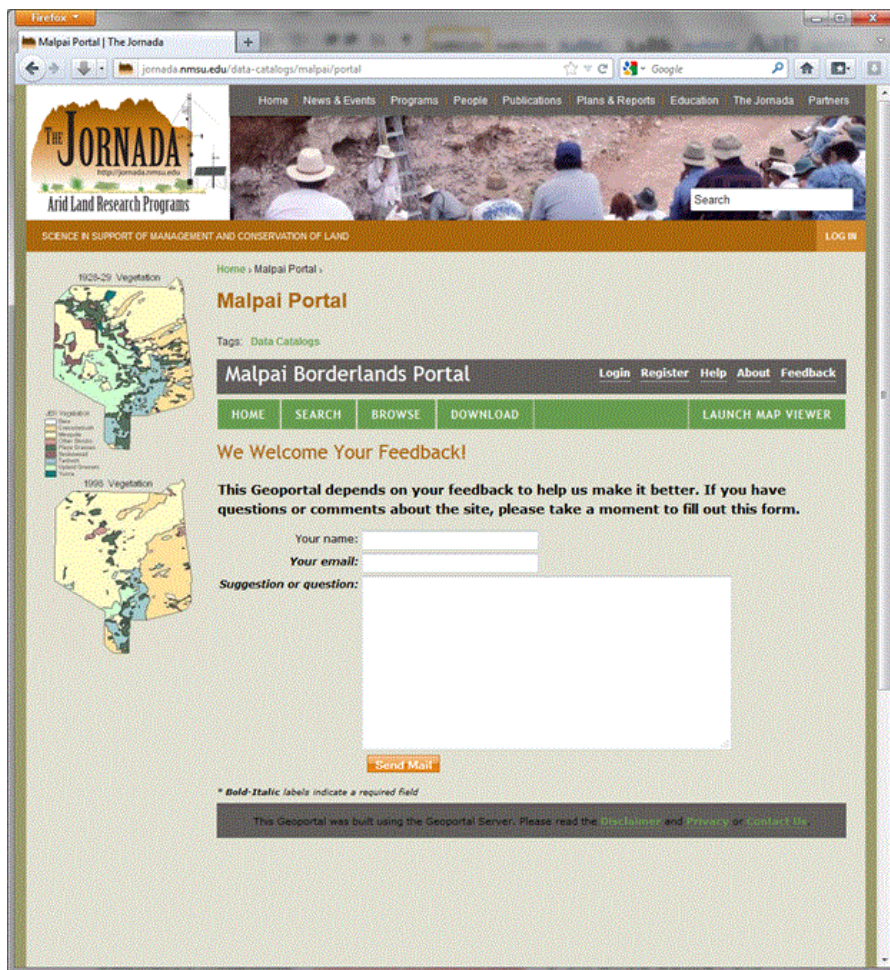


If a user self-registers with the portal, by default the user can save search terms and edit their profile information within the portal.

If a user is subsequently added to one or more groups by a portal administrator, the user will also inherit their group permissions. For instance, if a user is added to the geoportal publisher group, the user will have an Administration link added to the portal menu bar when the user is logged into the portal. The user can then use the administration interfaces to validate and publish new content to the Malpai portal.

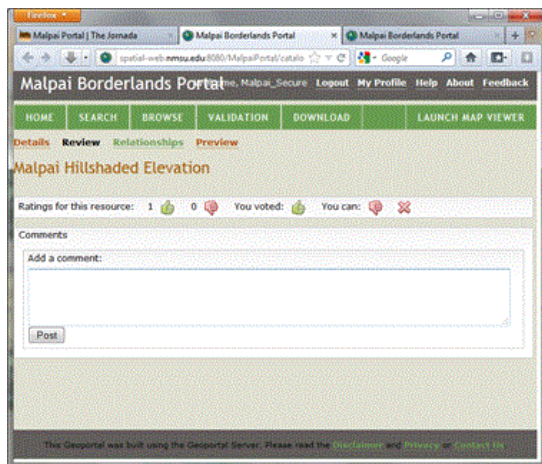
Note: User authentication for the Malpai Portal and Jornada website (Drupal content management system) has not been integrated. Use the Register and Login links within the portal (not the header Login menu) to register and login to the Malpai Portal once you have registered.

Feedback/ Contact Us page:



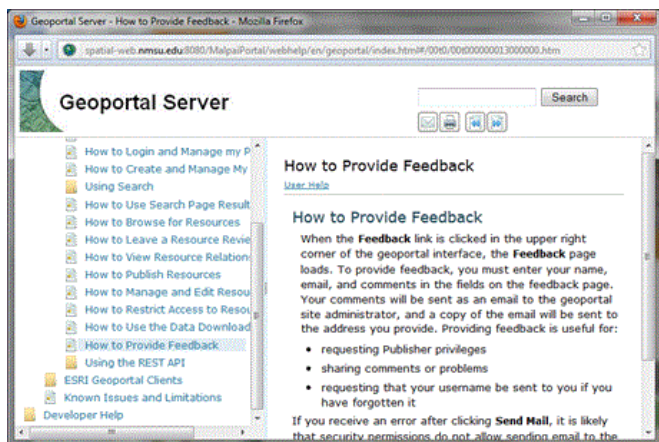
The feedback and contact us page allows users to post feedback about the portal to the portal administrator.

Thumbs:



While logged in, a user can enter content approval (up, down) and comments by selecting the thumbs icon (👍) within any content description.

Help:



The help page allows users to access information related to using and administering the geoportal.

Additional resources:

- Malpai geoportal: <http://jornada.nmsu.edu/data-catalogs/malpai/portal>
- Geo.Data.gov website: <http://geo.data.gov/>
- Geoportal project page: <http://geoportal.sourceforge.net/>
- Geoportal wiki: http://sourceforge.net/apps/mediawiki/geoportal/index.php?title=Main_Page

Preparing Spatial Data and Associated Metadata for the GeoNIS

edit

Theresa Valentine (AND)

A primary charge of the LTER Network is making data products widely available on-line. Traditionally this system has focused on tabular data and left spatial datasets to be organized in separate systems, with different methods and standards for documentation. The GIS Working Group has been working on resources to better integrate spatial data into the network and help site Information Managers, researchers, and students with the creation and access to these datasets and associated metadata.

Metadata is often defined as data about data. Wikipedia defines it as “structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use or manage an information resource”. The standard for geospatial data has been the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM). In 2010, the FGDC endorsed 65 non-Federally authored standards for metadata and are moving towards an international ISO suite of standards (FGDC standards). They recommend the following: “If you have a metadata collection whose contents can be accessed as XML or metadata management software that supports ISO metadata... consider converting your FGDC metadata in XML format into the ISO XML format ... and use an ISO metadata editor tool to create and update it.” If this wasn’t complicated enough, the Standard for LTER is the Ecological Metadata Language (EML). In addition, the commercial GIS software companies have created their own metadata documenting systems, and EML and other metadata standards continue to evolve, adding new versions to keep track of. Keeping up with all the changes can be difficult and frustrating for Information Managers and GIS specialists and it is often difficult to crosswalk between the different standards without losing track of where the most current information is located and compliance with Network policies.

It is however, important to remember that creating valid EML will insure that site spatial data is discoverable in searches of the Network Information System (NIS) and at local sites. This article will provide an abridged guide to the steps needed to create valid EML documents for spatial data, while keeping complete metadata within your local GIS databases, and dealing with the legacy of CSDGM metadata in existence. The workflow will also help the user prepare data packages for inclusion into the GeoNIS database (see article on the GeoNIS).

What is Spatial Data?

Webopedia defines “spatial data” as the following: “Also known as geospatial data or geographic information it is the data or information that identifies the geographic location of features and boundaries on Earth, such as natural or constructed features, oceans, and more. Spatial data is usually stored as coordinates and topology, and is data that can be mapped. Spatial data is often accessed, manipulated or analyzed through Geographic Information Systems (GIS). “ LTER sites generate the above but also have collections of remotely sensed imagery, aerial photography, computer models, historical maps, visualizations, and study site locations. The GeoNIS will be able to incorporate all LTER data that have been referenced to real-world coordinates.

Summary of workflow:

The following is a brief summary of the workflow involved in creating metadata for spatial data. The complete workflow can be found at the project [website](#). The final project will be a complete package of data and metadata products ready for inclusion into the LTER GeoNIS.

1. Create metadata: FGDC or Esri format. The place to start is with a metadata editor for documenting the spatial data. Some commercial GIS software have built in metadata editors (ArcGIS), however you can use stand alone tools. A listing of tools is maintained at the FGDC website: <http://www.fgdc.gov/metadata/geospatial-metadata-tools>. The tool is not as important as the amount of detail preserved and using the LTER Best Practices for EML guidelines for titles, abstracts, methods, placing of URL's, etc...(http://im.lternet.edu/im_practices/metadata/guides) as a guide.
2. Export to FGDC or ArcGIS metadata xml file: The metadata tool should allow the user to export to different formats. The important message is that the documentation needs to be in xml format. One of the following two file type options are needed for the next step:
 - a. FGDC CSDGM xml file
 - b. ArcGIS metadata xml file
3. Customize esr102eml21.xsl stylesheet and prepare for transformation. The stylesheet is the tool you need to transform you xml metadata document into EML. You will need to edit the default stylesheet to meet the needs of your site. This will allow you to automate some of the repeating information, and machine generated identifiers associated with you site. This editing is done with an xml editor such as Oxygen. The site Information Manager should be able to help with this step. A couple of edits that should be considered:
 - 3.a. The Intellectual Rights EML section. Here is where you can express the data usage policies. In the absence of such source in the original

metadata, the stylesheet will populate the intellectual rights with the LTER Network Data Policies. If you think you need special policies, this is the section you need to edit. Further guidance is found near the stylesheet corresponding section.

3.b. The scope for EML's package ID. Since Esri and FGDC are completely oblivious to this identifier, it needs to be hard coded in the stylesheet. Do your site a favor and change the scope to "knb-lter-yoursiteacronym", and you'll save yourself a bit of post editing. The rest of the packageID, the revision and numeric identifier, require post-edit work.

4. Complete the transformation. Once your stylesheet has been updated, you can run the program to transform the data into an EML xml document. There are several options for running the transformation and they are all documented on the project page. There might be some formatting errors at this stage that need to be corrected.

4.a The creator/metadataProvider/contact details. Esri tends to lump the first name, middle and last name in one field and one tag, but EML has separate placeholders for first name (givenName) and last name (surName). Since the XSLT cannot decide which one is the last name, it places all the string into the mandatory last name. Please fix it accordingly. You may have to perform these edits in several places on the resulting EML. Caution, as this will not be flagged as an error by any editor or validating tool.

4.b The identifier and revision part of the packageID. You need to assign these numbers according to the Metacat LTER and site protocols.

5. Run xml document through the EcoInformatics Parser and correct errors. The new EML xml file will need to be checked for errors using the EcoInformatics Parser available at: <http://knb.ecoinformatics.org/emlparser/>. The parser will check your EML document and report any formatting errors. You then go back and correct errors. It's important to note that the parser is looking for formatting errors. It will not let you know if you have problems with your content (spelling, missing data, and incomplete entries).
6. Prepare your data for the geospatial data package. The idea of data packaging is to prepare spatial data sets that can be harvested for ingestion into the GeoNIS geospatial database. Best practices for the contents of a geospatial data package are included in the Best Practices document. http://im.lternet.edu/sites/im.lternet.edu/files/Best%20Practices%20for%20documenting%20geospatial%20data_2.docx
7. Prepare final documents for harvest by Metacat. The data package needs to be placed in the location specified in your EML document. The EML document will be harvested by Metacat and the GeoNIS workflows will download the data package, unpack it, and add it to the GeoNIS geospatial database.

Future plans:

Test the GeoNIS workflows to insure that data package meets requirements and validate spatial data EML for quality. This is critical to insure that the data and metadata can be ingested into a central repository, and that the spatial data is searchable through Network tools. Continue working with LTER information managers and GIS specialists to make sure that the Best Practices reflect the workflow and processes required to prepare valid EML documents for spatial data.

Create a stylesheet for Esri native metadata xml. The current transformation stylesheet should be modified to use native Esri metadata documents. This will remove the intermediate step of translating to FGDC format before moving to EML. This intermediate step can cause content to be dropped when using Esri metadata tools. The FGDC to EML translator is still valid for users who use non-Esri metadata tools and create FGDC CSDGM metadata files.

Automate update process. Most LTER sites would benefit from the development of a tool that would produce a new EML document when metadata is updated in ArcCatalog. This project would require funding for a programmer.

Look at other GIS metadata programs and metadata editors. The current effort has been primarily focused on working with Esri GIS software metadata tools, as most of the LTER sites have access to the software. There are a few sites that are using other programs, and a list of those resources would be beneficial.

Prepare for new ISO format. The FGDC move to an international format will cause some ripple effects through spatial data metadata tools. We have seen some of this through the recent changes in Esri software, as they become more ISO centric. New versions of software will require updates to the stylesheet and best practices.

Link to Project page:

http://im.lternet.edu/project/GIS_document

References:

Wikipedia Metadata standards reference: http://en.wikipedia.org/wiki/Metadata_standards

Webopedia : http://www.webopedia.com/TERM/S/spatial_data.html

FGDC Content Standard for Digital Geospatial Metadata (CSDGM): <http://www.fgdc.gov/metadata/geospatial-metadata-standards#csdgm>

FGDC standards: <http://www.fgdc.gov/metadata/geospatial-metadata-standards#fgdcendorsedisostandards>

Considerations for making your geospatial data discoverable through the LTER metadata catalog.

edit



Inigo San Gil (MCM, LNO)

Currently, some LTER geospatial data are not discoverable through the LTER Metadata Catalog. The native metadata formats that many in the GIS community use to document spatial data are different than the Ecological Metadata Language (EML) required by the LTER Metadata Catalog. In many cases, this native metadata format is the Esri version of the Federal Geographic Data Committee (FGDC) products. Because of this, users exploring the LTER metadata catalog may think that many sites do not work with GIS data. This is not an accurate representation of the spatial data resources available at LTER sites.

The following article highlights considerations and background information to help sites integrate their spatial data metadata into the existing LTER EML based metadata catalog. The article will also provide a brief history of the Esri to EML metadata crosswalk, background on Esri's approach to metadata, details on how the crosswalk works (including customization required), and future direction given the changing metadata picture along with costs and benefits. It is our hope that

some of the knowledge expressed here will aid with future developments of the transformation, and in particular the interoperability among the diverse information management platforms.

Several individuals have worked to develop a crosswalk between FGDC products and EML. The latest group has modified an existing program to take Esri formatted FGDC metadata and transforms it to create an EML document that is in compliance with the EML data structure, and meets LTER metadata best practices. The evolution of the crosswalk (also known as transformation) has been complicated by versioning changes within Esri products, with EML specifications, and with changes in the FGDC standards. Note that the transformation described in the article applies to converting existing metadata records that are stored within Esri ArcCatalog, metavist produced records, or similar tools geared towards the FGDC metadata specifications. All the existing metadata records need to be in XML formatted files. The process for creating documentation for spatial data is summarized in another DataBits article this issue: [Preparing Spatial Data and Associated Metadata for the GeoNIS](#).

The transformation is implemented by mappings the different information placeholders (tags) that exist in the Esri specification and their corresponding tags within the EML specification. This implementation is stored in a XSLT, also known as a stylesheet, which is a limited programming language that uses XML to create directives that map, move and manipulate content within an XML file.

Brief History and Status of the Esri to EML metadata crosswalk

The pre-2005 version of the transformation (ESRI2EML) stylesheet was developed at the Central Arizona (CAP) LTER site, and used by some other sites to prepare EML documents. However, it was designed for a particular LTER site, and sites had difficulty modifying the stylesheet to meet their needs. In 2005/06 an opportunity emerged to work on a suite of tools to make the FGDC based metadata more interoperable with the EML based catalog. The opportunity was framed by a cooperative agreement between LTER and now defunct USGS National Biological Information Infrastructure. Most of the goals of this cooperative agreement targeted leveraging efforts between the programs that would foster interoperability among the scientific communities. One such efforts focused on improving the XSLT based transformation of EML records into FGDC-compliant metadata. The possible mappings and correspondences between the XML-implementation of the FGDC metadata and the EML XML schema, version 2.0.1 were carefully studied. The details of the XML produced by LTER (San Gil et al, 2011) and FGDC records hosted at the NBII metadata clearinghouse were also studied. This in-depth knowledge of both specifications was used to create or enhance the reverse transformation, resulting in an enhanced Esri2EML stylesheet. The base mapping between formats was expanded to offer more flavors of the FGDC related products, which includes the Esri backend specification and the Biological Data Profile. The revamped version of this stylesheet was posted in the LTER Information Managers website [project pages](#) section. The resulting ESRI2EML products helped a handful of LTER sites to produce EML-based geospatial metadata.

Other evolutions affected the stability of the ESRI2EML crosswalk. Back in 2005, the Federal agencies continued pursuing a transition to the North American Profile (NAP) of the ISO backed standards, specifically; the USGS was adopting the ISO19115 and XML implementation ISO19139. At the time, Sharon Shin was coordinating the **practical steps** to finalize the UMLs to give final shape to the North American Profile, and there is no end in sight for the completed transition from FGDC to ISO. Considering the wealth of existing geospatial and other metadata records across the US Federal Agencies, nobody expected a smooth overnight transition. Why is this related to the ESRI2EML crosswalk? At this time Esri started integrating metadata workflows that were ISO compliant. Version 9.+ of ArcGIS had ISO compliant options, with FGDC standards being core to the metadata operations. Version 10.0 (released in 2010) of ArcGIS, the software package for Esri, saw significant changes to the core of their metadata management tools. The ISO standards are now at the core of their metadata management. These new changes rendered the previous ESRI2EML crosswalk and workflow partially obsolete.

The ESRI2EML work was placed in the back burner, with Theresa Valentine (AND), making a push to improve the crosswalk, fixing some bugs and padding gaps. LTER found a renewed interest in geospatial data from several fronts, including sociology, land use change, and projects such as Maps and Locals, which made an impact at the ASM2009. At the same time, EML released a new version (EML 2.1) with no changes in the geospatial sections, but with new constraints that forced a small rewrite of that end of the correspondences. A working group of Information Managers gathered at the LNO in 2010 to revise the LTER EML best practices, and made great strides in providing guidance and recommendations on documenting spatial data. The metadata changes in ArcGIS 10.0 came as a surprise to many in the GIS community. The XML scheme, editing environment, and even metadata were all changed, and as noted above, was based in a large part on the ISO standard. This resulted in some critical momentum to improve the stylesheet, and to document procedures to help sites.

Some background on the Esri approach to metadata

ArcGIS, a product of Esri (Environmental Systems Research Institute, Inc), consists of a suite of tools for working with GIS content including desktop, server, and web based applications. Most LTER sites have access to this suite of tools through connections with universities that have higher education site licenses. ArcGIS also includes an integrated metadata management system accessed primarily through ArcCatalog (the data management component of ArcGIS).

The team examined XML metadata records that are stored in ArcCatalog, for a window into how Esri handles metadata. To summarize: Esri's approach to metadata in the pre-version 10 flavor, was like FGDC on steroids. Esri's XML tags contain the same general tags and structure of the FGDCs standard. However, Esri added a wealth of tags to accommodate metadata that was deemed important for proper data flow using their products and data structure. Esri needs to tag datasets with unique identifiers that enable proper manipulation in databases. Also, Esri added sets of tags that are critical to geospatial functionality, some of them may have been missed by the Content Standard for Digital Geospatial Metadata (CSDGM). The CSDGM is the actual name of the government sponsored metadata representation, commonly refer to as FGDC. The suite of profiles and extensions were the preferred implementation prior to the NAP of ISO19115.

At version 10.0 of Esri's products, the XML representation of the metadata has increased in volume, and the underlining structure changed. FGDC format was dropped in favor of a shortened Esri standard that was critical for implementing a new search function within the software, along with an expanded metadata editing system based on ISO standards. A patch was developed that allowed the importing of FGDC documents and conversion to the ArcGIS Metadata format, along with a stylesheet to export ArcGIS format to FGDC format.

In addition to all the FGDC tags and Esri's own fields, there are ISO-like fields that appear in Esri's XML files. Many times, the new ISO XML tags (or fields) duplicate the same targeted information placeholder that the FGDC side offers. For example, the information about the data "distributor", which is covered at length by the FGDC, is now duplicated in the Esri backend, with the ISO branch that stored "distributor" information. To illustrate this example, refer to Figure 1, where the "Distribution" Information related information groups (tags) are highlighted. Expanding the respective placeholders for the "distribution" placeholders show the parallelisms, creating redundancies in the text. Furthermore, there are many more tags that are duplicated by virtue of merging two synergistic XML specifications such as the FGDC's and ISO19139. Several are visible upon inspection of the figure and a sample, pseudo-XML [schema is available here](#).

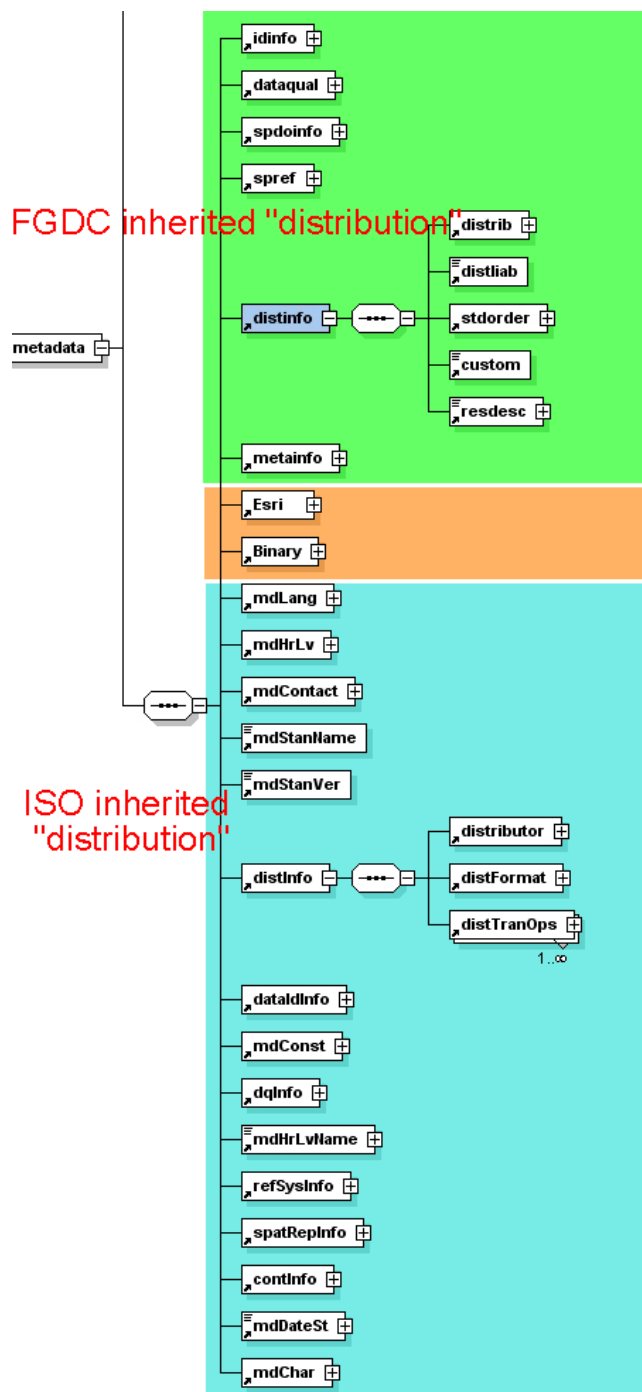


Figure 1. A reduced version of the hierarchical representation of the Esri 10 schema and section from the Content Standard for Digital Geospatial Metadata Workbook

Figure 1 is a screenshot of the Esri ArcGIS Version 10.0 schema, as distilled from an Andrews Forest LTER XML metadata record instance. Esri does not have an XML schema for their metadata available for distribution and the sample was created by filling in a sample metadata document, using all the possible entries, and exporting the resulting XML file to standard XML tools. The diagram is a visual complement to the high level description that follows. The XML root element is "metadata", which is also the root element of the FGDC schema. The top group of XML tags, surrounded by a green background, corresponds to official XML tags from the FGDC XML schema, while the small group of tags, with orange background, was added by Esri. These tags were present in versions prior to ArcGIS 10.0. The bottom group of XML tags, surrounded by aqua blue, is borrowed from the ISO19139 XML schema. A detailed file is available for download at: http://databits.lternet.edu/sites/databits.lternet.edu/files/esri10_xsd_RENAME_EXTENSION_TO_XSD.txt.

At first, Esri's metadata strategy of merging two synergistic standards may seem like a dangerous proposition. Accommodating both standards in this fashion nicely reflects the transition from FGDC/CSDGM/BDP to ISO, but there are clear drawbacks. One such inconvenience is the redundancies entered in the information location. The Esri metadata team was consulted for their insights, without success. Our conclusions about their merging strategies are derived from our own analysis and the use of Esri's tools. Esri is focusing their efforts at the application layer. When you use ArcCatalog, you may choose an ISO view (or skin) for manipulating metadata (default) or an FGDC skin. Both the ISO and FGDC skins have a similar look, but a different set of information is being gathered depending on the form used. The targets of these different forms may be the corresponding XML tags, either FGDC or ISO format. Likewise, you can export the records in both ISO compliant and FGDC compliant formats. In all, Esri assumes that no or few users will be handling the raw XML. Esri treats this XML as the vehicle to manipulate and transform some metadata in the backend, while at the same time; it complies with one of their most important clients, the US government

Details on the ESRI10toEML2.1.0 crosswalk creation process

The LTER community needs EML backed metadata to account for geospatial metadata through the metadata catalog, and many sites are using Esri GIS products to document their spatial data. The XML representations in ArcCatalog vary by use (ISO or FGDC), and may include hybridizations of both tags, purely FGDC, or ISO tags. Keep in mind that future versions of Esri may gradually deprecate the FGDC skin and replace it with the ISO, as the Federal agencies continue the slow transition to ISO-backed standards.

Given time and budgetary constraints, we set out to tackle improvements to the ESRI2EML crosswalk. We started with the aspects of the crosswalk that focused on those records manipulated through the FGDC-skin of Esri. These metadata records include all the legacy (pre-Esri 10) geospatial metadata documents, those documents produced with ArcCatalog-Esri 10 FGDC skin, and the documents that were in a FGDC format (non-Esri). We worked for about a week improving the crosswalk, including a one-day site visit where both invested the day exclusively in finding and correcting bugs and problems associated with the existing crosswalk. Better documentation was also programmed as part of the effort. Guiding resources are discussed in this Databits issue. We used both XMLSpy and oXygen XML editors to improve the standard, as well as ArcCatalog. For validation we used mainly the XMLSpy tools, but also the ecoinformatics parser that performs some extra validation checks. It is noteworthy to say that the previous version of this crosswalk mapped Esri 9 records to EML2.0.1, and in the newer release of EML (2.1), no empty XML tags are allowed. Since the Esri and FGDC tools are very lax, many records lack critical content. Because of the new EML constraints, the checks for content had to be tightened quite a bit.

The resulting stylesheet (ESRI10toEML2.1) was tested during the recent GEONIS working group held in Boulder. During one morning, participants volunteered to test the crosswalk on their site data. The exercise results were very interesting, as we came across uses that were unforeseen, and made the work challenging. A workflow was developed to integrate the new stylesheet with ArcCatalog, so you would have the option to export metadata as FGDC, ISO or EML. However, since ArcCatalog lacks an EML-skin linked to the backend XML product, and our crosswalk does not address the bulk of ISO tags/fields, the results were surprisingly disappointing. The largest problem was that some information entered into ArcCatalog was dropped during the export to FGDC step. The export to EML directly from ArcCatalog (without going to FGDC first) resulted in many un-mapped tags and needs significant work. Each potential process resulted in lost data that was needed within the final EML documentation. The best practices document (http://im.lternet.edu/project/GIS_document) describes the workflow, which has several steps and still may miss some critical metadata. The current best option is to export to FGDC format, and then transform to EML. The workflow could be simplified with a modified direct ArcCatalog to EML stylesheet.

Future directions

Ideally, the stylesheet should be improved in a way that would consider all observed uses of Esri metadata. However, this in practice is not cost effective. Both Esri's backend standards and EML are likely to change. The next version of ArcGIS (10.1) is scheduled for release in the next quarter. While significant changes to metadata aren't expected, there are always the unexpected results of a version change. The main goal of the stylesheet is to aid with the conversion of Esri formatted metadata into EML, and the best practices document is intended to guide the user with the process. It is worthwhile to explore some options for the future.

1. Provide a crosswalk from the ISO fields of Esri 10 to EML. The payoff is sizable. For one thing, it would avoid possible metadata leakage from Esri 10 encoded metadata. Many users may not really want to read any guidelines, and simply hit the "Export as EML", unbeknown of the perils of its limited use. Also, there is some chance that we would use the ISO2EML transform in other contexts.
2. Perform more debugging iterations to improve the quality of the metadata products. No matter how much effort we put into the crosswalk, there is always one more bug, or one more improvement. A list documenting those issues would be good for those who want to keep improving the crosswalk. It would be beneficial to prioritize the fixes, as some may be the effect of local practices, and it's important to keep the stylesheet as generic as possible so that many organizations could use it.
3. Improve the documentation and guidelines. The crosswalk is as good as its documentation. ISO and Esri do not enforce metadata. Mandatory fields are suggested in the interface, but the editing tools do not prevent you from saving and closing records that do not comply with the mandatory rules. EML is stricter, and the user has the right to know the challenges.
4. There are some LTER sites and other organizations that use non-Esri tools to develop metadata for their spatial data. It would be important to identify the users and their tools, and make sure that they can transform their metadata into EML.

Happy transforming!

References

San Gil, I; Vanderbilt, K. V. and Harrington, S. A.

An update on LTERMapS: Phase 2

edit

Adam Skibbe (KNZ) and Theresa Valentine (AND)



LTERMapS (LTER Map Services) was designed to be a multi-phased approach for the development of mapping tools accessible across the LTER network. Phase one of this concluded with the release of a Google Maps based tool for a quick look and SiteDB exploration at <http://www.lternet.edu/map/>. With its completion, the focus turned to phase two, a considerably more robust product for web-based cartography and data exploration.

The goal of phase two is to "Employ a standardized set of data and tools for all LTER sites (DEM, infrastructure, hydrography, structures, and high resolution aerial photography), as well as be modifiable to fit each site's specific needs. In addition to development of analytical tools..., phase two of LTERMapS will also allow for user submitted queries to harvest information and data and will integrate closely with the Network Information System (NIS) modules." Due to the complexity of this project, five LTER sites were chosen as pilots. This gave the team a manageable initial set of sites and data to develop standardized datasets.

A workshop was held at the LNO in November 2011 with the intent of building a beta application around these five pilot sites. The goals for the meeting were:

- Discuss options for integrating spatial data into PASTA
- Develop map templates for a common symbology and cartography for LTER Sites
- Design the backend database, and implement the software and hardware configuration at the LNO
- Develop internet mapping Web application using ESRI tools
- Construct web services and links to existing on-line resources

During the discussion of integration with PASTA it was decided to split LTERMapS Phase 2 into two products, the original on-line mapping tool, and a back end platform for dealing with these data, the GeoNIS. For phase two we constructed what requirements and template for an alpha product. A dedicated server was

set up, with appropriate software, at the LNO with remote access for LTERMapS team members. An Image service for pilot site Digital Elevation Models (DEM) is running on LNO server with all DEM data in the same projection. A JavaScript API skin is in place, and widgets are being added.

In addition to the work on phase two, LTERMapS phase one was upgraded to Google Maps version 3.0, though at the time of this release it has not yet been populated at the LNO. The team assisted the LNO staff by updating latitude and longitude data in SiteDB and updating the web entry form to reflect best practices.

The LTERMapS team met briefly at the GeoNIS meeting in Boulder, and continues to work in three major areas:

1. Development of a standardized backend database, (in collaboration with the GeoNIS project)
2. Application development with the Esri JavaScript API (front end web access to the data)
3. Server software and hardware requirements/support.

Additional pilot sites were added to the project because of interest, and opportunities to meet at the GeoNIS workshop. The following sites have provided data for inclusion into LTERmapS: AND, BES, BNZ, MCM, KNZ, GCE, JRN, NTL, VCR.

The GeoNIS: Adding Geospatial Capabilities to the NIS

edit

Aaron Stephenson (NTL)

Introduction

The LTER Network Information System (NIS, <https://nis.lternet.edu/>) is intended to provide a number of tools and services to promote data access and availability. These include standardized approaches to metadata management and data access, programs and workflows to create and maintain integrated derived datasets, and applications for data discovery, access, and use. These services will be enabled by the Provenance Aware Synthesis Tracking Architecture (PASTA) framework, the core component of the NIS that harvests site metadata and data into the NIS. The initial development of the NIS focuses on supporting well-documented tabular data only, leaving more complex data (such as geospatial data) to a later date. Many LTER sites have considerable geospatial data holdings; at some sites, geospatial data constitute the majority of their datasets. Rather than wait for PASTA to support geospatial data, which at this point doesn't have an implementation date, the LTER GIS Working Group intends to build a geospatial module for the NIS so that these data can be harvested, stored, and be made accessible through the NIS. This module is called the GeoNIS.

Vision

Why a GeoNIS? Location is everything. LTER researchers need to be able to search, access, and discover data sets and related research results across LTER and other areas of the world. Most projects need a geographic framework that helps place the project in context. What soils are similar or different? What is the elevation and aspect? Are study sites similar or do they contrast one another? The GeoNIS will help the LTER network build the geographic framework within and between our sites, and assist with the synthesis process.

The GeoNIS is intended to provide dynamic harvesting and archiving of site-based data and metadata, support value-added products, and include the ability to generate synthetically derived data products. The GeoNIS mirrors the design of PASTA with the additional capability to store and process geographic data. We will test the capabilities of the PASTA data cache to store spatial data before it's ingested into the GeoNIS.

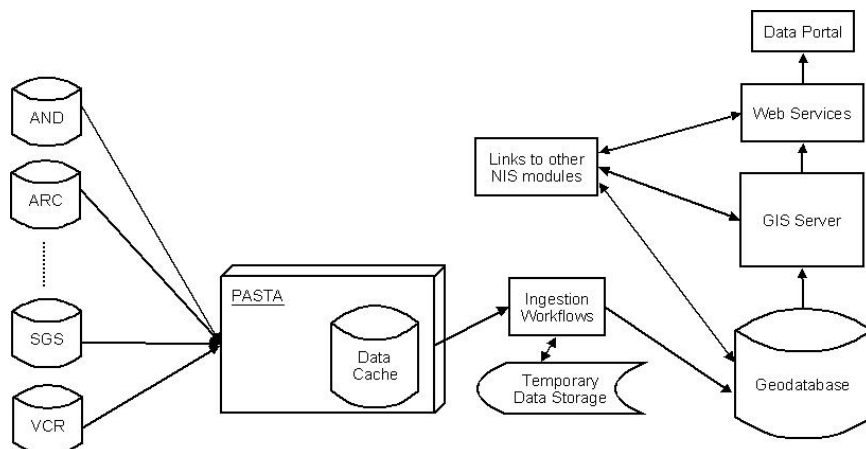
Using automated workflows triggered by an event listener, the GeoNIS will ingest geospatial data from the PASTA data cache into a geodatabase allowing data to be immediately useable by clients. Uses might include interactive mapping, geoprocessing (transforming and analyzing data to produce new data), or just simply making data available for filtering (location or attribute) and downloading only the portion of interest rather than the entire dataset.

Archiving of datasets will be a central piece of the GeoNIS. Each time a new version of a dataset is harvested by the PASTA harvester, that version will be ingested into the GeoNIS geodatabase. The goal is to enable every version of a dataset to be accessible to clients.

By integrating GIS data across sites into one repository, the GeoNIS will provide researchers with the ability to create new products and services, and provide the spatial framework for cross-site science. For example, locations where data collection took place can be coupled with information from external services (such as the Geographic Names Information System) to create a gazetteer that would be used to assign spatial keywords to LTER datasets. Another example is creating maps on demand by assembling LTER and non-LTER GIS resources via web mapping services.

Components

The GeoNIS will be comprised of several connected components, with future links to PASTA.



Site Data:

LTER sites will contribute to the GeoNIS by preparing EML compliant metadata and spatial data packages. The metadata will be harvested similar to other LTER metadata, and links to the associated spatial data packages within the metadata will be used to harvest the spatial data into the GeoNIS. The data packages will include the digital data files as well as GIS specific metadata if available. Access to both metadata formats will ensure that spatial data is included in LTER data catalog searches as well as within specialized GIS applications.

Ingestion Workflows:

Ingestion workflows, triggered by the PASTA event listener, consist of operations that extract, transform, and load spatial data from a variety of file formats into the GeoNIS geodatabase. These scripts will be written in Python so that both ArcGIS and operating system tools can be employed and used to automate the ingestion of site data into the GeoNIS. Through the use of the ArcGIS Data Interoperability Extension (<http://www.esri.com/software/arcgis/extensions/datainteroperability/index.html>), thousands of data formats can be converted and ingested by the GeoNIS.

Temporary Data Storage:

A set of local file folders on the GeoNIS server will be used for the temporary storage of spatial data files while they are being operated on by workflows.

Geodatabase:

An ArcSDE database will store geospatial data (both site data and synthesized data) for the GeoNIS. ArcSDE supports multiuser reading and editing, data versioning, and archiving.

GIS Server:

ArcGIS Server will provide the ability to create, manage, and distribute web services, which can be accessed by desktop, mobile, and web applications. Several kinds of services can be published, including OGC, KML, many kinds specific to ArcGIS (map, globe, image, geoprocessing, etc.), and more. A list is available online at <http://goo.gl/JzAlt>.

Web Services:

A variety of web services will be populated with GeoNIS data, allowing any number of clients to access the data. At first this will be limited to map and image services, but eventually geoprocessing services will also be developed.

Data Portal:

The GeoNIS data portal will provide the ability to discover and access GeoNIS data through a variety of interfaces: thematically grouped links, graphical mapping interface, and a textual search interface. This could be accomplished through the open source ESRI GeoPortal Server (<http://www.esri.com/software/arcgis/geoportals/index.html>) or directly in the NIS Data Portal, or some combination of the two.

Links to other NIS modules:

Other NIS modules could be linked to the GeoNIS, such as ClimDB, HydroDB, or SiteDB, for more in-depth analysis or for mash-up applications.

Additionally, Best Practices are being developed by the GIS Working Group for smooth operation of GeoNIS. Topics include data packaging, attribute definitions, symbology, coordinate systems, data structures, and much more.

Next Steps

One of the highest priority tasks for the LTER GIS Working Group is to request endorsements from IMExec, NISAC, and the Executive Board. Following that, work will begin in earnest on building the GeoNIS. Phases will include the building the geospatial software stack (database, server, web services, and applications), writing workflow programs (including connections to PASTA), building the data portal, and finally creating value-added products. The GIS Working Group invites any interested members of the community to assist with this project, especially members of the web services group, NIS Tiger Teams, and the various network database groups.

Commentary

Are LTER data online?

edit

Margaret O'Brien (SBC) and Don Henshaw (AND), IMC co-chairs

The perception of whether LTER data are online is being formed in part by whether data are discoverable and accessible through a single network portal, and in particular, our Network metadata catalog. Sites have progressively added metadata content in Ecological Metadata Language (EML) to this Metacat catalog since its establishment in 2005. Over the ensuing years, the style of contributions to this catalog has varied enormously. For some sites, the understanding was that even sparse EML was adequate to lead a user back to the site's catalog for data and more information; other sites provided EML that could directly deliver data along with metadata; and a few adopted EML as the basis of local site information systems as well as for network contributions. The bottom line though, is that although our varied practices reflect our local traditions, they have not gotten us closer to becoming a cohesive group. The goal of our network metadata catalog has not been clear and has suffered from lack of attention, and consequently - correct or not - the perception is that LTER data are not easily available online.

Our goal now is to change that perception. LTER data distribution policies are exceptional and our systems are widely admired and copied. The most recent views of the Network Metacat catalog show that only a few percent of datasets have no link to data at all. Many scientists outside the LTER have said "you mean I can just go to your website and download your data?!" Clearly, we have made great strides toward thoughtful and pragmatic data publication practices. However, the completeness of the metadata and the ease in which data are accessible is highly variable and uneven among sites and is in need of immediate improvement.

NSF has made increased data availability a high priority. They have emphasized that the Network catalog should provide access to site data - not just metadata, and that the volume should reflect all work at the sites, including data not intended for PASTA. These two expectations - simplified discovery for all data via the network portal and automated use of "PASTA-ready" data - are not necessarily conflicting or insurmountable, but will take commitment to network goals from all sites.

Here are recommendations from IMExec for making data more easily discoverable, accessible and usable through the network portals:

- Inventory your LTER-funded studies and catalog all site data sets. We still need to develop practices to clearly identify and justify different types of data, including Type II.
- Prioritize the development of data sets for inclusion in the LTER metadata catalog and for PASTA. Priorities should be driven by scientific questions.
- Improve metadata content to improve discoverability
 - Improve data set titles, abstracts, add LTER controlled vocabulary keywords
 - Improve data entity and attribute descriptions, and
 - Adopt network standards for URL construction and location
- Improve EML documents to comply with EML best practices, particularly in the placement of URL links to data sets at the entity level, allowing the network catalog to better represent site metadata and the automated processing of site data sets.
 - Become more familiar with EML best practices
 - Consider automated approaches to generating EML, e.g., DEIMS, Metabase
- Be prepared to simplify or remove web forms or cumbersome logins that might be obstacles to data access, per evolving Network recommendations.

There is obvious benefit to improving metadata and increasing the amount of data available through the network data portal. And while it may require significant resources, the Network became obligated to do so in 2003 when the LTER Coordinating Committee unanimously passed a motion to adopt "a tiered trajectory toward improved IM functionality for synthesis, and the trajectory increasingly incorporates common, structured metadata - the network adopts a general goal of improving each site's position in the trajectory". In the spirit of this tiered trajectory we plan to target certain site data sets to be "PASTA-ready" and establish specific scientific workflows to build value-added data products. As LTER is faced with demonstrating the potential of the NIS and justifying the investment, improvement in the quality of site data and metadata will go a long way towards illustrating its value.

News Bits

Selected Geospatial Data Projects and Site News

edit

The following are highlights from some of the geospatial activities currently happening at LTER sites around the network.

Andrews Experimental Forest (Theresa Valentine):

Recent data acquisition: The Upper Blue River drainage adjacent to the Andrews was flown with LiDAR in Fall 2011. This provides the site with extended 1 meter resolution digital elevation models (DEM) for the bare earth (see figure 1) and highest hits returns, and the raw point cloud data. The bare earth DEM has been used to generate new road and stream GIS layers, as well as help provide more accurate boundaries for gauged watershed studies.

Upper Blue River LiDAR

 HJ Andrews Boundary

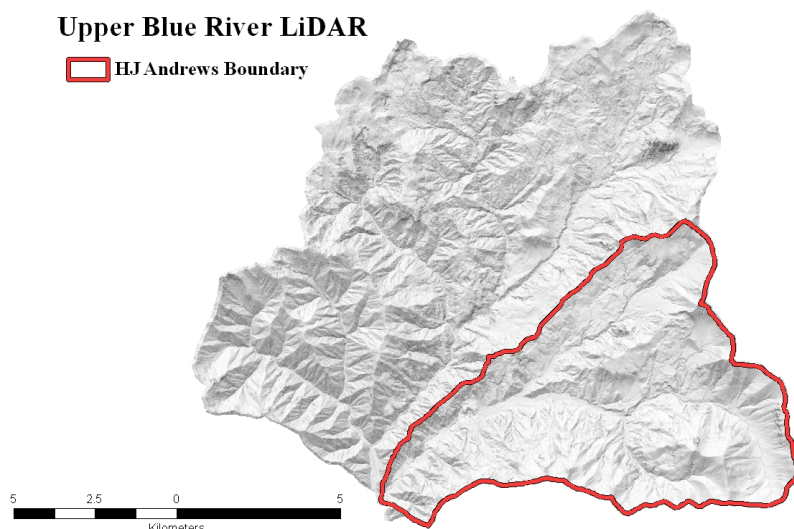


Figure 1. Upper Blue River Lidar

Digital Forest: Spatial Models of Vegetation Structure and Composition: The objective of this project is to spatially model current forest structure and composition of the Andrews. These models and data will then be used by other projects in LTER6 (e.g. water and carbon, modeling) to address questions related to the goals of the LTER. In addition, the spatial models of forest structure and composition will be used to understand how forest structure varies in relation to topography. The initial stage of this project will be to create and evaluate spatially models of canopy height and cover, which can be estimated directly from LiDAR. The second phase will be to model other forest structural features, such as biomass and basal area, which require that data from ground plots be used in conjunction with LiDAR. For more information: <http://andrewsforest.oregonstate.edu/lter/research/lter6/current.cfm?pid...>

Climate Data: Work is underway to develop improved spatial climate data sets (grids) to be used as input for modeling and analysis activities. These include new 1971-2000 mean monthly and annual precipitation grids at 50-m resolution prepared by C. Daly using PRISM (Fig. D-2). The mapping activity served as impetus for the digitizing, cleaning, and organizing of historical datasets collected over the past 60 years at HJ Andrews. Temperature measurements at 50 to 200 new sites within the HJA are being used to develop improved maps of temperature and explore relationships of temperature with topography and cold air drainage. <http://andrewsforest.oregonstate.edu/lter/research/lter6/cl/summary.cfm?...>

Useful modeling links:

iland: individual-based forest landscape and disturbance model: <http://iland.boku.ac.at/tiki-index.php>

RHESSys: a GIS-based, hydro-ecological modeling framework designed to simulate carbon, water, and nutrient fluxes. <http://fiesta.bren.ucsb.edu/~rhessys/>

Baltimore Ecosystem Study (Mark Kather):

Western School of Technology and Environmental Science Baltimore County Public Schools Land Stewardship Project:

On February 2, 2012, sophomores and juniors in the environmental technology magnet program at the Western School of Technology and Environmental Science (WSTES), teamed up to create a land stewardship plan for public school sites in Baltimore County. Teams selected individual schools based on the criterion that it is involved in pursuing Maryland "Green School" status. Students focused in on 5 elementary and 5 middle school sites.

A letter of introduction was prepared by each team and emailed to the green school contact person at the corresponding school. Base maps showing existing topography, roads and parking lots, buildings, streams, and vegetation were prepared and students visited their school and conducted a site analysis. Students then collaborated, they brain stormed, and they researched environmental best management practices that may be appropriate for school sites. They targeted the following strategies: no-mow areas, wooded area reforestation, rain gardens, environmental theme areas, erosion remediation, landscape enhancement, nature trails, and building shade buffers. Reflecting their school's site analysis, teams began to evolve their land stewardship plans. Unfortunately, due to the short time duration of this project, input from school representatives was not taken into account. However, in April teams will forward their land stewardship plans to their school and follow up opportunities can then be pursued. The schools will also be able to utilize the students' work with other materials necessary for the "Green School" application process.

These maps were compiled using ArcGIS 9.3 software. The GIS data was supplied by the Baltimore County Office of Information Technology in connection with the BCPS project on improving GIS education. Data layers used to construct the maps include topography, hydrology, roads, buildings, vegetation, tax parcel, and

aerial imagery. Students also used Google Earth as an additional aid to study their sites.

Jornada Basin (Ken Ramsey):

The Jornada has several GIS projects currently under development to support site and network integration efforts, including the LTER NIS and EcoTrends project. All long-term research datasets are being processed to relate research data tables to linked GIS features (research sites) within the enterprise geodatabase. This process includes adding identifier, key, and x,y coordinate columns to the comma-separated value data files delivered by the Jornada data catalog, geoportal, and the LTER Data Portal.

The Malpai project and Jornada geoportals are being populated and integrated with Drupal-based data catalogs. The Jornada interactive map is being updated and is based on the enterprise geodatabase. Research maps within the enterprise geodatabase (e.g., vegetation, soils, and geomorphology) are being processed to create EML following GeoNIS best practices.

Konza Prairie (Adam Skibbe):

Data management: At Konza we are currently working to update our GIS catalog. Using high resolution GPS units we have been re-collecting existing locations for ongoing LTER research projects with much greater accuracy. In addition to data found on our website, we will be able to offer a much more robust variety of research project locations once this is completed.

In addition we have been working with updating our aerial photography data holdings. To do this we have had several years of historic aerial photographs, dating back to the 1930s, scanned and are currently creating mosaics and rectifying these images.

GIS Research: There are a number research projects at Konza that are either focused on or heavily use GIS technology. One of these, though not specifically LTER funded, is a project which tracks the movement of our bison herd by using GPS collars. The tracked locations, along with knowledge of burn history, NDVI data, etc. are helping us to better understand how bison use the landscape.

North Temperate Lakes (Aaron Stephenson):

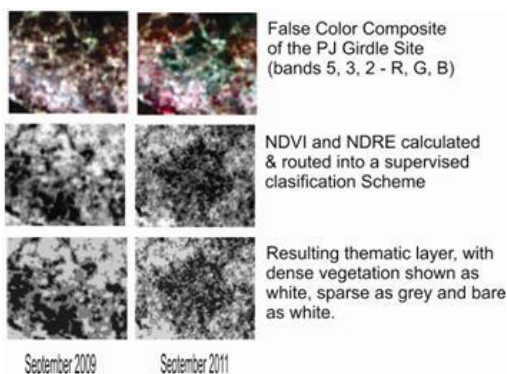
At NTL we are working towards implementing an open source geospatial software stack that will enable the discovery and use of geospatial data. We currently have a stack consisting of PostgreSQL + PostGIS, GeoServer, and Drupal. Drupal allows users to download entire geospatial datasets through a simple HTTP request to GeoServer. We set the default download format to zipped shapefile but GeoServer supports many different formats for the user to choose, something we hope to implement soon. Eventually we intend to have an interactive web mapping application built with the OpenLayers JavaScript API, possibly as a module within Drupal, which will allow the user to explore our geospatial data in a map context and download user-defined spatial extents of datasets. We are also considering building geoprocessing tools for the mapping application, an example of which might be calculating area and volume of a lake at a particular user-selected depth.

Type	Name	Title	Common Formats	All Formats
NTL:nhld_bathymetry_10ft	nhld_bathymetry_10ft		OpenLayers KML GML	Select one
NTL:nhld_bathymetry_5ft	nhld_bathymetry_5ft		OpenLayers KML GML	Select one
NTL:nhld_boundary	nhld_boundary		OpenLayers KML GML	Select one
NTL:nhld_census_2000	nhld_census_2000		OpenLayers KML GML	Select one
NTL:nhld_coarse_woody_debris_docks	nhld_coarse_woody_debris_docks		OpenLayers KML GML	Select one
NTL:nhld_coarse_woody_debris_logs	nhld_coarse_woody_debris_logs		OpenLayers KML GML	Select one
NTL:nhld_coarse_woody_debris_sdi	nhld_coarse_woody_debris_sdi		OpenLayers KML GML	Select one
NTL:nhld_coarse_woody_debris_trees	nhld_coarse_woody_debris_trees		OpenLayers KML GML	Select one
NTL:nhld_landuse_landcover_1930s	nhld_landuse_landcover_1930s		OpenLayers KML GML	Select one

Sevilleta (Mike Friggins):

Currently at SEV we are developing GIS capabilities in terms of providing modern SDE geodatabase driven processing and access for site spatial data and metadata. Perhaps more importantly, we are leveraging remote sensing GIS data to answer questions about ecosystem processes.

SevLTER Co-PI Marcy Litvak, Dan Krofcheck (UNM PhD candidate), Andrew Fox (NEON) and Amy Neuenschwander (UT-Austin) will utilize waveform LiDAR data acquired in 2011 to characterize vegetation structure and estimate above ground biomass within the tower fetches of a network of 8 eddy covariance towers in NM and TX. This represents an ecological gradient from black grama desert grassland on up through spruce-fir coniferous woodlands. The analysis, funded by the NASA ROSES Carbon Cycle program, is intended to reduce uncertainties regarding regional carbon dynamics in the Southwestern US by coupling a more accurate estimate of vegetation structure using full-waveform LiDAR to direct measurements of ecosystem-atmosphere carbon exchange from the towers. The LiDAR and tower data will be incorporated into the Community Land Model (CLM), a land surface model (LSM) using a model-data fusion (MDF) framework to improve regional carbon budgets and predict the response of C dynamics in semi-arid ecosystems to changing climate and disturbance.



Litvak, Krofcheck and colleagues from Idaho State University are also exploring means of measuring large changes in ecosystem structure, remotely, and relating those measurements to the change in function of the ecosystem as measured in-situ using eddy covariance techniques. In this case the analysis is constrained to piñon-juniper woodlands found at a pair of tower sites and the GIS data used is a time series of 42 images for both sites using the RapidEye constellation of satellites and classified using NDVI (Normalized Difference Vegetation Index) as a proxy for LAI and NDRE (Normalized Difference Red Edge) as an approximate for Chlorophyll concentrations calculations. At one of the PJ sites, all adult piñon in the tower fetch were girdled to simulate piñon mortality. The massive change in ecosystem structure following the selective mortality of the piñon overstory in the manipulation site has resulted in some drastic changes in ecosystem function which researchers hope can be detected using vegetation indices from the time series GIS dataset.

Virginia Coastal Reserve (John Porter):

The Virginia Coast Reserve uses GIS coupled with remote sensing to track changes in a "high-speed landscape." This includes time-series of GIS layers, some taken from historical maps, others from recent aerial photos or satellite images. On Hog Island (a primary study site) GIS was used to delineate the boundaries between land deposited during different time intervals to create a chronosequence that is used extensively in vegetation and soil studies. In another recent study, GIS was used to examine the rates of marsh-edge erosion in relation to the surrounding landscape. In 2010 the site had flown its first comprehensive LiDAR survey (previous surveys had been restricted to beach areas). Previous comprehensive elevation data lacked sufficient vertical resolution to be usefully applied on a landscape where vertical changes of 10 cm can represent the difference between grassland and shrubland.

Good Tools And Programs

Interactive Cartographic Almanac

edit

Jamie Hollingsworth (BNZ)

The Interactive Cartographic Almanac (ICA) project was funded by the LNO for the creation of a dynamic mapping tool for network-wide maps specifically to be use in presentations and publications. Jamie Hollingsworth, Bonanza Creek LTER, received funding to develop the map products. These maps enhance our ability to describe and depict the LTER network of sites to funding sources, interested public, and LTER and non-LTER affiliated scientists.

The maps were created in ArcGIS ArcMap software and exported to Adobe Illustrator, for output as a PDF file. An user can open the PDF file in Adobe Acrobat, turn layers on or off, and add or remove labeling and other cartographic symbols. The map layout is North America, with Alaska and Puerto Rico included. Antarctica and MCR sites are provided as insets. The products from this project are customizable high-resolution professional maps that can be exported from Adobe Acrobat as image files for use in other programs, such as Microsoft PowerPoint, and publications. The resources gathered for this atlas will be used to develop base layers for the LTERmapS interactive web application.

The pdf files can be accessed at : http://137.229.80.178/bob_map/.

The Drupal Ecological Information Management System (DEIMS) As A Tool For Many Tasks

edit

Eda Melendez-Colom (LUQ)

Introduction

For more than three years, a group of LTER information managers have been dedicated to the task of developing an information management system with the use of DRUPAL, a content management system (CMS) widely used in the USA, including government agencies (<http://www.drupal.com/>) as well as in Europe (<http://groups.drupal.org/europe>). The Drupal Ecological Information Management System (DEIMS) is the result of this effort (<http://code.google.com/p/deims/>).

When Marshall White and Ignigo San Gil, from the LNO, presented DRUPAL in one of our annual meetings, I knew that this was exactly the kind of system I needed to develop a dynamic and interactive website for Luquillo LTER (LUQ) and to transform the LUQ information management system (IMS) into a complete database driven system. The LUQ IMS had been a conglomerate of different documents developed and maintained with Word, Excel, QPRO and Paradox and distributed among several computers. Today, LUQ has created a DEIMS which has transformed LUQ IMS into a centralized, dynamic, and database driven system. In this article, I will also refer to it as LUQ website-IMS.

Main Features of the LUQ website-IMS

The advantages that DRUPAL, and the LTER version of it, have brought to LUQ have been more than expected. LUQ needed to modernize their website as well as its underlying IMS. Table 1 highlights the main features LUQ IMS needed and how they were implemented in the LUQ DEIMS website-IMS.

Table 1. Initial functionalities for the LUQ DEIMS website-IMS

Functionality Needed	Explanation	Example(s) in LUQ DEIMS website-IMS
Interactivity /allows the search functions	allows the user to interact with the website to search for specific kinds of information	Data sets that can be searched by contact person: http://luq.lternet.edu/data/bycontact People search: http://luq.lternet.edu/people/view Publications search: http://luq.lternet.edu/biblio/year
dynamic displays of content	allows the developer to produce up-to-date lists of content without having to edit more than one table or web page where the information is displayed. The content is entered into a MYSQL database and provides	URL dynamically generated when the user is looking for all publications dated 2012: http://luq.lternet.edu/biblio/year? keys=&lastname_op=or&lastname=All&title=&tid=All&biblio_year_op=%3D&biblio_year[value]=2012&biblio_year[min]=&biblio_year[max]=

	views to design displays of information.	
Dynamic generation of EML Packages	uses a customized module developed by the DEIMS group ^(a) , the site can list and produce new or revised versions of EML packages automatically, as the metadata is updated in the system	EML Package for a germination experiment data set: http://luq.lternet.edu/eml_view/4728
structured IMS	All the website-IMS information is centralized, structured, and stored in tables (content types).	A structured display for LUQ IMS: http://luq.lternet.edu/InfoMng Display of the data and metadata for the germination experiment data set: http://luq.lternet.edu/data/luqmetadata87
data curation ^(b)	The developer displays the contents in many ways using views, modules, and/or scripts to participate in web services such as Metacat and PASTA. ^(c)	EML packages accessible and visible in one of the LUQ DEIMS: http://luq.lternet.edu/data/datacatalog
Easy backups	Ability to backup/mirror the site on different media and transport it to other servers.	Administrative task (not public)

^(a)The DEIMS group is organizing a workshop to develop the second version of this module to enhance functionality and make it compatible with the latest version of DRUPAL (7)

^(b)See <http://crln.acrl.org/content/71/4/194.full> for a reference of the concept of data curation.

^(c)The content types (tables of the metadata database in the DEIMS) and the EML DRUPAL module are down-loadable at:
<http://code.google.com/p/deims/downloads/list>

LUQ opted for a more conservative general design⁽¹⁾ of its website, while using modern features (slide shows) to add a visually attractive, graphic displays. The view of the metadata of a data set is simple and clear to the user⁽²⁾. Figures 1 is a screen shot of the LUQ Home Page and Figure 2 displays the metadata of a dataset.

LUQ DEIMS Home Page Snapshot
(Click on the picture for full screen image)

Figure 1. LUQ DEIMS website-IMS Home Page (<http://luq.lternet.edu/>)

LUQ DEIMS data set display and data file availability Screenshot
(Click on the picture for full screen image)

Figure 2. LUQ DEIMS data set display and data file availability (<http://luq.lternet.edu/data/luqmetadata87>).

DRUPAL has a significant learning curve for the beginner, but the rewards are many. There is a large community of developers that are willing to share and learn from each other. The on-line resources are great, and with patience and perseverance, one can ease the learning process. The LTER DEIMS group was already comfortable with this learning style, and immediately felt at ease communicating with the development community.

Additional Functionalities provided by the DEIMS

As we developed these functionalities in the LUQ DEIMS, we have learned that DRUPAL provides the developer additional features to establish security, make content easily available to users, and visually pleasing design features to modernize the site while making it more dynamic.

One of the most important tasks that LUQ Information Management has accomplished using DRUPAL is teaching data management concepts to the LUQ Schoolyard community. We want them to learn to enter content in a controlled, supervised environment (3). This help us manage the large amount of information being collected at the site and at the same time they experience the need and importance of documenting data.

DRUPAL is an excellent tool for teaching general website and specific LTER concepts. It allows the participant to observe the content they have created as soon as they click the SAVE button. This is accomplished through the use of views and the default entry forms (<http://luq.lternet.edu/IM/OnlineForms>). Users and managers no longer need to edit static HTML pages and special scripts are not necessary to see the final product on the screen.

We conducted three workshops in 2011 and 2012 for the LUQ Schoolyard teachers. The respective main objectives were: understanding the concept and the need for metadata, experience entering metadata information for their research sites into the system, and updating the individual school websites that were created for the participating schools (<http://luq.lternet.edu/outreach/schoolyard>). The workshops were very rewarding, and they allowed us to communicate the importance of information management to the teachers.

Three additional functionalities are being developed into the LUQ DEIMS that will enhance user's capability to *discover* and *connect* different types of information: (1) the use of keywords that characterize data sets, research projects, people, and publications, (2) the development of a DEIMS map that will associate accessible data sets lists to LUQ's main research plots areas, and (3) the incorporation of data into the LUQ DEIMS.

Box 1. Additional DRUPAL features descriptions

- **Use of Taxonomies:** The system allows the user to find and relate information that is not explicitly related by assigning the correct set of keywords to each of these types of information. This is done dynamically with VIEWS, a DRUPAL module that allows the developer to design displays of the website content. Related information is generated on the fly (http://luq.lternet.edu/data/datasets/duration_all)⁽⁴⁾. The final impact of this functionality will be demonstrated when a complete set of keywords to all the LUQ DEIMS is completed. In addition, DRUPAL provides the capability of displaying related information by using user defined taxonomies. LUQ will make these displays public when the assignment of keywords is completed and verified.
- **Mapping research areas and data sets lists:** Druple provides tools to develop an on-line map (http://luq.lternet.edu/luquillo_plots_map) that displays the research areas with their associated list of on-line data sets. The display will serve to integrate LUQ DEIMS with the site GIS data. A site was developed by the local Remote Sensing staff containing all the Spatial data sets (<http://maps.ites.upr.edu/>). The LUQ spatial metadata is under development with the collaboration of some members of the LTER LTERMapS (<http://news.lternet.edu/article320.html>) group and the LNO staff. All spatial metadata will be incorporated into LUQ DEIMS and be searchable with all the site data.
- **Incorporating tables of data:** Finally, DRUPAL 7 provides an enhanced module that allows the incorporation of data tables into the system. This will further enhance data discovery and foster data synthesis.

Closing Statement

DRUPAL may not be the one and only answer for what LUQ needed in order to develop our current dynamic and interactive website-IMS but we are certain that in the current state of this technology, a CMS is the answer to our needs. The DEIMS allows us to have a database driven information management system that serves as a platform to hold all the information and to organize it and display it in a web site the way we want. We are aware that in the future, either other CMS or any other system developed in the Information Technology Domain could outgrow DRUPAL and may provide us with the features we will need then.

References

Moriya, Brian. 2011. C Programming - Pointers. <http://www.exforsys.com/tutorials/c-language/c-pointers.html>.

Footnotes

⁽¹⁾The DEIMS is designed in a modular fashion and the design of tables or content types that hold the metadata are normalized. DRUPAL fields from an outside table are called referrers, also known as pointers (Moriya 2011). This feature enables the dynamic characteristic of DRUPAL Views since updates have only to be done in their original table.

⁽²⁾ We followed the LTER Website guidelines conventions and other network-wide recommendations, including labeling of the Data Menu section, accessibility of data files (no more than 3 clicks), LTER Core data keywords assigned to each data set, and a LUQ key findings section listed in the Research sub menu of the site. The site also uses DEIMS conventions; including using the same fields types and names of the core content types (Persons, Publications, Data Sets, Research Projects). This standardization of the metadata structure is essential for generating dataset EML packages using the customized module developed by the DEIMS group.

⁽³⁾Parts of the website-IMS are still under development and we have restricted access to various parts of the website using users role capabilities provided by DRUPAL. This feature provides added security and allows us to allow updates by selected users where appropriate.

⁽⁴⁾The searches in the actual LUQ DEIMS are of the AND type (all conditions specified by the user in the windows have to be met to produce an output). The DEIMS group have decided to migrate to the new version of DRUPAL (<http://drupal.org/drupal-7.0>). A new version of the EML module will be developed in DRUPAL 7.0. The new version has additional and more powerful capabilities.

Organization and Layout for an LTER Web site using a Content Management System

edit

Eda Melendez-Colom (LUQ)

Introduction

Web content structure: Giving some logical structure to content in a website is a way to facilitate the users' browsing. DRUPAL, an open source Content Management System (CMS), is being used by several LTER sites to manage and organize their websites. DRUPAL stores information in a database (MYSQL). Each page, story, form or any other content (called a *node*) becomes a record or a series of linked records saved in one or more tables in the database. These systems usually assign long, incomprehensible URLs to any content by default. I could not follow the same method I used when organizing the old Luquillo's (LUQ) website in the new design which reflects the website's Menu items.

DRUPAL allows the developer to assign Aliases (customized URLs) to each type of content allowing the designer to give a logical structure to the content in the website. These new URLs do not necessarily reflect the path where the content lies in the file system (like in the DOS environment) but you can certainly use them to give an apparent logical structure to the website.

Web user interfaces: Websites are designed to help data discovery and to download information. Menus, as well as bars, are navigation-type user interfaces. They are one of the elements of the website that reflect its pages design and layout. In DRUPAL, tabs and panels can be also used to display the website content

in addition to menus.

A user can browse through a web site by: (1) following a hierarchical menu that leads them to the wanted information; (2) clicking on links provided by pop-ups that eventually lead them to a wanted product or information, and (3) using a search engine that provides them with windows to enter or select keywords that produce lists of links to the information in the site.

Apparently, high tech organizations like to use hierarchical menus and bars placed at the top or left panels of the page. Many websites' designs use this type of layout.

On the other hand, websites that are designed by or for younger users tend to use more panels and tags, and are often animated. Panels provide concurrent windows displayed in the same screen, each of which links to a specific set of products or information.

Terms: In my effort to organize and document the information in LUQ's website, I encountered the difficulty of trying to use the correct terms to describe the objects, functions or processes. When using "Webopidia"⁽¹⁾ to look for the correct definitions, I realized that terms like "directory" have different meaning to what they used to have in a DOS system. Main Menu, and other more recent ones like "search engine", have different subtle meanings for different communities. Therefore, I inevitably must define beforehand the terms I will be using to avoid some confusion. **Box 1** serves this purpose.

Box 1. Definitions of terms used in this article

- Directory - a Unix/DOS path that leads to a location of a document or another directory i.e., a (sub)directory). (Example: C:\uqweb\data\uqmetadata87)
- File system - a Unix/DOS-like hierarchy of folders or directories
- Search engine - "A program that searches documents for specified keywords and return a list of the documents where the keywords were found".⁽²⁾
- Main Menu - A list of links that is displayed at the top of all pages. The menu items display labels that refer to the type of content contained in the submenu or web page they refer to.

Structure for an LTER Web Site

LTER web sites purpose: An LTER web site's main purpose is to share information with the scientific community and public in general. We also want to foster collaboration and scientific synthesis. These activities are mainly facilitated by the completeness of the metadata in the web site and the way we organize it. Web Design guidelines⁽³⁾ have been developed by the Information Managers Committee (IMC) to facilitate this⁽⁴⁾.

Content Management Systems like DRUPAL provide content types, Views and Themes that allow us to organize content and design web sites that interconnect all types of metadata attributes and information. We organize the metadata using several content types which in turn, are closely related to our EML standards. The navigation and design and layout of our sites should make it easier for the user to find the information they want.

When looking for scientific data, users usually have a specific kind or type in mind. A menu with a structure should help the focused user to find what they want faster. In addition, search engines should allow them to narrow the list of possible links to select from. Navigation is related to structure. It is a way to use menus and DRUPAL's nodes to give DRUPAL websites a virtual classical file system structure. If we give a URL path to our documents in our websites and are consistent with the path we choose, not only will this give some "structure" to the site but it will also be easier to upload and find documents in it.

Methods. Every document's URL path could identify its type by using the structure given to the web site through the Main menu and a submenu item. The menu, submenu and the document's title (node's Title, Subject, Name, etc) should be given a label that describes their content. For example, all data sets in the LUQ website have been given the path /data/luqmetadata#, where # is the catalog number assigned to the data set when it is incorporated into the LUQ IM system. The catalog page that lists all the LUQ data sets has the path /data/datacatalog and the main menu item leading to it is called DATA.

Figure 1 contains the screenshot of LUQ present structure. The DATA menu item is being displayed on the screen; the other 6 menu items are shown as well.

Figure 1. Current LUQ DATA screen and other Main menu. (*Click on image to link to LUQ original image page*)

LUQ Data Page Displaying the LUQ Data Sets Catalog

The website as an archive. It is important for LTER websites to serve the community as a place where people can find old as well as current information.

The challenges presented are that our websites become very dense and designing a layout that allows users to find all the additional documents that cannot be displayed in the website Home Page. The task for maintaining such a website is huge and would need a permanent webmaster hired in each of our LTER sites. Since this is not plausible in most of our sites, we need to find solutions with the resources that we actually have. Box 2 contains suggestions that might help us meet these challenges.

Box 2. Suggestions for designing user-friendly websites

- Have a group of users (preferably students and scientists of the site) help develop useful Menus with meaningful labels
- Develop pages (or sub-sites) pertaining to specific groups from your community so they eventually can edit their content with limited users privileges
- Develop a public page holding the old news from the site that will serve as a historic reference at any time.
- The menu, tags, and panels accompanied by effective search engines can help the user find information they are looking for in a friendly manner.

How: Main menu items with submenus can be used to lead the user to all the principal types of information in a website. In Drupal, other navigational features can be used in addition to the Main menu ("Primary links"). Secondary menu ("Secondary links"), Navigation menu and boxes that are associated with specific paths can be used in different ways to display content in all or selected pages of the website. For example, a box containing user agreements and a disclaimer from the site can be displayed in any page listing data sets. **Figure 1** displays this box in the right panel of the screen.

Search engines in pages displaying specific types of content assist the users to find specific kinds of information. Along with the assignment of meaningful keywords, search engines may help the user discover information that are deeper in the hierarchy of the Menu system or displayed by subsequent pages. **Figure 1** also displays the use of search engines.

Final suggestions: As we create content in the websites we should be aware under which menu item we want our documents to be located and what content type we want them to belong to. The later is a somewhat complex concept in the DRUPAL environment, but when creating content, is better to choose first the menu/submenu you want the new content to have as parents and then choose a good label that shortly describe its content.

We should select a set of menu items, along with their submenu sets, that clearly and completely contain all the information we are holding in the site. News, calendars special global links and Help/FAQ links could appear in panels or blocks that could be set to appear in specific web pages or sections of the web site.

The sites should reflect the consensus that IM workgroups⁽⁵⁾ have developed for the IMC. These agreements should be documented in the work groups' final Terms of References (ToR), reports and guidelines.

The LUQ Drupal Web site has been designed to follow these guidelines to the extent that it has been possible. Constant revision is necessary to make sure that the guidelines are followed.

References and Notes

(1) Webopedia is an online "encyclopedia" of web terms. (<http://www.webopedia.com/>)

(2) Web Site: Web Developers Notes. Article's Title: "Understanding the importance of web site navigation: What is web site navigation?" Year: 2001-2010. Navigation: Web development tips and tricks/Web design tips and tricks. Article's URL: http://www.webdevelopersnotes.com/tips/webdesign/web_site_navigation.php3

(3) Web Site: IMC Web site. Article's Title: "Guidelines for LTER Web Site Design and Content". Year: 2007. Navigation: IM Guide/LTER IM Guidelines/. Article's URL: http://intranet.lternet.edu/im/im_requirements/webdesign_guidelines

(4) The main goal in preparing the "Guidelines for LTER Web Site Design and Content" was to make sure that all sites web sites' content provide their information in an efficient manner such that users could get to the data in no more than three to five clicks. It was intended to foster the appearance of a Network of LTER sites. That is, using the terms defined in this article, we want to make sure that the navigation provided in our web sites allow the user to get to the information they are looking for in no more than five clicks and to find a design and layout that will give the user the awareness of being on an LTER Network's web site. The later has been achieved at the Network level using siteDB as the common framework that allows us to present all US LTER sites basic information using a standard layout. (<http://www.lternet.edu/sites/>)

(5) Workgroups that have worked and/or are currently working with web standard issues and the way the IM Committee make decisions and govern are: Web designers group, Web guidelines group, and Governance Work Group (GWG).

Good Reads

Images and Internet Mapping Websites

edit

There are several impressive images and internet mapping websites that highlight some of the latest technological advances in acquiring, processing, and distributing geo-spatial data. Several are highlighted here:



The Blue Marble: The "Blue Marble" image of Earth snapped by the crew of Apollo 17 in 1972 is one of the most famous photos ever taken. When it appeared, we all suddenly saw the world in a much different way. In the years since, NASA has added other "Blue Marble" photos to its collection, and has used technology to enhance and sharpen the images. Today the space agency unveiled what it's calling the "most amazing high definition image of Earth — Blue Marble 2012." This one was taken "from the VIIRS instrument aboard NASA's most recently launched Earth-observing satellite — Suomi NPP," NASA says, and is a "composite image [that] uses a number of swaths of the Earth's surface taken on January 4, 2012." http://www.nasa.gov/vision/earth/features/bm_gallery_1.html

Wind Map: An invisible, ancient source of energy surrounds us—energy that powered the first explorations of the world, and that may be a key to the future. This map shows you the delicate tracery of wind flowing over the US right now. <http://hint.fm/wind/>



Biked any Good Maps Lately?: Michael Wallace, a Baltimore-based artist, uses his bike to capture GPS locations that he uses to make interesting images. <http://bigthink.com/strange-maps/562-biked-any-good-maps-lately>

Our Far South: A Morgan foundation project aimed at raising New Zealanders' awareness of the area south of Stewart Island. They have a nice interactive map that highlights their expeditions. <http://www.ourfarsouth.org/follow-the-trip.aspx>

Landsat Imagery: Available from Esri, including the Change Matters Viewer. You can do a search and see images from two years, and the NDVI change for the range of years. <http://www.esri.com/landsat-imagery/viewer.html>

FlickrMap: Free web application for exploring geo-tagged Flickr photos inside a map viewer. <http://tamchuen.com/labs/apps/flickrmap/>

NANOOS Visualization System (NVS): NANOOS is the Northwest Association of Networked Ocean Observing Systems. The area of interest is Washington, Oregon, and Northern California. They have a great internet mapping site, with graphs and other imagery. <http://www.nanoos.org/nvs/nvs.php>

How Institutional Factors Influence the Creation of Scientific Metadata

edit

M Mayernik, A Batcheller, and C Borgman, 2011. How Institutional Factors Influence the Creation of Scientific Metadata. iConference 2011 proceedings (<http://dl.acm.org/citation.cfm?id=1940818>), ISBN: 978-1-4503-0121-3.

This good read introduces the concept of institutional frictions relating to creating and managing scientific metadata. The author's compare institutional metadata frictions documented at 3 research networks; the Center for Embedded Networked Sensing, the Earth Sensor Grid, and the Long-Term Ecological Research network. The authors discuss the 3 research networks and demonstrate the types of metadata frictions encountered by each network, which include frictions related to metadata standardization, available time, data sharing, and lack of human resources. The paper has been valuable in framing discussions at the Jornada related to metadata creation and management and its relevance to site and network research activities.

Paper abstract:

Access to high volumes of digital data offer researchers in all disciplines the possibility to ask new kinds of questions using computational methods. Burgeoning digital data collections, however, challenge established data management and analysis methods. Data management is a multi-pronged institutionalized effort, spanning technology, policies, metadata, and everyday data practices. In this paper, we focus on the last two components: metadata and everyday data practices. We demonstrate how "frictions" arise in creating and managing metadata. These include standardization frictions, temporal frictions, data sharing frictions, and frictions related to the availability of human support. Through an illustration of these frictions in case studies of three large, distributed, collaborative science projects, we show how the degree of metadata institutionalization can strongly influence data management needs and practices.

A Comparison of two information management systems

edit

Lotz, T., J. Nieschulze, J. Bendix, M. Dobbermann, and B. König-Ries. 2012. Diverse or Uniform? – Intercomparison of two major German project databases for interdisciplinary collaborative functional biodiversity research. *Ecological Informatics*, 8:10-19. DOI:10.1016/j.ecoinf.2011.11.004

This paper compares the data management systems of two ongoing mid-sized collaborative biodiversity research projects. The two projects share general features which also are present in LTER sites, in that they deal with heterogeneous data contributions from diverse and rotating groups of scientists and students, and have a mission to share data among collaborators and curate for long term preservation. Similar functionality in these two systems has evolved despite independent development with entirely different infrastructures. One system is built on ASP/.NET with an XML database, and the other on open source Apache/Tomcat and MySQL with EML export. Both projects plan additional features for advanced querying and data integration using hierarchical vocabularies and/or ontology. The authors summarize the characteristics of data acquisition/upload, metadata management, database infrastructure, and data presentation in several tables which facilitate easy comparison. Most readers - especially LTER information managers - will feel compelled to compare these two systems both to their local site IM systems and to our Network's.

Calendar

Events: Summer/Fall 2012

edit

Event: Quantifying Uncertainty in Wet Atmospheric Deposition Workshop

Location: HJ Andrews LTER Site, Blue River Oregon

Date: May 21-22, 2012

Website: <http://intranet2.lternet.edu/content/quantifying-uncertainty-wet-atmospheric-deposition>

In this Synthesis Working Group, we propose to develop and apply approaches to estimating uncertainty in precipitation inputs of nutrients for several LTER sites.

Event: Data Acquisition from Remote Locations

Location: Sevilleta Field Station, New Mexico

Date: June 10-15, 2012

Website: <http://sevifs.unm.edu/workshop>

We are pleased to announce our upcoming training workshop, *Data Acquisition from Remote Locations*, sponsored by the Long Term Ecological Research Network Office and the University of New Mexico Sevilleta Field Station. This intensive workshop will be held June 10-15, 2012 at the Sevilleta Field Station in central New Mexico. We will employ a combination of field demonstrations, lectures, hands-on activities, and discussions focusing on three general topics related to the acquisition and handling of environmental data typical at LTER sites: wireless telemetry and networking, photovoltaic power for instrumentation, and connecting field instrumentation to and programming Campbell dataloggers.

Event: Veg-DB: Developing a cross-site system to improve access to vegetation synthetic databases

Location: Harvard Forest LTER, Petersham, MA

Date: June 18, 2012 - June 20, 2012

Website: Veg-DB

Developing a cross-site system to improve access to vegetation synthetic databases.

Event: Society for Conservation GIS Annual Conference

Location: Asilomar Conference Grounds, Monterey, CA

Date: July 19-22, 2012

Website: <http://www.scgis.org/>

The theme for the 2012 SCGIS Conference is "Building Resilience." Topics range from communication and public understanding of science, to how interdisciplinary cooperation can offer solutions to seemingly unyielding problems. Made up of conservationists, geographers, scientists, students, managers, educators, and more, SCGIS is a diverse group of professionals dedicated to using GIS to achieve conservation goals. If you share similar interests and are willing to learn and share, please consider this the conference for you.

Event: Esri International User Conference

Location: San Diego Convention Center, San Diego, California

Date: July 23-27, 2012

Website: <http://www.esri.com/events/user-conference/index.html>

Join with over 12,000 Esri software users to experience the power of "where" in action. Learn how to extend the use of your Esri software to deliver positive returns across your organization.

Event: Ecological Society of America 97th Annual Meeting

Location: Portland, Oregon

Date: August 5-10, 2012

Website: <http://www.esa.org/portland/>

The Theme for the 2012 ESA conference is "Life on Earth: Preserving, Utilizing, and Sustaining our Ecosystems".

Event: Information Managers Meeting

Location: YMCA of the Rockies in Estes Park, Colorado

Date: September 9, 2012

The annual IMC meeting is scheduled for the day prior to the start of the All Scientists Meeting.

Event: 2012 LTER All Scientist Meeting

Location: YMCA of the Rockies in Estes Park, Colorado

Date: September 10th to the 13th, 2012

Website: <http://asm2012.lternet.edu/>

The 2012 All Scientists Meeting (ASM) will once again be held at the YMCA of the Rockies in Estes Park, Colorado from September 10th to the 13th. The Program Committee, made up of a broad range of people representing the whole LTER community, has worked hard to create a meeting that is both focused and open for scientific interactions at many levels.

Plenary presentations will focus on the ASM theme of "The Unique Role of the LTER Network in the Anthropocene: Collaborative Science across Scales". There are plans for over 75 Working Group meetings in seven working group sessions, over 400 posters, four evening mixers, and pre-ASM meetings for information managers, graduate students, education representatives, and the LTER Executive Board. Ample free time is integrated within the program to allow for ad-hoc scientific interaction as well. Logistics for the meeting are handled by the LTER Network Office in collaboration with The Schneider Group, a company specializing in meeting organization.

Event: ForestSAT 2012

Location: Oregon State University, Corvallis, OR

Date: September 11-14, 2012

Website: <http://2012forestsat.com/>

ForestSAT 2012 is the 5th in a series of international conferences promoting scientifically based understanding of how spatial analysis technologies can help describe and monitor forested systems. The purpose of the ForestSAT 2012 conference is to promote scientifically-based understanding of how spatial analysis technologies can help describe and monitor forested systems.

Event: EIMC - SilviLaser 2012

Location: Vancouver, Canada

Date: September 16-19, 2012

Website: <http://silvilaser2012.com/>

SilviLaser 2012 is the twelfth international conference focusing on applications of laser systems for forestry and forest landscapes. Previous conferences have taken place in Canada, Australia, Sweden, Germany, USA, UK, Japan, and Finland. The return to Canada is aimed at bringing together research scientists and practitioners from around the world to share their experience in the development and application of LiDAR for forest and vegetated environments.